Energy Markets

Risk Management, Optimal Liquidation, and Derivatives

Dissertation zur Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für Mathematik und Wirtschaftswissenschaften
an der Universität Ulm

Vorgelegt von
Dipl.-Math. oec., MSc., Kevin Metka
im Januar 2011
Amtierender Dekan: Prof. Dr. Werner Kratz

1. Gutachter: Prof. Dr. Rüdiger Kiesel
2. Gutachter: Prof. Dr. Ulrich Rieder

Tag der Promotion: 24.03.2011
Acknowledgements

Apart from the many contributing individuals, this thesis would not have been possible without the financial support through a scholarship, granted by the EnBW Trading GmbH. In particular, I would like to thank Dr. Sven-Olaf Stoll and Klaus Wiebauer for their scientific support and trust in my work. In addition, they suggested many of the practical problems addressed in this thesis. By answering uncounted questions about energy markets, they contributed as a helpful and valuable source.

I am especially indebted to my Doktorvater, Professor Dr. Rüdiger Kiesel, who has been conducive to this thesis in multiple ways. (1) He initiated the joint work with the EnBW Trading GmbH, establishing the financial support. (2) Technical, as well as conceptional issues were always discussed and resolved. This held the scientific work on track. (3) His confidence in my work kept me motivated and made me adhere to the main objectives, which constitute the contributions of this thesis.

Professor Dr. Ulrich Rieder contributed by sharpening my senses for technical details. This enhanced the thesis considerably.

Further thanks go to the members of the Institute of Mathematical Finance at the University of Ulm, for the many fruitful discussions about and beyond the scope of mathematical finance. In particular, I would like to thank my mate Andreas Rupp for the many refreshing and insightful comments.

A special thank goes to Dr. Reik Börger, who motivated the practical question about retail contracts in Chapter 6 considerably, and to Stefan Ehrenfried, who gave valuable comments on Chapter 5 and 6.

Last but not least, my special thanks go to my mother, Beate Heine, her Husband Jürgen Heine, my brother Danny Metka, and my Sister Linn Bartscht for their infinite patience over the years.
# Contents

## Acknowledgements

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
</tr>
</tbody>
</table>

## 1 Introduction to Energy Markets

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Energy Markets</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Current Issues in Modeling Energy Markets</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Relevant Markets and Portfolio Setup</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Structure of the Thesis and Main Results</td>
<td>10</td>
</tr>
</tbody>
</table>

## 2 A Multivariate Commodity Analysis

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>15</td>
</tr>
<tr>
<td>2.1.1 Related Literature &amp; Contribution</td>
<td>16</td>
</tr>
<tr>
<td>2.1.2 General Setup</td>
<td>18</td>
</tr>
<tr>
<td>2.2 Multivariate Data Analysis</td>
<td>19</td>
</tr>
<tr>
<td>2.2.1 Volatility Patterns in Energy Markets</td>
<td>19</td>
</tr>
<tr>
<td>2.2.2 Volatility Models</td>
<td>24</td>
</tr>
<tr>
<td>2.2.3 Marginal Distributions</td>
<td>34</td>
</tr>
<tr>
<td>2.2.4 Dependence Structure</td>
<td>35</td>
</tr>
<tr>
<td>2.3 Application to Risk Management</td>
<td>36</td>
</tr>
<tr>
<td>2.3.1 Portfolio Setup</td>
<td>37</td>
</tr>
<tr>
<td>2.3.2 Backtesting</td>
<td>39</td>
</tr>
<tr>
<td>2.4 Conclusions &amp; Discussion</td>
<td>40</td>
</tr>
</tbody>
</table>

## 3 Optimal Liquidation Strategies

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Problem Formulation</td>
<td>47</td>
</tr>
<tr>
<td>3.2 Market Dynamics</td>
<td>51</td>
</tr>
<tr>
<td>3.3 Solution Approaches</td>
<td>55</td>
</tr>
<tr>
<td>3.3.1 Static Liquidation</td>
<td>55</td>
</tr>
<tr>
<td>3.3.2 Dynamic Liquidation</td>
<td>56</td>
</tr>
</tbody>
</table>

## 4 Optimal Static Liquidation Strategies

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>57</td>
</tr>
<tr>
<td>4.1.1 Related Literature &amp; Contribution</td>
<td>57</td>
</tr>
<tr>
<td>4.2 Optimal Liquidation as a Convex Optimization Problem</td>
<td>58</td>
</tr>
<tr>
<td>4.2.1 General Structural Properties</td>
<td>58</td>
</tr>
</tbody>
</table>
4.2.2 Conditions for Optimal Solutions of \((P')\) .................................. 63
4.2.3 Conditions for Optimal Solutions of \((P)\) ................................. 66

4.3 Solution Methods and Case Study ............................................. 67
4.3.1 A Cutting Plane Algorithm ............................................... 68
4.3.2 Further Properties of the Management Rule \((\text{PR}, \alpha)\) ........ 68
4.3.3 Negative Market Drift .................................................. 70
4.3.4 Positive Market Drift .................................................. 72
4.3.5 No Market Drift ....................................................... 74
4.3.6 Comparison of Different Models for a Fixed Management Rule . 75
4.3.7 General Remarks ....................................................... 76

4.4 Conclusions & Discussion ................................................... 77

5 Optimal Dynamic Liquidation Strategies .................................. 79
5.1 Introduction ................................................................. 79
5.1.1 Related Literature & Contribution ................................. 79
5.1.2 General Formulation as a Stochastic Dynamic Program .......... 81
5.2 Solving \((P')\): Explicit Optimal Liquidation Strategies ............ 83
5.2.1 The Bellman Equation and Existence of Optimal Solutions .... 83
5.2.2 Frictionless Markets ................................................ 86
5.2.3 Markets with Liquidity Risk ....................................... 93
5.3 Solving \((P)\): Extended Settings and Numerical Schemes ........ 110
5.3.1 The Bellman Equation and Existence of Optimal Solutions under
a Target Wealth .......................................................... 111
5.3.2 An Extended Grid Algorithm ..................................... 115
5.3.3 Case Study .......................................................... 120
5.4 Conclusions & Discussion ................................................ 125
5.5 Extensions: Margin Modeling ............................................ 126

6 Pricing of Structured Retail Contracts .................................... 129
6.1 Introduction ................................................................. 130
6.1.1 Related Literature & Contribution ................................. 130
6.1.2 The Retail Electricity Price ......................................... 131
6.1.3 Retail Contract Structure ......................................... 132
6.1.4 Customer Behavior ................................................ 133
6.2 Valuation Framework for Structured Contracts ..................... 134
6.2.1 Analysis of Payment Flows ......................................... 134
6.2.2 Market Dynamics .................................................. 137
6.2.3 Endowing Payment Flows with Call Rights ...................... 138
6.2.4 Further Properties of \(PV_{t_0}\) ...................................... 144
6.2.5 Illustrative Example ................................................ 147
6.3 Case Study ................................................................. 151
6.3.1 Example Contracts ................................................ 151
6.3.2 Calibration Results ................................................ 153
6.3.3 Analysis of Contracts without Call Rights ...................... 154
Chapter 1

Introduction to Energy Markets

‘There is not the slightest indication that energy will ever be obtainable from the atom.’

Albert Einstein


Energy is a physical quantity that describes the amount of work that can be performed by a force. This definition is commonly attached to the term ‘energy’. From an economic point of view this term is more difficult to specify. Many people understand energy as a class of energy sources, which are comprised of coal, natural gas, crude oil, and many more. These commodities have a long history as a source of energy with a wide range of applicability. In order to organize the distribution of these commodities, markets have emerged in the past centuries.

About 10-15 years ago, many governments initiated the liberalization of their local electricity markets. Electricity is not an energy source, but more an abstract presence and flow of electric charge. Nowadays, it is understood as a type of energy, and thereby constitutes a group together with energy sources. The electricity markets are highly organized and electricity trading has demonstrated its right to exist, at least in Germany. Only recently, a third member joined the group of energy sources and electricity: Carbon dioxide (CO₂), which is a natural byproduct of many combustion processes. The introduction of a carbon market is a typical example of a market, designed by political authorities, to control the CO₂ emissions. This is predominantly an issue in the industry. The idea of a CO₂ reduction has been manifested in the Kyoto protocol, which serves as a political justification for this market scheme. A replacement protocol has not yet been worked out and several attempts to do so failed in the past. The most

---

1 * March 14, 1879 in Ulm, Germany; † April 18, 1955 in Princeton, USA. German Physicist.
recent effort took place in Mexico in December 2010. The Cancún Climate-Change Summit\textsuperscript{2} achieved at least an adherence to the Kyoto protocol, however, on a voluntary basis.

This introductory paragraph described how we summarize energy sources, electricity, and carbon dioxide: The generic term is energy. Sometimes, people also use the term commodities. Now, what is so special about energy markets?

The energy markets are naturally interconnected. Thus, firms acting in these markets need to be aware of the uncertainty reflected in, e.g., prices, demand, exchange rates, etc. This exposure to uncertainty can be labeled as risk. Since firms intend to gain or maximize their profit, while at the same time limiting the downside risk, it is of utmost importance to identify and understand the sources of uncertainty. The peculiar properties of commodities along with quite heterogeneously designed energy markets make an analysis indispensable and challenging.

While markets for energy sources are well-examined, see e.g. Geman [61], electricity and emissions markets still exhibit many features that lack a deeper understanding and need therefore further exploration. Apart from the statistical properties, it is important to describe the markets accurately with regard to risk management. This includes the correct quantification of risk, as well as optimal trading strategies, from a portfolio’s perspective. For utility companies it is also important to run their power plants optimally, which includes the consideration of market prices of electricity, fuels, and CO\textsubscript{2} emissions. This consideration makes clear that acting in energy markets has to be viewed as a multicommodity task. I.e., all the sub-markets have to be modeled and understood simultaneously. Another aspect that comes along with liberalized and highly connected energy markets is the urge to introduce and price new financial products. These range from simple plain vanilla options over complicated structured products, such as swing options, to products for retail customers, e.g. electricity delivery contracts. The peculiarities of the energy markets have to be taken into account when it comes to pricing and hedging of these products.

In the past years, practitioners encourage the discussion about energy markets by introducing practically relevant questions. This issue is progressively reflected in this thesis as the core motivation for each chapter stems from a practical question. Dealing with practical issues puts a decisive task in the pole position: These issues are usually formulated in a heuristic fashion, not necessarily concise. This makes the treatment even more challenging as the problems have to be translated into mathematical frameworks in the first place. A clear and precise problem description is essential before turning to the development of solution methods. In other words, one major contribution in this thesis, which can not be emphasized sufficiently, is the concise and mathematical problem formulation.

\textsuperscript{2}The summit webpage can be assessed via www.global-energy.org/international/cancun-climate-summit
1.2 Current Issues in Modeling Energy Markets

As the first period of the emissions trading system (EU-ETS) has shown, introducing novel market designs does not necessarily yield the desired results in form of the behavior of the market participants. A few researchers (e.g. Carmona et al. [32]) showed that a comprehensive analysis reveals quite some important insight, and thus should be conducted prior to the installation of such schemes. It is not only the uncertainty of political sovereigns when it comes to such market designs. Other issues are also highly delicate. For instance, the electricity grid is still in the hand of a few major utilities, not only in Germany. This gives them an additional market power, because smaller providers depend on their discretion.

The multi-dimensional nature of energy markets raises the need for adequate multivariate models. While in a single-asset case, the focus is on distributional properties and serial dependence concepts, in the multi-asset case an additional factor comes into play. It is the interdependencies of the markets and their importance as they constitute a large portion of uncertainty. Many markets are interconnected and can influence each other. Thus, it is necessary to understand and describe these effects with a view towards risk management and pricing of derivatives. The difficulty of these approaches are ultimately reflected in the very few recent publications. Many researchers still focus on the one-dimensional analysis of typical power plants, such as gas, coal, and hydro, see Fleten et al. [58] and Kiesel et al. [82] for some recent publications. Some of the few multi-dimensional approaches can be found in Stoll & Wiebauer [116] and Andresen et al. [9].

Another issue already addressed above is the pricing of newly designed derivatives. The peculiar properties of energy markets call for more complex payoff structures that have to be priced. This is especially important for hedging purposes. Moreover, the statistical properties have to be modeled appropriately and pricing methods have to be fast and accurate in order to obtain reasonable prices for the derivatives. A typical example are swing options in energy markets, see Kluge [84] and Kiesel et al. [82].

1.3 Relevant Markets and Portfolio Setup

Throughout this thesis we put ourselves into the position of a utility firm. It can be regarded as a typical asset portfolio. This portfolio can be divided into two large components:

1. Power plant parks
2. Customers

Plant parks consist of coal-fired power plants, gas-fired power plants, nuclear plants, and in recent days, more and more renewable energy power plants are being built (wind
CHAPTER 1. INTRODUCTION TO ENERGY MARKETS

parks, hydro plants, and solar). Conventional plants need to be enlivened with oil products. Consequently, the primary sources to produce energy are coal, natural gas, nuclear, and oil. Recently, renewable energy sources encounter a sharp increase in the demand, mainly due to subsidiaries on a government level. Figure 1.1 depicts the energy mix for Europe and Germany, respectively, as of 2009. Since nuclear power and renewable energy sources are not tradable freely, their financial risk is difficult to assess, and therefore not a part of this thesis.

![Figure 1.1: Sources of power generation in Europe (left panel) and in Germany (right panel). Source: BDEW, IEA.](image)

Coal constitutes the largest position as a source for electricity production not only in Germany, but Europe-wide, as well as world-wide. This motivates us to consider a coal-fired power plant as a representative portfolio of a utility firm. Essentially, three commodities are of interest when it comes to the examination of such a power plant: Electricity, coal, and CO\textsubscript{2} emissions. But what constitutes their relationship?

How does a Coal-Fired Power Plant work?

Consider a coal briquet, which is buried somewhere near Richards Bay in South Africa. This is a large hub for coal and supplies the entire world. The briquet is mined and loaded up a so-called collier, which is a vessel for coal. This vessel pilots Rotterdam as it is the main hub for coal in central Europe. The briquet is transshipped to an inland navigation vessel and on its way via the Rhine to a river-near coal-fired power plant in western Germany. Enlivening the plant with oil prepares the combustion process of the coal briquet. This process dissolves away carbon, which reacts chemically with dioxygen and yields CO\textsubscript{2} as the byproduct. The produced heat from the combustion is used to boil water, which vaporizes and the steam drives a turbine. This turbine converts the energy of the raising steam into mechanical energy to drive a generator, which produces electricity. The electricity distribution network is supplied with the generated current. The efficiency of a coal plant determines how much energy (bound in the briquet) is
converted into electricity. We encounter the notion of *power plant efficiency* throughout this thesis.

Let us return to the portfolio view. In Table 1.1, the typical positions of a coal-fired power plant are listed.

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Quantity</th>
<th>Price</th>
<th>One-Day Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseload 2011</td>
<td>10,000,000 MWh</td>
<td>49.92 EUR/MWh</td>
<td>1.3%</td>
</tr>
<tr>
<td>Peakload 2011</td>
<td>5,000,000 MWh</td>
<td>76.50 EUR/MWh</td>
<td>0.9%</td>
</tr>
<tr>
<td>Coal API #2</td>
<td>-1,000,000 t</td>
<td>97.18 USD/MWh</td>
<td>1.9%</td>
</tr>
<tr>
<td>Currencies</td>
<td>-97,180,000 USD</td>
<td>0.69 EUR/USD</td>
<td>0.8%</td>
</tr>
<tr>
<td>CO$_2$ Emissions</td>
<td>-1,000,000 t</td>
<td>15.71 EUR/t</td>
<td>2.4%</td>
</tr>
</tbody>
</table>

Table 1.1: Typical multicommodity portfolio of a utility company, used for illustration. Baseload and peakload positions, coal, and CO$_2$ emissions are 2011 forward contracts. The prices are taken from 09/30/2010 and the volatility is computed historically from the corresponding log-returns data.

The differences among baseload and peakload electricity, coal API #2, and CO$_2$ emissions are described below. The currency position is necessary, since coal is noted in USD and in order to hedge against the currency risk, utilities usually include foreign currencies in their portfolios. The one-day volatility is calculated using historical log-returns data and is listed for a first information. One can already see that emissions prices have a significant impact on the total volatility of a portfolio. Analysts and traders agree that this is partly due to the presence of speculators (foremost investment banks and hedge funds) in emissions markets. Apart from emissions, the other large risk positions are induced by coal and electricity. To simplify matters, the FX positions are omitted when we analyze a portfolio. Thus, the focus is on electricity, coal, and CO$_2$ emissions. We describe these markets and the data sets briefly in the sequel, as it is a central element of this thesis.

### Electricity

In Germany, electricity is freely tradable since 1998. A well-established exchange is located in Leipzig: The *European Energy Exchange (EEX)*. By the end of 2010, a large variety of energy-related products are being traded quite liquidly. This justifies taking market data, which is quoted at the EEX, as a basis for the case studies and numerical examples. The basic electricity products traded are spot contracts (day-ahead) and forward, resp. futures contracts with delivery periods of one month, one quarter, and one year. Furthermore, one distinguishes between baseload and peakload contracts, which cover different hours throughout the day. Base hours are the arithmetic average of hour 1am until hour 24am. Peak hours are the arithmetic average of hour 9am until hour 8pm. For the long-term planning, a utility usually considers a portfolio of a certain
power plant and intends to liquidate the position in the forward markets, as these markets offer the opportunity to realize a revenue at an early stage. As long as the position is still open, the utility is exposed to price uncertainty. We consider in Chapter 4 and 5 a utility that intends to liquidate the portfolio positions in forward contracts with yearly delivery. Chapter 2 analyzes the forward markets with a view towards risk management. For reasons that become clear later, we consider a comprehensive data set comprised of the latest three delivery years, namely 2009, 2010, and 2011. Table 1.2 summarizes the main features of the contracts and the data sets subject to scrutiny.

<table>
<thead>
<tr>
<th>Contract Specifics</th>
<th>2009 Contracts</th>
<th>2010 Contracts</th>
<th>2011 Contracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contract Volume</td>
<td>8784 MWh</td>
<td>8784 MWh</td>
<td>8784 MWh</td>
</tr>
<tr>
<td>Minimum Price</td>
<td>34.00 EUR/MWh</td>
<td>37.90 EUR/MWh</td>
<td>39.55 EUR/MWh</td>
</tr>
<tr>
<td>Maximum Price</td>
<td>90.15 EUR/MWh</td>
<td>89.00 EUR/MWh</td>
<td>89.67 EUR/MWh</td>
</tr>
<tr>
<td>First Delivery Day</td>
<td>01/01/2009</td>
<td>01/01/2010</td>
<td>01/01/2011</td>
</tr>
<tr>
<td>Delivery Period</td>
<td>Year 2009</td>
<td>Year 2010</td>
<td>Year 2011</td>
</tr>
<tr>
<td>Data Set Total</td>
<td>1321 Days</td>
<td>1516 Days</td>
<td>1485 Days</td>
</tr>
<tr>
<td>Data Set Analysis</td>
<td>626 Days</td>
<td>626 Days</td>
<td>626 Days</td>
</tr>
</tbody>
</table>

Table 1.2: Baseload year futures 2009, 2010, and 2011 contract information. Source: EEX.

We denote the time-\(t\) electricity forward price by \(F_{t_0}(t, T_0) := F_t(t, T_0, T_1)\), where \(t < T_0\) and \([T_0, T_1]\) denotes the delivery period. In our case, this period corresponds to one year. We restrain the parameter \(T_1\) for notational convenience.

**Coal**

Coal has a 200 year history as a source of energy and it is still in the focus of attention. The reasons are self-evident: It is cheap to mine, cheap to transport, cheap to store, and most importantly, available in large amounts. In 2007, 26.5% of the world’s energy consumption has been served by coal and experts\(^3\) believe that the importance of coal will rather rise than drop in the coming decades. The world coal market is highly organized. Most of the coal is traded at hubs, e.g. the Amsterdam-Rotterdam-Antwerpen (ARA) region, Richards Bay in South Africa, or Newcastle in Australia. We consider the arithmetic mean of bid/ask forward prices on the Argus McCloskey’s Coal Price Index #2 (API #2), offered by 4 different brokers: Evolution, Spectron, Amerex, and Argus. The API #2 index itself is again an average of the prices of several types of coal all shipped to the ARA area; and from there to Germany via the Rhine. The ARA area is the main hub for imported coal in Europe. The API #2 coal needs to satisfy several

\(^3\)See the website of the World Coal Institute: www.worldcoal.org
1.3. RELEVANT MARKETS AND PORTFOLIO SETUP

criteria, such as a heat value of more than 5800 kcal/kg and a sulfur content of less than one percent. Also note that futures on coal do not have a delivery period as in the case of an electricity future, but rather delivery happens at one distinct day. The contract specifics are summarized in Table 1.3.

<table>
<thead>
<tr>
<th>Contract Specifics</th>
<th>2009 Contracts</th>
<th>2010 Contracts</th>
<th>2011 Contracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contract Volume</td>
<td>1 (Metric) Ton</td>
<td>1 (Metric) Ton</td>
<td>1 (Metric) Ton</td>
</tr>
<tr>
<td>Minimum Price</td>
<td>59.95 USD/t</td>
<td>67.25 USD/t</td>
<td>67.25 USD/t</td>
</tr>
<tr>
<td>Maximum Price</td>
<td>217.38 USD/t</td>
<td>214.00 USD/t</td>
<td>211.00 USD/t</td>
</tr>
<tr>
<td>First Trading Day</td>
<td>01/03/2006</td>
<td>05/02/2006</td>
<td>05/02/2006</td>
</tr>
<tr>
<td>Delivery Day</td>
<td>01/01/2009</td>
<td>01/01/2010</td>
<td>01/01/2011</td>
</tr>
<tr>
<td>Data Set Total</td>
<td>749 Days</td>
<td>920 Days</td>
<td>1145 Days</td>
</tr>
<tr>
<td>Data Set Analysis</td>
<td>626 Days</td>
<td>626 Days</td>
<td>626 Days</td>
</tr>
</tbody>
</table>

Table 1.3: ARA coal API #2 year futures 2009, 2010, and 2011 contract information. Source: EEX. The prices for the 2009 contract are taken from the trader Evolution, since the EEX traded futures price series are not sufficiently long.

Coal is usually traded in USD per ton. In order to eliminate the currency effects on these prices, we convert them into EUR per ton by using the corresponding forward exchange rate for the specific days. We denote the time-\( t \) coal forward price by \( F_2(t, T_0) \), where \( t < T_0 \).

**CO\(_2\) Emissions**

Natural by-products of combustion processes are emissions in form of carbon dioxide (CO\(_2\)). In recent years, several countries have been thinking about this issue as more and more people believe it contributes to global warming. Nevertheless, it might be reasonable to reduce the emission of pollutive gases in any case. One major effort in this direction was the Conference of the Parties to the UNFCCC (COP 3) in Kyoto, which resulted in voluntary agreements of the participating countries, called the Kyoto Protocol. To fulfil the obligations resulting from the ratification of this protocol, the European Union launched the so-called European Climate Change Programme (ECCP) in 2000 and its second version in 2005. Among the many measures taken to restrict the emission of climate relevant gases, the ECCP provides the implementation of CO\(_2\) emission allowances (EUA) trading through a framework called Emission Trading Scheme (EU-ETS). The EU-ETS ensures that every country develops a so-called National Allocation Plan (NAP) each year. This NAP assigns a certain amount of emissions allowances to every facility emitting a considerable amount of CO\(_2\) for free. If such a facility intends to emit more CO\(_2\) than it owns allowances, it needs to buy additional allowances in the market.
The design of the EU-ETS implemented a first period from 2005 to 2007 and a second period from 2008 to 2012, where allowances for the first period cannot be transferred to the second period. The generous NAP’s for the first period led to an oversupply, and therefore to a price collapse from 29 EUR per ton CO$_2$ to 0.04 EUR per ton CO$_2$. For the second period, the allocation plans are designed more strictly. Our data series is part of the second period, i.e. the futures for the years 2009, 2010, and 2011 traded at the EEX. These futures are on 1000 emission allowances, each representing the permission to emit one ton of CO$_2$. The contract specifics are again summarized in Table 1.4.

<table>
<thead>
<tr>
<th>Contract Specifications</th>
<th>2009 Contracts</th>
<th>2010 Contracts</th>
<th>2011 Contracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contract Volume</td>
<td>1000 EUA</td>
<td>1000 EUA</td>
<td>1000 EUA</td>
</tr>
<tr>
<td>Minimum Price</td>
<td>12.72 EUR/EUA</td>
<td>8.39 EUR/EUA</td>
<td>8.86 EUR/EUA</td>
</tr>
<tr>
<td>Maximum Price</td>
<td>30.48 EUR/EUA</td>
<td>31.55 EUR/EUA</td>
<td>32.65 EUR/EUA</td>
</tr>
<tr>
<td>Delivery Day</td>
<td>01/01/2009</td>
<td>01/01/2010</td>
<td>01/01/2011</td>
</tr>
<tr>
<td>Data Set Total</td>
<td>820 Days</td>
<td>1075 Days</td>
<td>1330 Days</td>
</tr>
<tr>
<td>Data Set Analysis</td>
<td>626 Days</td>
<td>626 Days</td>
<td>626 Days</td>
</tr>
</tbody>
</table>

Table 1.4: Second period European carbon futures 2009, 2010, and 2011 contract information. Source: EEX.

We denote the time-$t$ emissions forward price by $F_3(t, T_0)$, where $t < T_0$.

Naturally, not the entire price series of the forward contracts can be used for a statistical analysis due to the following reasons. First, all data sets (i.e. all three commodities) have to have the same length, and preferably all prices should be noted on the same day. Usually, historical data is often sparse and missing values are annoying obstacles. Secondly, in the beginning of the trading period the traded volume is still fairly low (sometimes even zero), which gives the prices not too much significance. Third, utility companies do not trade in forward markets until the very last trading day (i.e. the end of the year). Usually, the hedge concept of firms investing in forwards ends a few months prior to delivery (around August/September). Remaining open positions are then placed optimally in the spot market. Hence, we fix the last trading day by the end of September in the year right before the delivery starts. The data set is then spanned by the interval consisting of 626 days prior to the end of September, which corresponds to two and a half years. This yields 9 data sets consisting of all three delivery years and all three commodities. We use the notion of a forward and futures contract interchangeably as the subtle differences do not play any role in this thesis, see Hull [71] for further discussions.

It is important to understand the distinct nature of a futures market. On the one hand, one can look at the historical price series $t \mapsto F(t, T)$ for $t \in [0, T]$. On the other hand,
one can also look at the so-called forward curve $T \mapsto F(t, T)$. For a fixed time $t$, there are several contracts available with different maturities $T_1 < T_2 < \ldots$. In Chapters 2, 4, and 5 we work with the historical price series of specific contracts.

Let us focus on the specifics of a coal-fired power plant. For our purposes it is important to account for the efficiency of such a plant as it determines (1) the quantity of coal needed to produce a certain amount of electricity and (2) the emissions amount of carbon dioxide ($\text{CO}_2$) produced during the combustion process. Considering such a plant as a portfolio, the efficiency rate ultimately determines the portfolio weights. Different efficiency rates might yield substantially different portfolios with individual statistical properties. For this reason, we compare an efficient coal plant with an inefficient one. Table 1.5 lists the specifics of each power plant type. The conspicuous fact that more mass in form of $\text{CO}_2$ is emitted than in form of coal is fed can be explained by the combustion process as the carbon is dissolved away and reacts chemically with dioxygen.

<table>
<thead>
<tr>
<th>Output / Input / Capacity</th>
<th>Efficient Coal Plant</th>
<th>Inefficient Coal Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output Electricity</td>
<td>1 MWh</td>
<td>1 MWh</td>
</tr>
<tr>
<td>Input Coal ($h_1$)</td>
<td>0.3 t</td>
<td>0.4 t</td>
</tr>
<tr>
<td>Output Emissions ($h_2$)</td>
<td>0.8 t</td>
<td>1.0 t</td>
</tr>
<tr>
<td>Total Capacity $Q_{\text{tot}}$</td>
<td>500 MW</td>
<td>250 MW</td>
</tr>
</tbody>
</table>

Table 1.5: Comparison of an efficient and inefficient coal-fired power plant. Source: EnBW. The total capacity $Q_{\text{tot}}$ is not necessarily the total capacity of the power plant but merely the available capacity for the utility as plant sharing is common in Germany.

From a utility’s standpoint, the so-called *clean dark spread* is essential, i.e. the linear combination

$$S_t := F_1(t, T_0) - h_1 F_2(t, T_0) - h_2 F_3(t, T_0),$$

It determines the revenue from selling the long electricity position of the specific power plant (neglecting fixed and additional costs) and can be seen as a profit & loss function ($\text{P&L}$). $h_1$ and $h_2$ denote the heat rates and are taken from Table 1.5. Assuming prices from $t = 09/30/2010$ for the 2011 contracts, we can calculate the value of the clean dark spread for an (in-)efficient coal plant:

$$S_{\text{eff}}^t = 49.92 \text{ EUR/MWh} - 0.3 \cdot 73.48 \text{ EUR/t} - 0.8 \cdot 15.71 \text{ EUR/t}$$
$$= 15.31 \text{ EUR/MWh}$$

$$S_{\text{ineff}}^t = 49.92 \text{ EUR/MWh} - 0.4 \cdot 73.48 \text{ EUR/t} - 1.0 \cdot 15.71 \text{ EUR/t}$$
$$= 4.82 \text{ EUR/MWh}$$

The difference in the revenue is striking and shows the necessity to consider different portfolio settings. To obtain the revenue from selling in yearly contracts, one has to
multiply the spread by the maximum capacity $Q_{\text{tot}}$ of the power plant (assuming it can be fully liquidated) and again multiply this number by 8760 hours, since one year has 8760 hours. For instance, the efficient coal plant in Table 1.5 constitutes the following portfolio:

- Long position: $500 \text{ MW} \cdot 8760 \text{ h} = 4,380,000 \text{ MWh}$ in yearly electricity contracts.
- Short position: $0.3 \text{ t/MWh} \cdot 4,380,000 \text{ MWh} = 1,314,000 \text{ t}$ in yearly coal contracts.
- Short position: $0.8 \text{ t/MWh} \cdot 4,380,000 \text{ MWh} = 3,504,000 \text{ t}$ in CO$_2$ certificates.

Having introduced the most important notions, we conclude this first introductory chapter by summarizing the main contributions of this thesis.

1.4 Structure of the Thesis and Main Results

This thesis consists of three parts. Part 1 (Chapter 2) examines statistical properties of energy markets and the impact on risk management application. This is daily business for a utility company, and thereby a major research field. We extend recent studies and show how to describe the markets more accurately. Part 2 (Chapters 3-5) raises a question closely linked to the risk management issue from part 1. It examines how a multivariate portfolio can be liquidated optimally under various constraints and market setups. By focusing on the spread, this can be formulated as a single-asset liquidation problem with the advantage that the underlying dynamics can explicitly be described by Gaussian models, since negative spread prices are possible. Part 3 (Chapter 6) takes another party into consideration. There are firms and individuals outside the energy sector, which interact with these markets as well. We mean predominantly the heavy industry and private households, which consume electricity for their daily business/life. We show how to price electricity delivery contracts as this has not yet been addressed in the academic literature.

The core motivation behind all chapters is to carry out the entire modeling process. This includes identifying the problem and putting it down with mathematical tools. Apart from the precise description of the task and the derivation of theoretical results, it is of interest to us to carry out numerical experiments and to interpret the outcomes economically.

**Contribution of Part1/Chapter 2 - A Multivariate Commodity Analysis**

This chapter focuses on a precise statistical description of market prices and on the consequences for risk management application. A utility firm is exposed to uncertain prices and in order to hedge against these fluctuations one needs models that describe the prices accurately. An underestimation of the exposure to uncertainty can have severe consequences for the profit & loss of such a firm.
Many models focus on appropriate distributions combined with a suited dependence concept, such as copulae. However, they neglect the nonstationary nature of the relevant markets. In addition, maturity effects are regarded as a stylized fact present in many commodity markets. It can be misleading to rely on the common notions without scrutinizing the markets in depth. While maturity effects are pronounced in electricity forward contracts with monthly delivery, this effect is barely observable in yearly contracts. Furthermore, each market might be nonstationary, but the resulting portfolio (e.g. the clean dark spread) might become stationary. This is reasonable from a naive market perspective, since the commodities are tied deterministically through the efficiency rates. However, in recent years, more and more speculators are present in these markets driving the spread into different directions. Thus, relying on stationarity can be very misleading. Earlier studies, such as in Börger et al. [24] had only very limited access to data. Now, three years later, we can conduct more comprehensive studies. The advantage is that in 2008 the entire financial world encountered turbulent markets, which gives us very interesting data sets.

It is our scientific contribution to account for the nonstationary nature of recent energy data. We give empirical evidence by examining appropriately prepared historical data sets. The exploratory analysis identifies a speculation bubble (references on this topic are Dema [42], Einloth [51], Jarrow & Protter [72], and Endo & Yamaguchi [53]), encountered by the commodity markets in 2008. The resulting clean dark spread has been distorted and is until today no longer stationary. Recent studies performed well as long as the profit & loss time series was stationary. However, this depends ultimately on the portfolio composition and the specific delivery year. We suggest several time-dependent volatility models and regard characteristics such as heteroscedastic effects in forward markets. We fit the models to the data and examine their statistical performance. The result is that time-dependent volatility models can capture heteroscedasticity well, however, there are also differences among these extended models. The decisive question is, what is the impact on risk management application.

Comparing the different model approaches in a risk management application reveals the dominance of our suggested volatility models. The stationary model (i.e. constant volatility) serves as a benchmark. It can severely fail to predict the so-called value-at-risk adequately. Accounting for time-dependent fluctuations in commodity prices yields the overall insight that models, which include this type of volatility, forecast the value-at-risk conservatively. It appears that the statistical performance of the time-dependent volatility models in terms of describing the heteroscedastic effects well, comes along with a better performance in risk management application. However, there are also significant differences among the extended models. Some tend to show high persistence and others show a too strong reaction to changing markets situations. We suggest to apply these models with caution, however, their right to exist is beyond all question.
CHAPTER 1. INTRODUCTION TO ENERGY MARKETS

Contribution of Part2/Chapter 4 - Optimal Static Liquidation Strategies

Utility companies seek to maximize the revenue from the clean dark spread by liquidating the entire portfolio position over a fixed time period. To our knowledge, this has barely been addressed in the academic literature, not even in other single-asset markets. In addition, the energy markets suffer from illiquidity and other peculiarities that have to be taken into consideration. In a first step, we identify the liquidation problem as a convex optimization problem, derive several structural properties, and prove the existence of solutions. Furthermore, we recognize the analogy to a typical asset allocation problem in the spirit of Markowitz. The decisive difference to our setup is the term structure of the imposed model, which calls for a careful examination. A case study shows by what incentives trivial block strategies can be altered and we quantify benefits and risks induced by the underlying dynamics.

A realistic risk constraint is a target wealth that has to be attained with a certain probability at the end of the liquidation period. We refer to this constraint as a management rule as it often stems from a practical application and is determined in a heuristic fashion. We develop a numerical tool that gives the utility criteria to reasonably fix such target revenues upfront. By imposing two different market models, we examine the resulting liquidation strategies and quantify the revenues using realistic scenarios.

The key insight is that the utility seeks the optimal tradeoff between a maximal expected wealth and the risk not to achieve the target wealth. The drift component induced by the model has a large impact on trading strategies in a risk-neutral setting, however, the tradeoff yields a shifting towards faster selling strategies.

Contribution of Part2/Chapter 5 - Optimal Dynamic Liquidation Strategies

A static liquidation framework does not incorporate changing market conditions as time passes by. This shortcoming can be circumvented by identifying the liquidation task as a Markov Decision Problem (MDP). In a first step we formulate a general stochastic dynamic program. We show that the different market setups fulfill an integrability condition, which justifies deriving the Bellman equation. Furthermore, existence theorems are derived and we find specific model settings that allow for a derivation of explicit analytic formulae for the strategies and value functions. These formulae give detailed and valuable insight into the optimal liquidation behavior.

In a second part, the framework is extended by the target wealth constraint (‘management rule’, mentioned above). The formulation of the Markov Decision Problem from the first part is readily transferable, only slight modifications are necessary. This also holds for the existence of the Bellman equation and of optimal solutions. We demonstrate the difficulty of calculating liquidation strategies explicitly. However, even approximate formulae allow for useful economic interpretations. To gain further conclusions, we develop a numerical algorithm that allows us to examine the liquidation strategies under various
settings. The optimal strategies depend on the state variables and model parameters, which makes the results more difficult to interpret. However, we can still quantify the revenues from trading in the market and draw conclusions about optimal behavior.

In the end, we give an outlook on future work. We are convinced that the concept of dynamic programming to examine optimal liquidation behavior is appropriate. Scrutinizing the risk constraints in more depth, in combination with more realistic models is an auspicious approach for forthcoming research.

Contribution of Part 3/Chapter 6 - Pricing of Structured Retail Electricity Contracts

A classical research field in mathematical finance is the pricing of contingent claims. As already mentioned in the introductory paragraph, especially in energy markets there are many new products being developed, which calls for a sound treatment of these objects. We analyze the German electricity market for retail products and identify structured components within these contracts, such as rights to cancel the agreement from a customers side. The specific design of these contracts combined with the peculiarities of the electricity market makes a detailed analysis necessary. This has not yet been addressed in the academic literature.

It is worth to mention that we encountered practical challenges in collecting all relevant contract specifics. Apart from the publicly advertised features, there are other components important, which are not readily available. We solved these issues by using additional information sources such as telephone hotlines or we decided to hold the framework general enough to include missing pieces of information.

In order to analyze the contracts, we develop a risk-neutral pricing framework, which includes the assumption of an appropriate market model. We calibrate this model to infer market parameters. Furthermore, the contract structure requires the formulation of the framework via the Snell Envelope, which can be seen as the Bellman equation for this problem. Structural properties are derived where needed. The dynamic programming framework cannot be solved explicitly and we suggest to use a numerical scheme in form of the least-squares Monte-Carlo algorithm. This allows us to assign a monetary value to the structures in the contracts. Taking real-world examples, we show whether or not the retail contracts are priced correctly within our pricing framework. In a second step, we show how to price these contracts initially, which is a valuable information for utility companies.

In the end, we suggest how to extend the pricing framework. A major shortcoming is the assumption of a deterministic customer demand. It is well-known that volume risk accounts for a major risk position in a utility’s portfolio. We present an approach how to model the demand reasonably, such that we can still calculate quantities that we are interested in. Furthermore, we draw the parallels to the deterministic demand case to
validate the approach we imposed. We list the next steps of highest priority, which can
be addressed in future work.

**Thesis Structure in a Nutshell**

This thesis has a strong economic motivation. We consider a utility firm as a portfolio and this economic view ties all chapter conceptually. On the other hand, we employ quite different mathematical concepts to come to conclusions. Thus, this mathematical view separates the chapters. These two points of view are essential and we put this thesis in a nutshell as depicted in the following scheme.

<table>
<thead>
<tr>
<th>Starting Point</th>
<th>Power Plants</th>
<th>Customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>Chapter 2</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>Economic View</td>
<td>Exposure to uncertain market prices under a fixed liquidation strategy, effects on risk management</td>
<td>Determination of optimal, static liquidation strategies</td>
</tr>
<tr>
<td>Mathematical View</td>
<td>Statistical model, which incorporates heavy-tailed distributions, time-dependent volatility, and copulae</td>
<td>Convex optimization, using discrete-time models, Karush-Kuhn-Tucker conditions</td>
</tr>
</tbody>
</table>
Chapter 2

A Multivariate Commodity Analysis

2.1 Introduction

Individuals and firms acting in deregulated energy markets are exposed to several sources of uncertainty. We labeled the exposure to uncertainty as risk. Prior to deregulation, risk was widely understood as the failure to deliver electricity to the customer. Thus, the main focus of regulated utilities was to ensure the secure delivery of electricity. The reasons for a failure of delivery were mostly congestion or imbalances in the electricity demand and supply. Major risk sources are comprised of

- Market risk
- Volume risk
- Credit risk
- Operational risk

to name just a few. In this current chapter, we focus solely on the market risk, which are basically comprised of electricity, fuel, and CO$_2$ emissions prices, interest and exchange rates.

A typical utility company can be regarded as a portfolio consisting of long positions in electricity and short positions in fuels and CO$_2$ emissions. In this current chapter we intend to examine the performance of several statistical models with a view towards risk management. We extend recent approaches and show that they describe the behavior of commodity prices not precisely enough. The performance in risk management application is assessed by means of backtesting. The closing of the open portfolio positions takes place in the forward markets, predominantly via yearly delivery contracts. This is the usual approach in practice, since utility firms intend to secure their earnings as early as possible to minimize the risk induced by uncertain prices.
2.1.1 Related Literature & Contribution

Statistical analysis’ and application to energy markets has been studied extensively in recent years. However, while examining each market separately was the focus in publications such as Knittel & Roberts [85] and Kat & Oomen [78], the multicommodity nature of energy markets has been recognized later. One of the first references is Kat & Oomen [79], and Börger et al. [24] conducts a cross-commodity analysis using multivariate models by combining insight about fat-tailed distributions (Barndorff-Nielsen [11], McNeil et al. [98], and Eberlein & Prause [49]) and copulae (Dias [44], Cherubini et al. [36], and Härdle & Okhrin [68]). Dependence among several asset classes is an important ingredient in risk analysis. Since a utility company faces uncertainty in fuel prices, as well as output (electricity) prices, it is crucial to grasp the interdependence of all (or at least, of the most important) commodities that constitute the high-dimensional risk factor. In the past, the Gaussian distribution of financial log-returns was a common assumption in multivariate models. In various markets, especially energy, this is not very often consistent with the historical prices. Copulae allow for an extension of dependency models to non-ellipticity and for separation of margins from the dependency. These novel models with more flexible dependence structures and arbitrary marginal distributions, in contrast to the multivariate Normal model, showed a significant improvement of statistical properties. Börger et al. [24] examined thoroughly the impact on risk management application by conducting a comprehensive backtesting study for several multivariate models. Mainly, the flexible statistical model (i.e. a fat-tailed distribution for the margins coupled by a suitable copula) outperforms more rudimental models, such as the multivariate Normal model.

However, the approaches in [24] neglect the nonstationary nature of energy data (in particular, heteroscedasticity), they exclusively study distributional and dependence properties in a static framework. In addition, maturity effects are regarded as a stylized fact present in many commodity markets. It can be misleading to rely on the common notions without scrutinizing the markets in depth. While maturity effects are pronounced in electricity forward contracts with monthly delivery, this effect is barely observable in yearly contracts. Furthermore, each market under consideration might be nonstationary, the resulting portfolio (e.g. the clean dark spread) might become stationary. This is reasonable from a naive market perspective, since the commodities are tied deterministically through the efficiency rates. However, in recent years, more and more speculators are present in these markets driving the spread into different directions. Thus, relying on stationarity can be very misleading.

It is our scientific contribution to account for the nonstationary nature of recent energy data. We give empirical evidence by examining appropriately prepared historical data sets. The exploratory analysis identifies a speculation bubble (references on this topic are Dema [42], Einloth [51], Jarrow & Protter [72], and Endo & Yamaguchi [53]), encountered by the commodity markets in 2008. The resulting clean dark spread has been distorted and is until today no longer stationary. Recent studies performed well as long
as the profit & loss time series was stationary. However, this depends ultimately on the portfolio composition and the specific delivery year. We suggest several time-dependent volatility models and regard characteristics, such as heteroscedastic effects in forward markets. We fit the models to the data and examine their statistical performance. The result is that time-dependent volatility models can capture heteroscedasticity well, however, there are also differences among these models. This decisive question is, what is the impact on risk management application.

Comparing the different model approaches in a risk management application reveals the dominance of our suggested volatility models. The stationary model (i.e. constant volatility) serves as a benchmark. It can severely fail to predict the so-called value-at-risk adequately. Accounting for time-dependent fluctuations in commodity prices yields the overall insight that models, which include this type of volatility, forecast the value-at-risk conservatively. It appears that the statistical performance of the time-dependent volatility models in terms of describing the heteroscedastic effects well, comes along with a better performance in risk management application. However, there are also significant differences among the extended models. Some tend to show high persistence and others show a too strong reaction to changing markets situations. We suggest to apply these models with caution, however, their right to exist is beyond all question.

Our suggested framework can also be extended by accounting for term structures in the dependencies among several assets (so-called dynamic copula models). This has already been addressed in other markets, see Dias [44], Fantazzini [57], and Serban et. al [113]. We do not impose any of these models for two reasons: (1) From a theoretical point of view, these models are not very well examined. Not much is known about them. Dynamic copula approaches are basically statistical models applied to a data set. However, the effort necessary to formulate a clean framework and the difficulty in applying it to real data becomes evident in Kallsen & Tankov [118] and Tankov [117]. (2) The benefit for a firm is questionable. Static models that capture the distributional properties well, perform already satisfactory when applied to various markets.

The rest of the chapter is organized as follows:

Section 2.1.2 introduces and describes the framework that is used for the statistical analysis of the markets. We establish the connection to existing approaches. The careful discussion of the data sets has already been carried out in Section 1.3. Section 2.2 is the core of this chapter, where we examine the markets under consideration with respect to intertemporal changes. Based on the observations, we suggest several time-dependent volatility models and assess their fitting quality in a case study. In Section 2.3, we show how to apply the suggested model approaches to risk management application and conduct a backtesting procedure to examine their performance. Section 2.4 concludes.
CHAPTER 2. A MULTIVARIATE COMMODITY ANALYSIS

2.1.2 General Setup

The general setup is closely related to forward curve modeling. In a classical Black model (see Black [19]), the dynamics are given by

\[ \frac{dF(t, T_0)}{F(t, T_0)} = \sigma \cdot dW(t), \]

where \( \sigma > 0 \) and \( \{W(t)\}_{t \geq 0} \) is a standard Brownian motion. We consider the corresponding discrete time version and define the logarithmic return series of the three commodities by

\[ t \mapsto r_i(t), \]

\[ r_i(t) := \log \left( \frac{F_i(t, T_0)}{F_i(t-1, T_0)} \right), \quad (2.1) \]

for \( i = 1, 2, 3 \triangleq \{\text{electricity, coal, emissions}\} \) and \( t = 1, \ldots, T = 625 \). \( \log \) denotes the natural logarithm. Let us discuss the recent model approaches. The most basic model setup is the multivariate Gaussian model:

\[ r(t) = \mu + \Sigma \cdot \epsilon(t), \quad (2.2) \]

where \( r(t) = (r_1(t), r_2(t), r_3(t)) \) is the vector of the data set, \( \mu = (\mu_1, \mu_2, \mu_3) \) is the vector of the means, \( \Sigma \) is the constant covariance matrix of dimension \( 3 \times 3 \), and \( \epsilon(t) = (\epsilon_1(t), \epsilon_2(t), \epsilon_3(t))^t \) is the transposed vector of independent standard Gaussian random variables. The calculation of risk numbers such as value-at-risk is straightforward via, e.g. the Delta-Gamma methodology (McNeil et al. [98]). The Gaussian distribution of financial log-returns is a common assumption in multivariate models. It is, however, not very often consistent with historical. Copulae allow for an extension of dependency models to non-ellipticity and for separation of margins from the dependency. In extended studies by, e.g. Börger et al. [24], the authors studied a more flexible approach by separating the margins from the multivariate distribution. The coupling was performed by a suitable copula. More specifically, this model approach can be described by

\[ r_i(t) = \mu_i + \sigma_i \cdot \epsilon_i(t), \quad i = 1, 2, 3 \]

\[ G(\epsilon_1, \epsilon_2, \epsilon_3) = C_{\theta}(G_1(\epsilon_1), G_2(\epsilon_2), G_3(\epsilon_3)). \quad (2.3) \]

\[ (2.4) \]

We dropped the time index of \( \epsilon_i \) for notational convenience. \( \epsilon_i \) are standardized i.i.d. increments with zero mean and variance equal to one. \( \mu_i \) is the constant mean of commodity \( i \) and \( \sigma_i \) is the constant volatility of commodity \( i \). \( C_{\theta} \) is a copula belonging to a parametric family \( \mathcal{C} = \{C_{\theta} : \theta \in \Theta\} \). \( G_i(.) \) denote the marginal distribution functions and they depend on the parameter vector \( \theta \) of the specific marginal distribution. \( G(.) \) denotes the resulting multivariate distribution function. In [24] it was shown how the Normal Inverse Gaussian (NIG) distribution outperforms other elliptic distributions.
2.2 MULTIVARIATE DATA ANALYSIS

Statistically and that the flexibility of the copula approach captures the risk positions more realistically. However, both approaches (2.2) and (2.3)-(2.4) are stationary models applied to, as we see in the sequel, nonstationary data. They simply ignore the nonstationary nature of commodity prices. Thus, the task is to study the impact of an additional time-dependent (and possibly stochastic) volatility function \( \sigma_i(t) \). More specifically,

\[
\begin{align*}
  r_i(t) &= \mu_i + \sigma_i(t) \cdot \varepsilon_i(t), \quad i = 1, 2, 3 \quad (2.5) \\
  G(\varepsilon_1, \varepsilon_2, \varepsilon_3) &= C_\theta(G_1(\varepsilon_1), G_2(\varepsilon_2), G_3(\varepsilon_3)) \quad (2.6)
\end{align*}
\]

\( \mu_i \) is as above. Let \( F_t \) be the available information at time \( t \).

\[
\sigma_i^2(t) = \mathbb{E} \left[ (r_i(t) - \mu_i)^2 \big| F_{t-1} \right]
\]

is the conditional variance given \( F_{t-1} \). For the subsequent analysis we confine the framework to the NIG distribution for the margins and the t-copula for the dependence structure as they performed best in recent studies. This gives us the necessary focus on the impact of different volatility models for the margins on risk management application.

2.2 Multivariate Data Analysis

The analysis in this current section focuses on the data set for the 2011 contracts. Corresponding tables and graphs for the 2009 and 2010 contracts can be found in the Appendix A.2. We asserted above that earlier studies applied stationary models to nonstationary data. In the next section we account for this assertion and give evidence for nonstationarity.

2.2.1 Volatility Patterns in Energy Markets?

Before we focus on detecting volatility patterns across several time periods we examine briefly the stationarity of the log-returns.

Figure 2.1 depicts the autocorrelation function (ACF) for price series of power base compared to the corresponding log-returns. Usually price series exhibit strong autocorrelation and a first step to remove serial dependencies is to consider logarithmic returns. The first difference of the log-price appear slightly correlated, which is also supported by a Ljung-Box test, compare Table 2.1.

The null hypothesis is

\[ H_0: \text{The data is random (i.e. the correlations are zero).} \]

and the test statistic is given by
CHAPTER 2. A MULTIVARIATE COMMODITY ANALYSIS

Figure 2.1: Autocorrelation function (ACF) of price series of power base (left panel) compared to their log-returns (right panel), contract with delivery in 2011.

<table>
<thead>
<tr>
<th>Commodity</th>
<th>p-Value</th>
<th>Statistic</th>
<th>$H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Base</td>
<td>0.06%</td>
<td>46.91</td>
<td>Rejected</td>
</tr>
<tr>
<td>ARA Coal</td>
<td>0.31%</td>
<td>41.59</td>
<td>Rejected</td>
</tr>
<tr>
<td>CO₂ Emissions</td>
<td>0.00%</td>
<td>72.19</td>
<td>Rejected</td>
</tr>
</tbody>
</table>

Table 2.1: Ljung-Box test of the log-returns to show that the ACF of the log-returns are not equal to zero.

\[
Q = T(T + 2) \sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{T-k},
\]

where $T$ is the sample size, $\hat{\rho}_k = \frac{\text{Cov}(r(t),r(t-k))}{\text{Var}[r(t)]}$ is the sample autocorrelation with lag $k$, and $h$ is the number of lags being tested. Under the null hypothesis, $Q \sim \chi^2_{1-\alpha,h}$, where $\alpha$ is the confidence level. Thus, the critical region for rejection is $Q > \chi^2_{1-\alpha,h}$. The rejection of $H_0$ advocates a model with time-dependent mean $\mu_i(t)$ in (2.5). For instance, autoregressive moving average (ARMA) models are popular approaches. However, to keep things simple we focus solely on volatility effects. As argued above, it is not only the nonstationary nature of the margins that are essential to examine but also the stationarity of the profit & loss time series. We introduced two different coal plants with a different degree of efficiency. These constitute two profit & loss (P&L) time series. The efficient P&L time series is referred to as $P&L_{eff}$ and the inefficient P&L time series is referred to as $P&L_{ineff}$.

Figures 2.2, A.1, and A.2 depict the log-returns of the 2011, 2009, and 2010 contracts and the resulting P&L’s $P&L_{eff}$ and $P&L_{ineff}$, respectively.

There is no maturity effect (sometimes referred to as Samuelson effect, see Samuelson [110]) observable in the 2010 and 2011 contracts. Considering the 2009 contract in iso-
2.2. MULTIVARIATE DATA ANALYSIS

Figure 2.2: Log-returns of forward contracts with delivery in 2011 and the resulting P&L’s $P\&L_{\text{eff}}$ and $P\&L_{\text{ineff}}$, 625 trading days: 04/15/2008 - 09/30/2010.

The alleged maturity effect is a precursor of the impending speculation bubble encountered in 2008 (references on this topic are Dema [42], Einloth [51], Jarrow & Protter [72], and Endo & Yamaguchi [53]). It becomes obvious when we compare the log-returns of all three delivery years. The increase in volatility occurs during the same time period (04/2008 - 02/2009) and appears detached from the specific contract. Table 2.2 supports the conjecture of the absence of the maturity effect. We list the annualized volatilities for the first half of the trading period (first 312 trading days) and for the second half (last 313 trading days) of the trading period for the years 2009, 2010, and 2011.

<table>
<thead>
<tr>
<th>Year</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Half</td>
<td>Second Half</td>
<td>First Half</td>
</tr>
<tr>
<td>2009</td>
<td>9.29%</td>
<td>18.96%</td>
<td>12.49%</td>
</tr>
<tr>
<td>2010</td>
<td>17.82%</td>
<td>26.38%</td>
<td>29.17%</td>
</tr>
<tr>
<td>2011</td>
<td>23.41%</td>
<td>15.54%</td>
<td>35.47%</td>
</tr>
</tbody>
</table>

Table 2.2: The volatility across the years and contracts shows very different patterns. This is mainly due to the speculation bubble in 2008.

However, maturity effects are present in contracts with shorter delivery (at least in the case of electricity), see for instance the implied volatility in Table 6.2 in Section 6.3.2. The log-returns of the emissions certificates show the highest volatility compared to the
other two commodities. A common explanation for this phenomena is the abiding presence of speculators, such as investment banks and hedge funds. Considering the time series of $P_{\text{eff}}$ and $P_{\text{ineff}}$ in Figure A.1 explains why earlier studies performed well in risk management application, even though they neglected the nonstationary nature of the marginal log-returns. This resulting clean dark spread was close to being stationary. This is a natural consequence from the deterministic tying of the commodities through the heat rate parameters $h_1$ and $h_2$, respectively (see Table 1.5). In other words, the speculation bubble is also reflected in the P&L time series and the markets are beleaguered by more than just utility firms.

![Figure 2.3: Annualized, rolling volatility with a window of 100 days. All commodities exhibit volatility clustering.](image)

We continue the discussion by calculating the annualized rolling volatility with a window size of 100 trading days. Figure 2.3 depicts the anticipated clustering effects for all three commodities with delivery in 2011. The emissions prices exhibit the highest volatility. The volatility term structure is anything but increasing close to maturity. This nonparametric method shows already that imposing a time-dependent volatility model should be taken into consideration. While a rolling window volatility might give us a qualitative insight into possible volatility term structures, there are well-developed parametric models that provide a more comprehensive statistical toolbox. These are generalized autoregressive conditional heteroscedasticity (GARCH) models, developed by Engle & Bollerslev [21], [54], [56]. In order to determine GARCH effects, we look at the squared log-returns of the commodities. Figure 2.4 shows the squared log-returns for all three commodities. It indicates the presence of GARCH effects in all considered cases.
To prove statistical significance we again employ the Ljung-Box test on the squared log-returns and the squared P&L time series. As expected, the null hypothesis is rejected at all reasonable levels, see Table 2.3.

Table 2.3: Ljung-Box test of the squared log-returns for all commodities and portfolios under consideration. The table lists the p-values with the corresponding test statistics in parentheses.

<table>
<thead>
<tr>
<th>Year</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
<th>P&amp;L_{eff}</th>
<th>P&amp;L_{ineff}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(255.93)</td>
<td>(379.00)</td>
<td>(302.17)</td>
<td>(101.70)</td>
<td>(89.78)</td>
</tr>
<tr>
<td>2010</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(333.28)</td>
<td>(369.63)</td>
<td>(423.05)</td>
<td>(212.80)</td>
<td>(243.15)</td>
</tr>
<tr>
<td>2011</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(165.34)</td>
<td>(254.41)</td>
<td>(483.75)</td>
<td>(286.81)</td>
<td>(402.81)</td>
</tr>
</tbody>
</table>

Even though, tests for stationarity have to be considered with caution, we employ a test that focuses on both, independent and stationary increments. For this sake, Brock, Dechert, and Scheinkman developed such a test in 1987, and therefore it is named BDS test, see Brock et al. [27]. The test statistic is asymptotically normally distributed under the null hypothesis, and as reported in Brock et al. [26], the approximation is
not affected by skewness and heavy tails in the data (as observed in the markets under consideration). The null hypothesis is

\[ H_0: \text{The data is i.i.d.} \]

versus an unspecified alternative and Table 2.4 lists the results.

<table>
<thead>
<tr>
<th>Year</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
<th>P&amp;L_{eff}</th>
<th>P&amp;L_{ineff}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(11.81)</td>
<td>(10.14)</td>
<td>(5.56)</td>
<td>(8.60)</td>
<td>(7.65)</td>
</tr>
<tr>
<td>2010</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(10.64)</td>
<td>(9.76)</td>
<td>(7.95)</td>
<td>(8.02)</td>
<td>(8.79)</td>
</tr>
<tr>
<td>2011</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(7.97)</td>
<td>(8.35)</td>
<td>(8.79)</td>
<td>(8.67)</td>
<td>(10.13)</td>
</tr>
</tbody>
</table>

Table 2.4: BDS test on log-returns \( r_1(t), r_2(t), r_3(t) \), and the portfolios under consideration. The table lists the p-values with the corresponding test statistics in parentheses.

Tests for stationarity are usually not very reliable (e.g. KPSS test), as well as unit root tests (e.g. Dickey-Fuller test). For completeness we note that both reject the null hypothesis for all three commodities and \( P&L_{eff} \) and \( P&L_{ineff} \), respectively. The test statistics in Table 2.4 for the P&L's in 2009 are the lowest which supports again the reasonable results in earlier studies. In Section 2.3 we conduct the risk management analysis on all three delivery years 2009, 2010, and 2011. In turns out that stationary models perform satisfactorily for the 2009 contracts, but fail severely for the 2010 and 2011 contracts, respectively.

The object of the subsequent section is to examine different volatility models that describe the data accurately. To assess the fitting quality of each model, we need to check whether the residual \( \varepsilon_i \) are stationary. We do not address the issue of a proper distribution for the margins, as this has been extensively studied in the past. In fact, we use the insight that NIG distributions are well suited for commodity markets and that the coupling via copulae results in tractable (for risk management purposes) and flexible multivariate models. The key idea is to make the margins stationary by devolatilizing them. The clean dark spread becomes stationary as linear combinations of stationary times series are again stationary, see Taylor [119].

### 2.2.2 Volatility Models

Let us consider the multicommodity model (2.5) introduced above:

\[
r_i(t) = \mu_i + \sigma_i(t) \cdot \varepsilon_i(t),
\]

(2.8)
2.2. MULTIVARIATE DATA ANALYSIS

for \( i = 1, 2, 3 \) \( \triangleq \) \{electricity, coal, emissions\} and \( \sigma_i(t) \) might be any (possibly random) volatility model. The sequence \( \{\varepsilon(t)\}_{t=1}^T \) consists of standardized i.i.d. NIG random variables. The ordinary least-squares estimator of \( \mu_i \) is given by the mean of the log-returns series:

\[
\bar{r}_i := \frac{1}{T} \sum_{t=1}^T r_i(t).
\]

In what follows it suffices to assume \( \sigma_i(t) \varepsilon_i(t) = r_i(t) - \bar{r}_i \). The devolatilized residuals are calculated by

\[
\hat{r}_i(t) := \frac{r_i(t) - \bar{r}_i}{\sigma_i(t)}, \quad (2.9)
\]

for \( i = 1, 2, 3 \) \( \triangleq \) \{electricity, coal, emissions\} and \( t = 1, \ldots, T = 625 \). For the following introduction of several volatility models we omit the subscript \( i \) for notational convenience, however, we always mean all three commodities. Since the innovations in (2.8) are non-Gaussian, the parameter estimation via maximum likelihood becomes basically a quasi-maximum likelihood (QML) procedure. This, however, has no noticeable impact and can be treated as the Gaussian case. See White [123] for further properties of QML estimators. Let \( f(., \zeta) \) be the NIG-density of \( \varepsilon(1) \) and \( \zeta \) the parameter vector of the NIG distribution. Furthermore,

\[
l_t (r(t) - \bar{r}, (r(s))_{s=1,\ldots,t-1}, \vartheta, \zeta),
\]

denotes the log-density of \( r(t) - \bar{r} \), given the observations \( r(1), \ldots, r(t-1) \), and \( \vartheta \) denotes the parameter vector of the specific volatility model \( \sigma(t) \). We obtain for the log-likelihood function, given the data \( r(1), \ldots, r(T) \):

\[
L (r; \vartheta) := \sum_{t=1}^T l_t (r(t) - \bar{r}, (r(s))_{s=1,\ldots,t-1}, \vartheta, \zeta)
= \sum_{t=1}^T \log f \left( \frac{r(t) - \bar{r}}{\sigma(t)}, \zeta \right) - \log \sigma(t).
\]

The parameters are estimated using routines implemented in Matlab. The log-likelihood function is calculated based on Gaussian increments. Consequently, the log-likelihood function becomes

\[
L (r; \vartheta) = \sum_{t=1}^T \log f \left( \frac{r(t) - \bar{r}}{\sigma(t)}, \zeta \right) - \log \sigma(t)
= -T \log \sqrt{2\pi} - \frac{1}{2} \sum_{t=1}^T \frac{(r(t) - \bar{r})^2}{\sigma^2(t)} - \sum_{t=1}^T \log \sigma(t).
\]
The constant $-T \log \sqrt{2\pi}$ can be omitted in the optimization procedure, because it does not affect the location of the optimal values $\hat{\vartheta}$. Thus, we merely consider

$$L(r; \vartheta) = -\frac{1}{2} \sum_{t=1}^{T} \frac{(r(t) - \bar{r})^2}{\sigma^2(t)} - \sum_{t=1}^{T} \log \sigma(t). \quad (2.10)$$

$r$ denotes the vector of all log-returns. After all, (2.10) has to be maximized over the parameter set $\vartheta$ that belongs to the model underlying $\sigma(t)$.

1. Constant Volatility

Constant volatility is the simplest model. It serves as a benchmark for the more advanced approaches. This volatility is estimated by

$$\hat{\sigma}(T) = \frac{1}{T-1} \sum_{t=1}^{T} (r(t) - \bar{r})^2,$$

given the data $r(1), ..., r(T)$, and $\bar{r}$ denotes the mean of the log-return series. The variance of this estimator is given by

$$\text{Var}[\hat{\sigma}(T)] = \frac{1}{T} \left( \mu_4 - \frac{T-3}{T-1} \sigma^4 \right),$$

where $\mu_4$ is the fourth centralized moment and $\sigma^4$ is the quadratic variance of the random variable $\varepsilon(1)$. For the risk management analysis we need to calculate the one-day forecast of the volatility. In other words, given $\hat{\sigma}(t_0)$, $t_0 \leq T$, we need to calculate $\sigma(t_0 + 1)$. In this simple case we set

$$\sigma(t_0 + 1) = \hat{\sigma}(t_0).$$

2. Exponentially Weighted Moving Average (EWMA)

The constant volatility estimator puts equal weights on all past observations, regardless how far back these observations might be. To put more weight on more recent observations, which yields a volatility model that can react more quickly to changing market situations, we propose the following model known as exponentially weighted moving average (EWMA):

$$\hat{\sigma}^2(T) = \frac{(r(T) - \bar{r})^2 + \cdots + \lambda^{T-1} (r(1) - \bar{r})^2}{1 + \cdots + \lambda^{T-1}}$$

$$= \frac{\lambda - 1}{\lambda^T - 1} \sum_{t=0}^{T-1} \lambda^n (r(T-t) - \bar{r})^2$$

$$= \frac{\lambda - 1}{\lambda^T - 1} \sum_{t=1}^{T} \lambda^{T-t} (r(t) - \bar{r})^2, \quad (2.11)$$
2.2. MULTIVARIATE DATA ANALYSIS

Given the data \( r(1), \ldots, r(T) \), and \( 0 < \lambda < 1 \). For \( \lambda \to 1 \) we obtain with l'Hospital

\[
\hat{\sigma}^2(T) = \frac{1}{T} \sum_{t=1}^{T} (r(t) - \bar{r})^2,
\]

i.e. a biased version of the equally weighted case from 1. For any \( t = 1, \ldots, T \), (2.11) can be written recursively:

\[
\hat{\sigma}^2(t) = \frac{\lambda - 1}{\lambda^t - 1} (r(t) - \bar{r})^2 + \frac{\lambda^t - \lambda}{\lambda^t - 1} \hat{\sigma}^2(t - 1)
\]

for \( t \) large \( \approx (1 - \lambda) (r(t) - \bar{r})^2 + \lambda \hat{\sigma}^2(t - 1) \). (2.12)

This is a very convenient formula and we use it for our analysis. The variance of the estimator is given by

\[
\text{Var} [\hat{\sigma}(t)] = 2\hat{\sigma}^4(t) \left( \frac{1 - \lambda}{1 + \lambda} \right),
\]

where \( \hat{\sigma}^4(t) \) is calculated as in 1. Using a specific value for \( \lambda \) yields \( \hat{\sigma}(t) \) and the volatility forecast is calculated by

\[
\sigma(t_0 + 1) = \hat{\sigma}(t_0).
\]

The choice of \( \lambda \) is a delicate issue and well discussed in the literature. Unfortunately, we do not have a proper statistical method to determine an optimal \( \lambda \) in the sense of a maximal likelihood. Risk Metrics\textsuperscript{TM} [1] suggest to use \( \lambda = 0.94 \). In this case, the first (most recent) squared periodic return is weighted by 6%. The next squared return is simply a \( \lambda \)-multiple of the prior weight. In this case 6% multiplied by 94% = 5.64%. And the third prior day’s weight equals 5.30% and so on. This ensures a variance that is weighted or biased toward more recent data.

We use \( \lambda = 0.94 \) in a first step, however, we argue that this number can be chosen more reasonably. The Ljung-Box test tests for weak stationarity and a high p-value advocates a model to be appropriate. Thus, we choose \( \lambda \) such that the p-value of the Ljung-Box test is maximized. While this is quite a heuristic approach, we see later that it leads indeed to better results in risk management application. More specifically, consider the test statistic

\[
Q = T(T + 2) \sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{T - k},
\]

where \( T \) is the sample size, \( \hat{\rho}_k = \frac{\text{Cov}[\hat{r}^2(t), \hat{r}^2(t-k)]}{\text{Var}[\hat{r}^2(t)]]} \) (where \( \hat{r}(t) \) is calculated according to (2.9)) is the sample autocorrelation with lag \( k \), and \( h \) is the number of lags being tested. Under the null hypothesis, \( Q \sim \chi^2_{1 - \alpha, h} \). Thus, the critical region for rejection is \( Q > \chi^2_{1 - \alpha, h} \). The optimized value for \( \lambda \) is a solution of the equation.
CHAPTER 2. A MULTIVARIATE COMMODITY ANALYSIS

\[ \lambda^* = \arg \max_{\lambda \in (0,1]} P \{ Q \geq q | H_0 \}. \]

This optimization procedure is equivalent to minimizing the test statistic \( Q \).

3. ARCH Volatility

The autoregressive conditional heteroscedasticity model (ARCH) has been introduced by Engle [54] in 1982. We consider the simplest case of an ARCH model:

\[ \sigma^2(t) = \alpha_0 + \alpha_1 (r(t - 1) - \bar{r})^2, \]  \hspace{1cm} (2.13)

where \( \alpha_0 \) is a real-valued constant and \( \alpha_1 \in (0, 1) \). \( \bar{r} \) is the mean of the according data series. This model can be estimated with maximum likelihood techniques by plugging (2.13) into (2.10). The estimated parameter set is given by \( \hat{\vartheta} = \{ \hat{\alpha}_0, \hat{\alpha}_1 \} \) and standard deviations are readily available. The unconditional volatility is given by

\[ \bar{\sigma} = \sqrt{\frac{\alpha_0}{1 - \alpha_1}}. \]

The one-day volatility forecast is calculated by

\[ \sigma(t_0 + 1) = \sqrt{\hat{\alpha}_0 + \hat{\alpha}_1 (r(t_0) - \bar{r})^2}. \]

4. GARCH Volatility

Let us consider the simple GARCH(1,1) model:

\[ \sigma^2(t) = \alpha_0 + \alpha_1 (r(t - 1) - \bar{r})^2 + \beta \sigma^2(t - 1), \]  \hspace{1cm} (2.14)

where \( \bar{r} \) is the mean of the according data series. Estimating and forecasting GARCH models has been studies extensively, see Bollerslev [21] and Alexander [3]. GARCH models of higher order did not show any statistical improvement in the case study. Thus, we examine the GARCH(1,1) case. The estimated parameter set is comprised of \( \hat{\vartheta} = \{ \hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta} \} \) and is estimated employing the BHHH algorithm due to Martin et al. [97]. Combining (2.10) and (2.14) yields the log-likelihood function for this model. This function has to be maximized over the set \( \vartheta \), subject to the constraints \( \alpha_0 > 0 \), \( \alpha_1, \beta \geq 0 \) and \( \alpha_1 + \beta < 1 \). The unconditional volatility is given by

\[ \bar{\sigma} = \sqrt{\frac{\alpha_0}{1 - \alpha_1 - \beta}}. \]

We require \( \alpha_0 > 0 \) and \( \alpha_1 + \beta < 1 \) as it ensures the positivity and existence of the unconditional volatility \( \bar{\sigma} \). However, in the estimation procedure we do not impose these constraints, because estimated parameters outside these bounds indicate that the GARCH model is inappropriate for the time series under consideration. Standard deviations for the estimated parameters are again readily available and the one day forecast is given by
\[ \sigma(t_0 + 1) = \sqrt{\hat{\alpha}_0 + \hat{\alpha}_1 (r(t_0) - \bar{r})^2 + \hat{\beta} \hat{\sigma}^2(t_0)}. \]

5. Asymmetric GARCH Volatility

We extend the ordinary GARCH model by introducing a leverage parameter \( \kappa \) that captures asymmetric volatility responses. It was initially suggested by Engle [55] and slightly modified by Glosten, Jagannathan, and Runkle [63] (we refer to this model as GJR-GARCH). The model has the form

\[
\sigma^2(t) = \alpha_0 + (\alpha_1 + \kappa \mathds{1}_{r(t-1) - \bar{r} < 0}) (r(t-1) - \bar{r})^2 + \beta \sigma^2(t-1). \tag{2.15}
\]

Parameter estimation can again be performed by plugging (2.15) into (2.10). The estimated parameter set is augmented by the additional leverage parameter and becomes:

\[ \hat{\vartheta} = \{ \hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}, \hat{\kappa} \}. \]

The parameter \( \kappa \) captures an effect where negative shocks have a greater volatility impact than positive shocks. In equity markets, a positive \( \kappa \) is common, and in contrast negative values for \( \kappa \) are frequently observed in commodity markets. Standard deviations of the estimated parameters are readily available and the unconditional volatility is given by

\[ \hat{\sigma} = \sqrt{\frac{\alpha_0}{1 - \alpha_1 - \beta - \frac{1}{2} \kappa}}. \]

The one-day volatility forecast is calculated by

\[ \sigma(t_0 + 1) = \sqrt{\hat{\alpha}_0 + (\hat{\alpha}_1 + \hat{\kappa} \mathds{1}_{r(t_0) - \bar{r} < 0}) (r(t_0) - \bar{r})^2 + \hat{\beta} \hat{\sigma}^2(t_0)}. \]

Case Study

The six previously introduced volatility models are now assessed with respect to their statistical performance. In other words, how stationary do the data series become after removing the time-dependent volatility? Table 2.5 lists the volatility estimates for model 1 and the values for \( \lambda \) for models 2. and 3. The conditional volatility models 4.-6. provide parameter estimates and the standard deviations are given in parentheses. The results for the years 2009 and 2010 are given in Tables A.1 and A.2, respectively.

As already observed in Figure 2.2, the emissions data series exhibits the highest annualized volatility with about 38%. This does also hold for the other delivery years 2009 and 2010. ‘Optimizing’ the p-value of the Ljung-Box test yields indeed different values for \( \lambda \) in the EWMA model. Figure 2.5 depicts the p-values against the varying values for \( \lambda \). We refer to this model approach as EWMA Optimized.

The GARCH volatility might settle down to the unconditional volatility level \( \hat{\sigma} \), under the assumption that no further market shocks occur. The error parameter \( \hat{\alpha}_1 \) measures
### Volatility Model Parameters

<table>
<thead>
<tr>
<th>Volatility Model</th>
<th>Parameters</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Volatility (p.a.)</td>
<td>$\hat{\sigma}$</td>
<td>20.17%</td>
<td>30.57%</td>
<td>37.81%</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>$\lambda$</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>$\lambda^*$</td>
<td>0.88</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>$\hat{\alpha}_0$</td>
<td>1.47e-004 (5.93e-006)</td>
<td>3.14e-004 (1.63e-005)</td>
<td>4.33e-004 (2.26e-005)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_1$</td>
<td>0.1077</td>
<td>0.1744</td>
<td>0.2388</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.0489</td>
<td>0.0538</td>
<td>0.0564</td>
</tr>
<tr>
<td></td>
<td>$\bar{\sigma}$</td>
<td>1.28%</td>
<td>1.95%</td>
<td>2.39%</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>$\hat{\alpha}_0$</td>
<td>3.89e-006 (1.90e-006)</td>
<td>1.14e-006 (1.13e-006)</td>
<td>1.39e-005 (5.80e-006)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_1$</td>
<td>0.1398</td>
<td>0.0627</td>
<td>0.1017</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.0275</td>
<td>0.0131</td>
<td>0.0205</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_1 + \hat{\beta}$</td>
<td>0.9806</td>
<td>0.9974</td>
<td>0.9759</td>
</tr>
<tr>
<td></td>
<td>$\bar{\sigma}$</td>
<td>1.42%</td>
<td>2.11%</td>
<td>2.40%</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>$\hat{\alpha}_0$</td>
<td>4.11e-006 (1.97e-006)</td>
<td>8.76e-007 (1.04e-006)</td>
<td>1.81e-005 (6.23e-005)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_1$</td>
<td>0.1291</td>
<td>0.0728</td>
<td>0.0506</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.0350</td>
<td>0.0159</td>
<td>0.0260</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_1$</td>
<td>0.8371</td>
<td>0.9386</td>
<td>0.8627</td>
</tr>
<tr>
<td></td>
<td>$\hat{\kappa}$</td>
<td>0.0252</td>
<td>-0.0267</td>
<td>0.0966</td>
</tr>
<tr>
<td></td>
<td>$\hat{\kappa}_1 + \frac{1}{2} \hat{\kappa}$</td>
<td>0.9788</td>
<td>0.9981</td>
<td>0.9616</td>
</tr>
<tr>
<td></td>
<td>$\bar{\sigma}$</td>
<td>1.39%</td>
<td>2.13%</td>
<td>2.17%</td>
</tr>
</tbody>
</table>

Table 2.5: Annualized volatility estimates for model 1 and values for $\lambda$ for models 2.-3. Furthermore, parameter estimates are provided for the volatility models 4.-6. applied to the 2011 contracts. Standard deviations are given in parentheses.
the reaction of the conditional volatility to market shocks. Thus, especially for electricity the volatility is very sensitive to market events as the value is about 0.14. The lag parameter $\beta$ measures the persistence in conditional volatility regardless of what happens in the market. Coal has a $\beta$ of over 0.9, which means that the volatility takes a long time to die out. $\alpha_1 + \beta$ determines the rate of convergence of the conditional volatility to the constant level $\bar{\sigma}$. In case of coal this value is quite close to 1, which means that the term structure of volatility forecasts are relatively flat. Figures 2.6-2.8 depict the sample autocorrelation functions of the squared devolatilized residuals for all commodities and all volatility models. The optimized EWMA model and the GARCH approaches seem to perform best in the sense that they capture the heteroscedastic effects accurately.

In order to assess a quantitative measure we again conduct the Ljung-Box on the squared devolatilized returns $\hat{r}_{1}^{2}$, $\hat{r}_{2}^{2}$, and $\hat{r}_{3}^{2}$, and the BDS test on the devolatilized residuals $\hat{r}_{1}$, $\hat{r}_{2}$, and $\hat{r}_{3}$. Tables 2.6-2.7 summarize the results. Compare also the delivery years 2009, 2010 listed in Tables A.3-A.4 in the Appendix.

The optimized EMWA model and the GARCH models clearly outperform the other approaches. This does also hold for the other delivery years 2009 and 2010. There is no model that performs best for all three commodities. For example, the GARCH(1,1) model yields the highest p-value for the 2011 CO$_2$ emissions contracts, while the optimized EWMA approach yields the highest p-value in case of electricity. In order to keep things comparable and to confine the effort we do not conduct a cross analysis. That means we do not take the best-performing volatility model for each commodity separately. Instead, when establishing a multivariate model we impose the same volatility model for each commodity.

In the next section, we briefly describe the fitting procedure of the marginal distributions. This is merely carried out for completeness.
Figure 2.6: Autocorrelation function of the squared returns of power base, ARA coal, and CO$_2$ emissions for the models (1) constant volatility and (2) EWMA Risk Metrics.

Figure 2.7: Autocorrelation function of the squared returns of power base, ARA coal, and CO$_2$ emissions for the models (3) EWMA Optimized and (4) ARCH.
2.2. **MULTIVARIATE DATA ANALYSIS**

Figure 2.8: Autocorrelation function of the squared returns of power base, ARA coal, and CO$_2$ emissions for the models (5) GARCH and (6) GJR-GARCH.

<table>
<thead>
<tr>
<th>Volatility Model</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO$_2$ Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Volatility</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(165.34)</td>
<td>(254.41)</td>
<td>(483.75)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>38.40%</td>
<td>38.27%</td>
<td>20.17%</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>95.52%</td>
<td>40.20%</td>
<td>36.43%</td>
</tr>
<tr>
<td></td>
<td>(10.63)</td>
<td>(20.92)</td>
<td>(21.57)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(142.82)</td>
<td>(166.70)</td>
<td>(166.12)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>72.87%</td>
<td>81.92%</td>
<td>42.63%</td>
</tr>
<tr>
<td></td>
<td>(15.80)</td>
<td>(14.22)</td>
<td>(20.51)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>80.59%</td>
<td>86.40%</td>
<td>65.03%</td>
</tr>
<tr>
<td></td>
<td>(14.47)</td>
<td>(13.30)</td>
<td>(17.04)</td>
</tr>
</tbody>
</table>

Table 2.6: Ljung-Box test p-values of the squared devolatilized log-returns $\hat{r}^2_1$, $\hat{r}^2_2$, and $\hat{r}^2_3$ for the delivery year 2011. The corresponding test statistics are given in parentheses.
### Table 2.7: BDS test p-values of the devolatilized log-returns $\hat{r}_1, \hat{r}_2,$ and $\hat{r}_3$ for the delivery year 2011. The corresponding test statistics are given in parentheses.

<table>
<thead>
<tr>
<th>Volatility Model</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Volatility</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(7.97)</td>
<td>(8.35)</td>
<td>(8.79)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>7.58%</td>
<td>37.01%</td>
<td>6.75%</td>
</tr>
<tr>
<td></td>
<td>(1.78)</td>
<td>(0.90)</td>
<td>(1.83)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>80.44%</td>
<td>56.29%</td>
<td>41.92%</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.58)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.17%</td>
</tr>
<tr>
<td></td>
<td>(5.81)</td>
<td>(5.45)</td>
<td>(3.14)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>45.84%</td>
<td>35.83%</td>
<td>91.96%</td>
</tr>
<tr>
<td></td>
<td>(-0.74)</td>
<td>(0.92)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>42.74%</td>
<td>27.67%</td>
<td>91.60%</td>
</tr>
<tr>
<td></td>
<td>(-0.79)</td>
<td>(1.09)</td>
<td>(-0.11)</td>
</tr>
</tbody>
</table>

#### 2.2.3 Marginal Distributions

As already mentioned earlier, fat-tailed distributions perform exclusively well in risk management application. Especially, the NIG distribution shows the best fitting quality. For this reason, we fit an NIG distribution to the devolatilized margins $\hat{r}_i(t)$, defined in (2.9). It is well-known that, e.g. GARCH models establish distributions with fatter tails than the Gaussian distribution. One might ask if a Gaussian model with GARCH volatility yields also acceptable results. This might indeed be the case, but requires a much more comprehensive case study. We merely intend to show that stationary models can severely fail in risk management application. Figure 2.9 shows for completeness the fitted NIG distribution to the residuals $\hat{r}_i(t)$ and for comparison a fitted Gaussian distribution.

![Histograms](image-url)

**Figure 2.9:** Histograms of the devolatilized residuals $\hat{r}_i(t)$ of each commodity and fitted NIG distribution (red line) and Gaussian distribution (blue line) for comparison, 2011 contracts.
The previous discussion showed how to treat the heteroscedastic nature of the markets subject to scrutiny. In recent studies, it has been proven useful to model the margins separately and couple them via copulae.

### 2.2.4 Dependence Structure

In Section 2.2.2 we proposed that the log-returns of the examined commodities follow a heavy-tailed distribution endowed with time-varying, possibly stochastic volatility, see (2.8). The innovations in (2.5) have a distribution function described by

\[ G(\varepsilon_1, \varepsilon_2, \varepsilon_3) = C_\theta(G_1(\varepsilon_1), G_2(\varepsilon_2), G_3(\varepsilon_3)), \]

where \( C_\theta \) is a copula belonging to a parametric family \( \mathcal{C} = \{ C_\theta, \theta \in \Theta \} \). We implement a representative from the elliptical family, namely the t-copula. We briefly review on relevant issues about copulae. For generality, we discuss the d-dimensional case.

The term *copula* goes back to Sklar [114]. Standard textbooks are Joe [74] and Nelson [100]. The most general definition is the following one.

**Definition 2.2.1** A copula is a multivariate distribution with standard uniform \((0, 1)\) margins.

Sklar’s theorem, which is the central result, connects copulae functions with distributions such that, on the one hand, every distribution function can be decomposed into its marginal distribution and a copula, and on the other hand, a copula is obtained from decoupling every continuous multivariate distribution function from its marginal distributions.

**Theorem 2.2.2** Let \( G \) be a multivariate, d-dimensional distribution function with margins \( G_1, ..., G_d \). Then there exists a copula \( C \) with

\[ G(x_1, ..., x_d) = C(G_1(x_1), ..., G_d(x_d)), \]  

for \( x_1, ..., x_d \in \mathbb{R} \). If all \( G_i \) are continuous, then \( C \) is unique. Otherwise, \( C \) is uniquely determined on \( \text{Ran} \ G_1 \times \cdots \times \text{Ran} \ G_d \). Conversely, if \( C \) is a copula and \( G_1, ..., G_d \) are univariate distribution functions, then the function \( G \) defined in (2.16) is a multivariate distribution function with margins \( G_1, ..., G_d \).

**Proof:** See Joe [74], p.41.

In other words, we are able to write a multivariate distribution as a function of a copula and several univariate marginal distributions. As a copula does not depend on the univariate marginal distributions, we can say that it captures completely all the information about the dependence among the variables. Formula (2.16) allows us to construct new multivariate distributions by combining various copulae with various...
marginal distributions. For an arbitrary multivariate distribution, we determine its copula from

\[ C(u_1, ..., u_d) = G\left(G_1^{-1}(u_1), ..., G_d^{-1}(u_d)\right), \]  

for \( u_1, ..., u_d \in [0, 1] \). \( G_i^{-1} \) denote the inverse marginal distribution functions. We have to assume that the random variables are continuous as it ensures the existence of the inverse functions.

Copulae can be categorized in several classes. The main classes are simple copulae (independence and perfect dependence), Archimedean copulae, and elliptical copulae. We focus on the latter class as it has outperformed the Archimedean copulae in recent studies, compare Börger et al. [24]. The first representative of an elliptical copula is the Gaussian copula. It represents the dependence structure of the multivariate Gaussian distribution. Coupling non-Gaussian margins (e.g. NIG margins) with a Gaussian copula yields so-called meta-Gaussian distributions. The second representative is the Student’s t-copula. It contains the dependence structure of the multivariate \( t \)-distribution. This copula has one more parameter than the Gaussian copula - \( \nu \), which corresponds to the degrees of freedom from the \( t \)-distribution. For \( \nu \to \infty \) we obtain the Gaussian copula. A low \( \nu \) indicates a strong dependence in the tails of the marginal distributions.

Copula Calibration & Sampling

The estimation of a copula-based multivariate distribution requires both, the estimation of the marginal parameters of \( G_1, ..., G_d \), and the estimation of the copula parameters \( \theta \). We employ the inference for margins method, which is a two-step estimation procedure. In the first step, the marginal parameters are estimated via ML and in the second step one employs a pseudo log-likelihood function to estimate the copula parameter \( \theta \). A detailed discussion about this method can be found in Joe & Xu [75]. In order to sample from the specific copula we use the conditional inverse method, see Embrechts et al. [52] and Devroye [43].

2.3 Application to Risk Management

The portfolio of an utility company consists of many risk positions. We argued in Chapter 1 that it is reasonable to choose a coal-fired power plant to analyze the exposure to uncertain market prices, compare Table 1.1. The choice of a power plant is arbitrary to a certain extent, but represents a quantity that is easy to interpret. Although power plants are exposed to several types of risk, such as operational risk, volume risk, and many more, in the subsequent analysis we cover the main financial risks in the following way. First, we represent the plant by certain financial futures positions. Thus, we neglect all options included in the timing of production, i.e. we assume that we run the plant at electricity baseload times and prices and do not incorporate the possibility of larger earnings when producing the energy at peakload times and prices. Second, risk
metrics such as value-at-risk assume a mark-to-market valuation, i.e. compare prices of the assets today with possible prices of the assets at the end of the time period. This relies on the fact that the owner of the portfolio is able to sell the assets at current market prices. However, the open positions in an energy portfolio can be very large. The market might simply be not able to absorb all positions in one block. In this case, one had to reduce the open positions step by step, such that the portfolio selling is performed over several periods. Taking this into account, one has to think about an optimal closing strategy, which is in the scope of Chapter 4 and 5, respectively.

2.3.1 Portfolio Setup

The profit & loss function of a portfolio determines its VaR. For this reason we need to know the multivariate distribution of this P&L. Let \( \omega = (\omega_1, \omega_2, \omega_3)' \in \mathbb{R}^3 \) denote the portfolio weights (i.e. long and short positions of the commodities) and \( F(t, T_0) = (F_1(t, T_0), F_2(t, T_0), F_3(t, T_0))' \) a non-negative random vector representing the forward prices of the commodities at time \( t \). The value \( V(t) \) of the portfolio at time \( t \) is given by

\[
V(t) = \sum_{i=1}^{3} \omega_i \cdot F_i(t, T_0). \tag{2.18}
\]

The random variable

\[
PL(t) := V(t) - V(t - 1),
\]

\[
= \sum_{i=1}^{3} \omega_i \cdot (F_i(t, T_0) - F_i(t - 1, T_0))
\]

\[
= \sum_{i=1}^{3} \omega_i \cdot F_i(t - 1, T_0) \left( \frac{F_i(t, T_0)}{F_i(t - 1, T_0)} - 1 \right)
\]

\[
= \sum_{i=1}^{3} \omega_i \cdot F_i(t - 1, T_0) \left( e^{ri(t)} - 1 \right), \tag{2.19}
\]

is called the profit & loss function and expresses the change in the portfolio value between two subsequent points in time. The distribution function of \( PL(t) \) is given by the multivariate distribution function \( G_{PL(t)}(x) := \mathbb{P}[PL(t) \leq x], x \in \mathbb{R} \). Let us formalize the notion of the VaR for the P&L.

**Definition 2.3.1** The VaR level \( \alpha \) of a portfolio \( \omega \) is defined as the \( \alpha \)-quantile of \( G_{PL(t)} \):

\[
VaR_\alpha(t) = G_{PL(t)}^{-1}(\alpha). \tag{2.20}
\]
It is clear that the distribution of $PL(t)$ depends on the three-dimensional distribution of the risk factors $r_i(t)$. The weight vector $\omega$ corresponds to the in Table 1.5 introduced (in-)efficient coal plants, i.e.

$$\omega_{\text{eff}} = (1, -0.3, -0.8)$$
$$\omega_{\text{ineff}} = (1, -0.4, -1.0).$$

In order to compute the VaR, we follow these steps:

1. The portfolio consists of power base, ARA coal, and CO$_2$ emissions and we consider the log-returns at time $t_0 \leq T$:

   $$r_1(t_0), r_2(t_0), r_3(t_0).$$

   Estimation of the volatility models described in Section 2.2.2. This yields

   $$\hat{r}_1(t_0), \hat{r}_2(t_0), \hat{r}_3(t_0),$$

   according to (2.9).

2. Simulate $N = 50,000$ Monte-Carlo scenarios for each one of the devolatilized log-returns $\hat{r}_1(t_0 + 1), \hat{r}_2(t_0 + 1), \hat{r}_3(t_0 + 1)$ over the time horizon $(t_0, t_0 + 1)$:

   (a) Sample from the copula passing $\hat{\theta}$ and the parameters $\hat{\zeta}$ from the NIG fit using Sklar’s theorem 2.2.2:

   $$C_{\hat{\theta}}(u_1, u_2, u_3) = G \left( G^{-1}_1 \left( u_1; \hat{\zeta} \right), G^{-1}_2 \left( u_2; \hat{\zeta} \right), G^{-1}_3 \left( u_3; \hat{\zeta} \right) \right).$$

   This yields $N \times 3$ data points, which follow the individual multivariate distribution.

   (b) *Rescale* the filtered residuals by using the fitted volatility models $\sigma_i(t_0 + 1) = \hat{\sigma}_i(t_0)$, i.e.

   $$r_i(t_0 + 1) = \hat{r}_i + \hat{r}_i(t_0 + 1) \cdot \sigma_i(t_0 + 1).$$

3. The P&L is computed for each scenario $j = 1, \ldots, N$ according to (2.19):

   $$PL(t_0) = \sum_{i=1}^{3} w_i \cdot F_i(t_0, T_0) \left( e^{r_i(t_0 + 1)} - 1 \right).$$

   This provides $N = 50,000$ values for $PL(t_0)$.

4. The 95% VaR is the 2500th ordered scenario and is denoted by $\text{VaR}_{0.95}(t_0)$.

To assess the quality of the different volatility models in risk management application one usually applies backtesting procedures. This is discussed in the next section.
2.3. APPLICATION TO RISK MANAGEMENT

2.3.2 Backtesting

Backtesting is a generic term for a procedure to assess the ability of risk models to predict the value-at-risk accurately. Campbell [30] provides a comprehensive summary of backtesting procedures. The volatility analysis in Section 2.2.2 was based on 625 data points. Now we proceed as follows. We estimate the proposed models for a fixed time frame of 250 trading days starting 27 months prior to delivery. Each model provides a forecast of tomorrow’s return distribution given the past data. Consequently, we can calculate the first value-at-risk number \( \text{VaR}_{\alpha}^{(251)} \) for the 251st day. Shifting the window one day into the future and repeating the procedure yields a second value-at-risk number \( \text{VaR}_{\alpha}^{(252)} \). Proceeding in this fashion, we arrive at \( \text{VaR}_{\alpha}^{(625)} \), which is in our cases three months prior to delivery of the corresponding yearly contract. The vector of value-at-risk numbers \( \text{VaR}_{\alpha}^{(251)}, \ldots, \text{VaR}_{\alpha}^{(625)} \) is then compared to the actual \( P&L \) occurred on day \( t = 251, \ldots, 625 \). This yields a hit sequence \( h_{\alpha}(t) \) defined by

\[
\begin{align*}
    h_{\alpha}(t) &= \begin{cases} 
    1, & \text{if } PL(t) \leq \text{VaR}_{\alpha}(t) \\
    0, & \text{if } PL(t) > \text{VaR}_{\alpha}(t).
    \end{cases}
\end{align*}
\] (2.21)

This sequence determines the losses in excess of the value-at-risk. We consider the 95% level of the one-day VaR and in a suitable model framework, the frequency of excessive losses should be close to 5%, i.e.

\[
\frac{1}{T} \sum_{t=251}^{T} h_{0.05}(t) \approx 0.05.
\]

According to Christoffersen [37], the hit sequence \( h_{\alpha}(t) \) should satisfy:

1. **Unconditional coverage**: \( \mathbb{P}[h_{\alpha}(t) = 1] = \alpha \), for all \( t = 251, \ldots, T \). A model inducing a lower probability can be labeled as too conservative whilst a higher probability means a too aggressive model.

2. **Independence**: For all \( i \neq j \), the pairs \( \{h_{\alpha}(t+i), h_{\alpha}(t+j)\} \) must be independent. If value-at-risk violations come in clusters, a good model should account for this by learning from the first violation and yielding a higher value-at-risk in the next step, avoiding further clustering.

Christoffersen [37] suggests tests to examine whether or not the models fulfill 1. and 2. A likelihood ratio test on unconditional coverage has been developed by Kupiec [89]. If we assume \( h_{\alpha}(t) \) to be an i.i.d. sequence, then \( \sum h_{\alpha}(t) \) follows a binomial distribution with parameters \( T - t_0 \) and \( \alpha \). Consequently, the maximum likelihood estimator for \( \alpha \) is given by \( \hat{\alpha} = \frac{1}{T-t_0} \sum_{t=t_0}^{T} h_{\alpha}(t) \). The null hypothesis is given by

\[
H_0 : \mathbb{P}[h_{\alpha}(t) = 1] = \alpha,
\]
and the test statistic is calculated by

$$K = \frac{(1 - \alpha)^n_0 \alpha^{n_1}}{(1 - \hat{\alpha})^{n_0} \hat{\alpha}^{n_1}},$$

with $n_0$ and $n_1$ being the number of zeros and ones in the sequence $h_\alpha(t)$. $-2 \log K$ is asymptotically $\chi^2_1$-distributed under the null hypothesis. To account for independence of violations we test against the alternative of a Markov chain. Consider the hit sequence $h_\alpha(t)$ as a binary first-order Markov chain with transition matrix

$$\Pi_1 = \begin{pmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{pmatrix},$$

with $\pi_{ij} = P[h_\alpha(t) = j|h_\alpha(t-1) = i]$. Independence corresponds to the transition matrix

$$\Pi_2 = \begin{pmatrix} 1 - \pi_2 & \pi_2 \\ 1 - \pi_2 & \pi_2 \end{pmatrix}.$$

$\pi_{01}$, $\pi_{11}$, and $\pi_2$ are estimated by maximum likelihood: $\hat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}}$, $\hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}$, and $\hat{\pi}_2 = \frac{n_{01} + n_{11}}{n_{00} + n_{10} + n_{01} + n_{11}}$. $n_{ij}$ denotes the number of observations $i$ that are followed by $j$. The test statistic is given by

$$C = \frac{(1 - \hat{\pi}_2)^{n_{00} + n_{10} \hat{\pi}_{01}^{n_{01} + n_{11}}}}{(1 - \hat{\pi}_{01})^{n_{00} \hat{\pi}_{01}^{n_{01} + n_{11}}} (1 - \hat{\pi}_{11})^{n_{10} \hat{\pi}_{11}^{n_{01} + n_{11}}}},$$

and $-2 \log C$ is again asymptotically $\chi^2_1$-distributed under the null hypothesis.

Finally, we can combine both tests above by considering the likelihood ratio test statistic $J = K + C$, which is as well asymptotically $\chi^2_2$-distributed under the null hypothesis. We refer to this test as a joint test. The backtesting procedure is performed for the delivery years 2009, 2010, and 2011 and for the volatility models described in Section 2.2.2. Note again that we fixed the specification of the marginal distributions (NIG) and the dependence structure (t-copula). Table 2.8 summarizes the test results and Figures 2.10-2.11 depict the backtesting performance for the different volatility models for the 2011 delivery contracts. The concluding discussion can be found in the next section.

### 2.4 Conclusions & Discussion

The significant improvement of VaR calculations under more flexible models (i.e. fat-tailed distributions and copulae) has been shown in recent studies, see Börger et al. [24]. However, a few authors suggest that the nonstationarity of time series also has to be taken into consideration as it might affect the calculation of risk metrics. Motivated by Eberlein et al. [48] and Härdle & Okhrin [68], we examine several volatility models, which are a reasonable choice for energy markets.
### 2.4. CONCLUSIONS & DISCUSSION

<table>
<thead>
<tr>
<th>Model</th>
<th># Violations</th>
<th>Kupiec LR</th>
<th>Markov</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contracts 2011 $P&amp;L_{eff}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>4</td>
<td>0.00%</td>
<td>76.87%</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.96)</td>
<td>(0.96)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>11</td>
<td>4.73%</td>
<td>31.90%</td>
<td>8.51%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.14)</td>
<td>(0.61)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>20</td>
<td>76.94%</td>
<td>39.02%</td>
<td>66.22%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.96)</td>
<td>(0.69)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>2</td>
<td>0.00%</td>
<td>88.34%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.99)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>15</td>
<td>35.80%</td>
<td>62.26%</td>
<td>58.06%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.66)</td>
<td>(0.89)</td>
<td>(1.54)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>14</td>
<td>23.96%</td>
<td>54.02%</td>
<td>41.51%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.50)</td>
<td>(0.83)</td>
<td>(1.33)</td>
</tr>
<tr>
<td><strong>Contracts 2011 $P&amp;L_{ineff}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>7</td>
<td>0.15%</td>
<td>60.53%</td>
<td>0.56%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td>(0.88)</td>
<td>(0.88)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>12</td>
<td>8.77%</td>
<td>38.76%</td>
<td>16.02%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.23)</td>
<td>(0.69)</td>
<td>(0.92)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>18</td>
<td>85.81%</td>
<td>88.26%</td>
<td>97.35%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.98)</td>
<td>(0.99)</td>
<td>(1.97)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>2</td>
<td>0.00%</td>
<td>88.34%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.99)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>15</td>
<td>35.80%</td>
<td>26.28%</td>
<td>35.01%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.66)</td>
<td>(0.53)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>13</td>
<td>15.02%</td>
<td>33.32%</td>
<td>22.24%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.36)</td>
<td>(0.63)</td>
<td>(0.98)</td>
</tr>
</tbody>
</table>

Table 2.8: Backtesting results and p-values of several tests for VaR adequacy. The test statistics are given in parentheses.
Figure 2.10: Real losses (indicated as bars) for the 2011 contracts compared to the $\text{VaR}_{0.05}(t)$, $t = 251, ..., 625$, efficient power plant.
2.4. CONCLUSIONS & DISCUSSION

Figure 2.11: Real losses (indicated as bars) for the 2011 contracts compared to the $\text{VaR}_{0.05}(t)$, $t = 251, ..., 625$, inefficient power plant.
CHAPTER 2. A MULTIVARIATE COMMODITY ANALYSIS

Having a comprehensive data set at hand provides the opportunity to explore the markets in more depth. The common perception of maturity effects in commodity forward markets is not valid in the case of yearly delivery contracts. Moreover, constant volatility time series models are also not suitable for a precise description of the margins. The reasons become lucid when we compare the P&L’s in Figures 2.2, A.1, and A.2, and the backtesting results in Figures 2.10, 2.11, A.3, A.4, A.5, and A.6. For the 2009 contracts the time series of the P&L has a significant increase in volatility towards the end of the trading period. We identified this effect as the initiation of the 2008 speculation bubble (references on this topic are Dema [42], Einloth [51], Jarrow & Protter [72], and Endo & Yamaguchi [53]). The constant volatility model performs best on this data set, since throughout most of the trading period the P&L time series behaves stationary. Only towards the end one can observe an increase in volatility. However, this does not effect the tails of the portfolio distribution too much, yielding more violations, to be sure, but these are within reasonable limits as confirmed by the Kupiec test. For the 2010 contracts the increase in volatility in the P&L occurs much earlier as is obvious from Figure A.2. The constant volatility model is not able to react fast enough to this heteroscedastic effect and consequently penetrates the P&L fairly often. On the other hand, the 2011 contract provides an opposite result where the constant volatility model severely underestimates the P&L. This is due to the early volatility increase in the P&L time series, which yields a very conservative estimation of $\hat{\sigma}$. The last 375 trading days that are used for the backtesting procedure behave quite stationary and the constant model does not learn fast enough that the volatility is quite low. It persists on a high level for quite some time, which can be observed in Figures 2.10 and 2.11.

The introduced ARCH model performs surprisingly poor for the 2011 contracts with very few violations. This can be explained by the very slow adaption to declining volatility. This is already supported by the bad performance in making the data series stationary, see Tables 2.6 and 2.7. However, in 2009 and 2010 this model seems to perform much better. Especially in Figures A.5 and A.6 one can see how quickly the model adapts to a rapid change in volatility. This provides very high Kupiec p-values.

The EWMA approach yields a strong improvement in the backtesting results for the years 2010 and 2011, respectively. In particular, the optimized EWMA outperforms the naive EWMA in every respect. However, the fairly strong fluctuation for the year 2010 yields too conservative VaR numbers, and hence an overestimation of risk. In Figures 2.10 and 2.11 one can see nicely the tight fit of the VaR against the profit & loss and a perfect forecast for the year 2011.

The GARCH models do not fluctuate as much as the EWMA approaches and yield about the same results in the test statistics. We would favor the simple GARCH(1,1) model as it outperforms the other approaches in an overall view. All model approaches, except the constant volatility model, have in common that they yield too few violations. We obtained 375 P&L values and imposed $\alpha = 5\%$. I.e. 18, respectively 19 violations give a perfect number. A model that provides a lower number of violations is more desirable
2.4. CONCLUSIONS & DISCUSSION

than a model that provides a higher number, from a risk perspective. In other words, the model forecasts worse scenarios than actually occur. After all, time-dependent, possible stochastic volatility has to be taken into account when modeling the risk positions of a utility company.

Let us now establish the connection to the next chapters. The calculation of risk numbers for specific portfolios implicitly requires that the open portfolio positions can be closed within the prescribed time frame. This time frame is usually comprised of ten working days (two weeks). However, energy, and especially electricity markets are severely limited in liquidity. In other words, it is sheer impossible to unwind a large electricity position within a few days. In practice, this can take up to several weeks. This shortcoming is a central assumption of the common VaR. Practitioners introduced the notion of liquidity-adjusted VaR. The first reference for energy markets is Burger et. al [28]. They assume a linear closing over several weeks of the long electricity position and show how an ordinary VaR can be adjusted to account for this closing strategy. In other words, this strategy is predetermined, however, not necessarily optimal (in some sense). Since utility companies usually seek to maximize their profit, one could ask for strategies that adhere to this objective. Naturally, this objective might be subject to risk constraints that reflect the risk aversion/appetite of the utility firm.

In the next three chapters we examine frameworks that are capable of handling exactly this issue. The risk management model developed in this current chapter is too complicated to be implemented for the subsequent analysis. We have to simplify the settings considerably to construct a tractable framework. In other words, the current chapter and the next chapters are connected on an economic basis, however, from a mathematical modeling point of view they are fundamentally different.
Chapter 3

Optimal Liquidation Strategies

3.1 Problem Formulation

The foremost purpose of a utility company is to provide and ensure the security of energy supplies, predominantly gas and electricity. We focus primarily on an electricity provider, and naturally such a provider owns power plants to produce electricity. Apart from generating electricity in own power plants, electricity can be actively traded at energy exchanges or over the counter. Regardless of these sources, the utility, and moreover the government has a strong interest in a secure delivery of electricity. A part of a utility’s portfolio can be labeled as a power plant park. This portfolio contains a natural long position in electricity and since only running power plants earn profit, there is immediately a short position in fuels (coal, gas, oil) and possibly emissions certificates, see also Table 1.1. Such positions are partly settled in the forward and futures market as they allow for an early lock in of profits. Furthermore, if the utility is able to liquidate a large part of its portfolio in forward markets, the exposure to fluctuating demand, predominantly occurring in the spot market, can be minimized and controlled.

In the subsequent chapters, we discuss and analyze how to liquidate a representative portfolio in the forward resp. futures market. That is, we consider merely the settlement in financial contracts, circumventing the peculiarities in the spot market induced by, e.g. physical delivery.

Consider the following setup: A utility company owns a coal-fired power plant and intends to liquidate the total capacity, denoted by $Q_{\text{tot}}$, of this plant. The quantity $Q_{\text{tot}}$ is measured in megawatt (MW) and a typical coal-fired power plant has a capacity of $Q_{\text{tot}} = 500$ MW, compare Table 1.5. This capacity corresponds to 4,380,000 MWh on a yearly basis. The aim is to sell $Q_{\text{tot}}$ in the forward/futures market in yearly contracts (which is not a limitation; contracts with shorter delivery, such as quarterly or monthly are also conceivable). We refer to this period as the delivery period and it is denoted by $[T_0, T_1]$. Furthermore, the liquidation period is comprised of a set of finitely many, equally spaced trading points $t = 0, ..., T$. Thus, we impose a discrete time setting.
and there are \( T + 1 \) points in time where the utility firm, resp. trader can act in the market. In practical application, the liquidation period is usually one year and \( T = 51 \) means that liquidation occurs on a weekly basis. As already introduced in Section 1.3, we denote the financial forward contract for electricity by \( F_1(t, T_0) \), which is measured in EUR per megawatt-hour (MWh), the financial forward contract for coal is denoted by \( F_2(t, T_0) \), which is measured in EUR per metric ton (MT), or short per ton (t), and the financial forward contract for CO\(_2\) emissions is denoted by \( F_3(t, T_0) \), which is measured in EUR per ton CO\(_2\). We assume that whenever a unit of electricity is sold, the corresponding unit (necessary for the production) of coal, as well as CO\(_2\) emissions certificates, is purchased. For the subsequent analysis it is irrelevant whether or not to include CO\(_2\) emissions, however, to deduce realistic model parameters we hold on to the notion of the clean dark spread. Distinguishing between an efficient and an inefficient coal plant spanned a reasonable scope in the risk management analysis conducted in Chapter 2. In the subsequent chapters we stick to an efficient coal power plant, and hence take \( h_1 = 0.3 \) and \( h_2 = 0.8 \) as the efficiency rates, compare Table 1.5. The clean dark spread is now defined by

\[
S_t := F_1(t, T_0) - 0.3 \cdot F_2(t, T_0) - 0.8 \cdot F_3(t, T_0).
\]

This implies that the entire liquidation task can be viewed as a single asset case, where the underlying dynamics are imposed on \( S_t \). Ramifications such as margin modeling are conceivable, but are not discussed here (see Section 5.5 for some comments on future work). The quantity of electricity liquidated at some time \( t \) is denoted by \( \pi_t \) and is measured in MW. This (possibly measurable) function belongs to some set \( \Pi_t \) and if not stated otherwise, we let \( \pi_t \in \mathbb{R} \), for all \( t = 0, ..., T \). Furthermore, we write \( \pi := (\pi_0, ..., \pi_T) \) and \( \Pi := (\Pi_0, ..., \Pi_T) \). The only constraint we always incorporate (and this has to be reflected in \( \Pi \)) is

\[
\sum_{t=0}^{T} \pi_t = Q_{\text{tot}},
\]

i.e. the open electricity position has to be fully liquidated by the end of the liquidation period. The objective is now to maximize the expected terminal wealth (or revenue) from selling electricity in the forward market (and at the same time, the purchasing of coal and emissions certificates). Using the expectation is appropriate in these application for the following reason. Asset allocation strategies (of which portfolio liquidation is a special case) are implemented over hundreds of problem instances, and thereby no single realization has a potential for severe impact on the company wealth. In these situations, the law of large numbers ensures that long-term average revenues are maximized as long as optimal risk-neutral strategies are employed. With the above, heuristic considerations we can state the liquidation problem as follows: Find an optimal liquidation strategy \( \pi^* = (\pi_0^*, ..., \pi_T^*) \) that solves

\[
(P') \quad \max_{\pi \in \Pi} \mathbb{E}_0 \left[ \sum_{t=0}^{T} S_t \pi_t \right],
\]
We write $E_t[.] := E[. | G_t]$ for the conditional expectation operator and $G_t$ is the natural filtration generated by the stochastic processes (introduced in Section 3.2) underlying the liquidation problem. The filtration contains all relevant information accrued up to and including time $t$. The accumulated wealth at time $t$ is denoted by

$$W_t := \sum_{n=0}^{t-1} S_n \pi_n, \quad (3.3)$$

$$W_0 := 0.$$  

We consider several ramifications and constraints of $(P')$. The certainly most severe constraint in the energy markets (especially in the case of electricity) is the illiquidity (in Chapter 5 we elaborate on this notion in more detail). In practice, the utility cannot liquidate the entire position $Q_{\text{tot}}$ at once, because the market is not able to absorb this quantity on short notice. Consequently, $Q_{\text{tot}}$ has to be divided up into several blocks of maximum size $Q_{\text{max}} < Q_{\text{tot}}$, which can be liquidated at a time, i.e. $\pi_t \leq Q_{\text{max}}$, for $t = 0, \ldots, T$. When we require $0 \leq \pi_t \leq Q_{\text{max}}$, for $t = 0, \ldots, T$, the set $\Pi_t$ becomes

$$\Pi_t = \left\{ \pi_t \in \mathbb{R} : \left( Q_t - (T - t) \cdot Q_{\text{max}} \right)^+ \leq \pi_t \leq \min \{ Q_t, Q_{\text{max}} \} \right\}$$

$$= \left[ \left( Q_t - (T - t)Q_{\text{max}} \right)^+, \min \{ Q_t, Q_{\text{max}} \} \right], \quad (3.4)$$

where $Q_t := Q_{\text{tot}} - \sum_{n=0}^{t-1} \pi_n = Q_{t-1} - \pi_{t-1}$ is the remaining quantity to be sold for the remaining $T - t$, e.g. weeks. For any real number $a$, $(a)^+ = \max\{0, a\}$. If $\pi_t$ is unconstrained, i.e. $\pi_t \in \mathbb{R}$, $(3.5)$ becomes

$$\Pi_t = (-\infty, \infty), \quad \text{for } t = 0, \ldots, T - 1$$

$$\Pi_T = Q_T, \quad (3.6)$$

where $Q_T = Q_{\text{tot}} - \sum_{t=0}^{T-1} \pi_t$. This again ensures that the boundary condition $Q_{T+1} = 0$ is met. Another additional condition when one considers profit maximization is a target wealth that has to be reached by the end of the liquidation period. This target wealth is usually fixed prior to the start of trading. We call this target wealth a planning result, denoted by $\text{PR}$. This quantity is measured in EUR or EUR/MWh and serves as a reference to the realized revenue from sales in the electricity market and purchases in the coal market, resp. emissions market. Since the achievement of $\text{PR}$ is uncertain, it is reasonable to provide a minimum probability level and determine optimal liquidation strategies that attain this level. This setup is also discussed in the sequel and we formulate the extended program as follows: Find an optimal liquidation strategy $\pi^* = (\pi^*_0, \ldots, \pi^*_T)$ that solves
The probabilistic constraint in (3.8) is basically a heuristic management rule and states that the accumulated wealth $W_{T+1}$ in (3.3) has to exceed the planning result $\overline{PR}$ with probability at least $1 - \alpha$. In the sequel we denote the management rule by the tuple $(\overline{PR}, \alpha)$. We are ultimately interested in examining properties of the programs $(P')$ and $(P)$, respectively. Chapter 4 treats the static case, meaning that all quantities are calculated based upon the information available at and including time 0. In other words, all quantities are deterministic constants. This approach does not allow for a dynamic adaption to changing market situations. This case is discussed in Chapter 5. Regardless of the approach, we intend to examine optimal behavior and structural properties of the central quantities. In other words, existence and structure of solutions are as interesting and important as methods to solve the maximization problems $(P')$ and $(P)$, respectively. The golden thread of the two subsequent chapters is summarized in Table 3.1.

<table>
<thead>
<tr>
<th></th>
<th>Static Approach (Chapter 4)</th>
<th>Dynamic Approach (Chapter 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(P')$ Existence</td>
<td><em>Lemma 4.2.1</em></td>
<td><em>Theorem 5.2.1 (Sufficient)</em></td>
</tr>
<tr>
<td></td>
<td><em>Corollary 4.2.3 (Sufficient)</em></td>
<td><em>Proposition 5.2.3 (Sufficient)</em></td>
</tr>
<tr>
<td></td>
<td><em>Karush-Kuhn-Tucker 4.2.4 (Sufficient &amp; Necessary)</em></td>
<td><em>Explicit Formulae</em></td>
</tr>
<tr>
<td>Solution Methods</td>
<td><em>Simplex-Algorithm</em></td>
<td><em>[ ]</em></td>
</tr>
<tr>
<td>$(P)$ Existence</td>
<td><em>Lemma 4.2.1</em></td>
<td><em>Theorem 5.3.1 (Sufficient)</em></td>
</tr>
<tr>
<td></td>
<td><em>Corollary 4.2.3 (Sufficient)</em></td>
<td><em>Proposition 5.2.6 (Sufficient &amp; Necessary)</em></td>
</tr>
<tr>
<td>Solution Methods</td>
<td><em>Cutting-Plane-Algorithm</em></td>
<td><em>Grid-Algorithm</em></td>
</tr>
</tbody>
</table>

Table 3.1: Summary of the scheme that we examine in the next two chapters.
There are common features underlying the subsequent discussion. First, we merely focus on the spread, neglecting properties of the individual markets such as different distributions, possible interdependencies or differences in the term structure (e.g. maturity effects). A necessary and certainly interesting extension is the analysis of optimal liquidation when the markets are modeled individually. This also extends naturally the discussion from Chapter 5 by introducing time to the problem. Furthermore, we model the spread with simple one-factor diffusion processes. This also allows for several extensions, which can be incorporated easily into the framework we develop below. However, for the discussion presented here it is sufficient to focus on these simple dynamics. The next section elaborates on possible market dynamics.

3.2 Market Dynamics

Modeling $S_t$ directly neglects the different distributional properties of the margins, as addressed in Chapter 2. Moreover, the dependence structure among these commodities is completely disregarded. One advantage of the direct approach is that the statistical and temporal properties of the spread can be modeled quite accurately. Another advantage is that the direct modeling leads rather to closed-form solutions, which can be useful, e.g. when it comes to option pricing under these models, see Margrabe [95]. An alternative approach models first the margins, which allows for a thorough and precise description of the distributional properties and then the dependence structure among them is taken into consideration. However, in general it is unknown, which marginal models bequeath certain properties that result in desired properties for the spread. For example, it is well-known that in commodity markets the spread exhibits mean-reverting properties. The margins (in futures markets), however, do not necessarily show this type of behavior. The question is which non-mean-reverting dynamics induce mean-reverting dynamics for the spread. Let us introduce some basic models that we use in Chapter 4 and 5, respectively.

The Brownian motion with drift is one of the simplest stochastic diffusion models and is described by the following differential equation:

$$dS_t = \mu dt + \sigma dW_t,$$

(3.9)

where $\{W_t\}_{t \geq 0}$ is a standard Brownian motion, $\mu$ is a real-valued constant, and $\sigma$ is a positive constant. The next lemma summarizes relevant properties.

**Lemma 3.2.1 (Properties of Brownian Motion)**

1. For $0 \leq s < t$, the solution of (3.9) is given by

$$S_t = S_s + \mu(t-s) + \sigma W_{t-s}.$$

(3.10)

2. The conditional distribution is given by

$$S_t | S_s \sim \mathcal{N} \left( S_s + \mu(t-s), \sigma^2(t-s) \right),$$

where $\{W_t\}_{t \geq 0}$ is a standard Brownian motion, $\mu$ is a real-valued constant, and $\sigma$ is a positive constant. The next lemma summarizes relevant properties.
and

\[ \text{Cov}[S_s, S_t] = \sigma^2 s. \]

**Proof:** Karatzas & Shreve [76].

The direction of the drift is determined by the sign of \( \mu \). Moreover, the variance is linearly increasing as time \( t \) goes to infinity. The discrete-time version is given by

\[ S_t = \mu + S_{t-1} + \varepsilon_t, \quad (3.11) \]

with \( \varepsilon_t \) being a Gaussian random variable with zero mean and variance \( \sigma^2 \). Furthermore,

\[ S_t | S_0 \sim \mathcal{N} \left( S_0 + t \mu, t \sigma^2 \right) \quad (3.12) \]

A slight ramification of the simple Brownian motion model is the Bachelier model, see Bachelier [10]. The dynamics are given by

\[ dS_t = \mu S_t dt + \sigma dW_t, \quad (3.13) \]

where \( \{W_t\}_{t \geq 0} \), \( \mu \), and \( \sigma \) are as before. Equation (3.13) is useful for the following reason. Consider the spread as defined in (3.1). Assume, we describe the dynamics of the margins by (omitting emissions certificates for simplicity)

\[
\begin{align*}
    dF_1(t, T_0) &= \mu F_1(t, T_0) dt + \sigma_1 dW_t^{(1)}, \\
    dF_2(t, T_0) &= \mu F_2(t, T_0) dt + \sigma_2 dW_t^{(2)},
\end{align*}
\]

with two Brownian motions \( W_t^{(1)} \) and \( W_t^{(2)} \) with correlation \( \rho \). When we choose

\[ \sigma = \sqrt{\sigma_1^2 - 2 \rho \sigma_1 \sigma_2 + h_1 \sigma_2^2} \]

and

\[ W_t = \frac{h_1 \sigma_2}{\sigma} W_t^{(2)} - \frac{\sigma_1}{\sigma} W_t^{(1)}, \]

we can recover the dynamics for \( S_t \) given in (3.13). Note that \( h_1 \) is the heat rate introduced in (3.1). The disadvantage of this approach is that the marginal distribution are Gaussian, and therefore negative values with positive probability can be attained. To circumvent this problem one usually models the margins by a geometric Brownian motion and approximates the distribution induced by the dynamics in (3.13), see e.g. Carmona [31]. The next lemma summarizes relevant properties of the Bachelier model.
3.2. MARKET DYNAMICS

Lemma 3.2.2 (Properties of the Bachelier Model)

1. For $0 \leq s < t$, the solution of (3.13) is given by
   \[ S_t = e^{\mu(t-s)} S_s + \int_s^t \sigma e^{\mu(t-u)} dW_u. \] (3.14)

2. The conditional distribution is given by
   \[ S_t|_{S_s} \sim \mathcal{N} \left( e^{\mu(t-s)} \frac{\sigma^2}{2\mu} \left( e^{2\mu(t-s)} - 1 \right) \right), \]
   and
   \[ \text{Cov}[S_s, S_t] = -\frac{\sigma^2}{2\mu} \left( e^{-\kappa(t-s)} - e^{-\kappa(t+s)} \right). \]

The next model we consider is more general, and therefore is able to describe more market properties. The dynamics are given by
\[ dS_t = \eta \left( \bar{S} - S_t \right) dt + \sigma dW_t, \] (3.15)
which is an arithmetic Gaussian Ornstein-Uhlenbeck process (see Ornstein & Uhlenbeck [103], Karatzas & Shreve [76], and Benth et al. [13]). The Bachelier model is a special case of the OU process: Set $\bar{S} = 0$ and $\eta = -\mu$. The following lemma summarizes important properties.

Lemma 3.2.3 (Properties of Gaussian OU-Process)

1. For $0 \leq s < t$, the solution of (3.15) is given by
   \[ S_t = \bar{S} + (S_s - \bar{S}) e^{-\eta(t-s)} + \int_s^t \sigma e^{-\eta(t-u)} dW_u. \] (3.16)

2. The conditional distribution is given by
   \[ S_t|_{S_s} \sim \mathcal{N} \left( \bar{S} + (S_s - \bar{S}) e^{-\eta(t-s)} \frac{\sigma^2}{2\eta} \left( 1 - e^{-2\eta(t-s)} \right) \right), \]
   and
   \[ \text{Cov}[S_s, S_t] = \frac{\sigma^2}{2\eta} \left( e^{-\eta(t-s)} - e^{-\eta(t+s)} \right). \]

Proof: See Appendix A.1.

The discrete-time version is given by (see Appendix A.1 for more details)
\[ S_t = (1 - \kappa) \bar{S} + \kappa S_{t-1} + \varepsilon_t, \] (3.17)
and
\( S_{t|S_0} \sim \mathcal{N}\left(\left(1 - \kappa t\right) \bar{S} + \kappa t S_0, \sigma^2 \kappa^2 t^2 - \kappa^2 - 1\right) \) \tag{3.18}

Note that we set \( \sigma := \sigma \epsilon \) and \( \kappa \in (0, 1) \). Interestingly, the monotonicity of the mean is determined by the initial value \( S_0 \). This is obvious from the first derivative w.r.t. \( t \) (assuming the mean is a continuous function in \( t > 0 \)):

\[
\left( S_0 - \bar{S} \right) \kappa t \log \kappa.
\]

On the other hand, the variance is always a monotonically increasing function in \( t \), which is obvious from the first derivative of the variance:

\[
\frac{2 \sigma^2 \log \kappa \kappa^2}{\kappa^2 - 1} > 0.
\]

In practical application it is always questionable how to determine the mean-reversion level \( \bar{S} \). Statistical procedures are readily available, but the true nature of \( \bar{S} \) remains latent. Usually, investors acting in the market have a perception about the mean level \( \bar{S} \), but do not really know the exact position at any time \( t \). Moreover, the term structure can be very complicated. In many markets one anticipates a seasonal shape. In some settings and for a Gaussian Ornstein-Uhlenbeck process, the optimal strategy \( \pi^* \) is determined by the sign of \( S_0 - \bar{S} \). It can be very misleading to assume a certain value for \( \bar{S} \) and to base decisions upon this mean-reversion level.

**Model Fitting and Parameter Scaling**

In the next two chapters we employ the discrete-time versions of the Brownian motion model and the Ornstein-Uhlenbeck process. The parameters are inferred by using real-world data. The data set used for estimation is comprised of the price series of the 2011 contracts of power base, ARA coal, and CO\(_2\) emissions. Assume we intend to liquidate the capacity of the coal-fired power plant introduced above. The delivery period is the year 2011 = \([T_0, T_1]\) and the liquidation period is some time prior to maturity, say September 2009 until September 2010. Table 3.2 summarizes the parameter estimates.

<table>
<thead>
<tr>
<th>Model</th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>( \kappa )</th>
<th>( S_0 ) in EUR</th>
<th>( \bar{S} ) in EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian Motion</td>
<td>-0.1, 0, 0.1</td>
<td>0.5</td>
<td>-</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Ornstein-Uhlenbeck</td>
<td>-</td>
<td>0.5</td>
<td>0.75, 0.85, 0.95</td>
<td>20, 15, 20, 25</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Initial model parameters that are used for the empirical analysis in the sequel. The granularity is daily. The data set used for estimation is comprised of the price series of the 2011 contracts of power base, ARA coal, and CO\(_2\) emissions.

From (5.14) we immediately have
3.3. SOLUTION APPROACHES

\[ S_t - S_{t-1} = \tilde{\varepsilon}_t, \]

with \( \tilde{\varepsilon}_t \sim N(\mu, \sigma^2) \). Both parameters can be estimated using maximum likelihood. The drift \( \mu \) is -0.1 EUR per day. To examine the impact of no drift and a positive drift on liquidation strategies, we also consider \( \mu = 0 \) and \( \mu = 0.1 \). The volatility is about 0.5 EUR. Scaling to a weekly or monthly granularity can be performed by multiplying \( \mu \) and \( \sigma^2 \) by 5 or 20, respectively. This follows directly from (3.12). As described in the Appendix A.1, the OU process can be estimated using the Yule-Walker equations (see Hamilton [65]). The volatility is about the same as for the random walk and the mean-reversion level is detected around 20 EUR. For a more comprehensive study we consider \( \bar{S} \in \{15, 20, 25\} \). The mean-reversion parameter is \( \kappa = 0.95 \), which corresponds to a half-life of about 14 days. Note that \( \eta = -\log \kappa \approx 0.05 \). We also examine different values for \( \kappa \) to see the impact on the liquidation strategies. Scaling \( \sigma^2 \) and \( \kappa \) is more involved. Consider (3.18) and to scale, for instance to a weekly basis we set \( \tilde{\kappa} = \kappa^5 \) and \( \tilde{\sigma}^2 = \sigma^2 \kappa^{10} \). Then, for a one week period

\[ S_{t|S_0} \sim N\left((1 - \tilde{\kappa}) \bar{S} + \tilde{\kappa}S_0, \tilde{\sigma}^2\right). \]

The initial spread price \( S_0 \) is simply taken from the historical spread price on day \( t = 0 \). After all, the parameter estimates should merely give a rough idea about the size of the model parameters as this yields realistic scenarios in the liquidation analysis.

### 3.3 Solution Approaches

Naturally, there are two ways to look at the liquidation problem: (1) One predetermines a liquidation schedule based upon the information available at time \( t = 0 \). (2) One takes future information into account and establishes a framework that incorporates these information appropriately.

#### 3.3.1 Static Liquidation

In a static framework, optimal liquidation is closely related to asset allocation. However, there are fine differences. Asset allocation has its origin in the seminal paper by Markowitz [96]. Holding on to our notation, the allocation problem can be formulated as follows. There are \( T \) risky assets \( S_t \) with weights \( \pi_t, t = 1, \ldots, T \) and variances \( \text{Var}[S_1], \ldots, \text{Var}[S_T] \) and one riskless asset \( S_0 \) with weight \( \pi_0 \). \( E_0[W_{T+1}] \) can be interpreted as the expected one-period return of the portfolio. In general, one assumes the assets to be multivariate normally distributed and allocates optimally in the sense that, e.g. the expected return is maximized under the constraint that a certain variance level is not exceeded. There are some differences to the program \( (P') \). First, the portfolio weights do not equal one in sum, but \( Q_{\text{tot}} \). This can be eliminated by normalizing the weights. Secondly, and more importantly, imposing a general covariance structure on the assets does not recover a Gaussian distribution for \( W_{T+1} \). However, modeling \( S_t \) via diffusions with i.i.d. increments yields indeed a Gaussian distribution for \( W_{T+1} \). Thus,
we do not have to make the assumption about normality.

The probabilistic constraint in \((P)\) is known in asset allocation problems and goes back to Telser [120]. Similar concepts have been suggested by Roy [109] and Kataoka [80]. In our case, we can show that this constraint is a second order cone, and since we imposed specific dynamics for \(S_t\), our analysis goes into a different direction. This is the subject of Chapter 4.

### 3.3.2 Dynamic Liquidation

A dynamic reaction to changing market situations requires a different problem formulation. A natural candidate to solve this issue is dynamic programming. Related known problems are knapsack problems or generalized optimal stopping problems. Dynamic liquidation has barely been treated in the academic literature. This might be due to the lack of closed-form solutions and the high complexity of the problems (in form of additional constraints). The general setup of a dynamic program requires the predefinition of (1) a finite or infinite horizon and (2) a discrete or continuous time formulation. Since we liquidate a portfolio subject to the boundary condition \(Q_{T+1} = 0\), we have naturally a finite horizon problem. The decision (2) is not as obvious. Continuous time formulations demand much more mathematical tools, however, the resulting partial differential equations can be embedded into a very elegant theory. On the other hand, discrete-time formulations are more intuitive and easier to handle. Sometimes optimal solutions are available explicitly. Since there is barely any literature on the liquidation problem, we choose to start with the discrete-time case. Future work might elaborate more on the continuous-time formulation.
Chapter 4

Optimal Static Liquidation Strategies

4.1 Introduction

When we speak of static liquidation we mean that the liquidation schedule is determined and fixed at time 0. This disables the flexible reaction to changing market situations, which is usually reflected in prices, interest/exchange rates, inventory levels, and other economic factors. We confine the discussion to uncertainty in prices, more specifically the spread prices $S_t$.

4.1.1 Related Literature & Contribution

Optimal liquidation can be regarded as a special case of asset allocation. There is a vast literature available starting basically with the seminal paper by Markowitz [96]. Since then, a tremendous group of researchers and practitioners have elaborated on ramifications and extensions, see for instance Almgren et al. [5], [6], [7]. However, liquidation problems, especially under risk constraints apart from mean-variance frameworks, have not been an issue, yet. We contribute with the current chapter to the existing literature by discussing a liquidation problem of practical relevance and its implication for optimal behavior in energy markets. We identify the task as a convex optimization problem, derive conditions under which optimal solutions exist, and give a verification theorem in form of the Karush-Kuhn-Tucker conditions (see Theorem 4.2.4 and 4.2.6). Furthermore, we conduct numerical experiments to show the sensitivities of the optimal strategies with respect to model parameters and the management rule $(PR, \alpha)$. The key result is the insight that risk averse traders tend to liquidate the open electricity position as fast as possible, regardless of the drift in the market (or induced by the model).

The sequel of this chapter is organized as follows:
CHAPTER 4. OPTIMAL STATIC LIQUIDATION STRATEGIES

The target function in (3.8) is a linear function that is maximized over the intersection of a compact and convex polytope and a cone of second order. In the absence of the management rule \((\text{PR}, \alpha)\), the optimization problem breaks down to a classical linear program, since the market constraints, reflected in \(\Pi\), can be written as a system of linear inequalities. This is derived in Section 4.2.1. In Section 4.2.2 we consider the simplified program \((P')\) and show that the Karush-Kuhn-Tucker conditions can be applied to verify the existence of optimal solutions \(\pi^*\). Taking the management rule into consideration leads to a more complicated set over which we maximize the target function. Yet, optimal solutions still exist, which is the scope of Section 4.2.3. In Section 4.3 we conduct a case study, which requires an algorithm to compute optimal solutions. Furthermore, we discuss the consequences of different underlying models driving \(S\) and how the choice of the risk level \(\alpha\) and the planning result \(\text{PR}\) influences the optimal strategies.

4.2 Optimal Liquidation as a Convex Optimization Problem

4.2.1 General Structural Properties

Consider the optimization problem \((P)\) introduced in Section 3.1. Let us assume for now that \(\pi_t \in \mathbb{R}, t = 0, ..., T\), i.e. a priori we do not rule out buying back the electricity. For any fixed \(t = 0, ..., T\), the quantities \(\pi_t\) are constants. Since we consider static strategies, the target function is obviously a linear function in \(\pi_t\):

\[
G : \mathbb{R}^{T+1} \rightarrow \mathbb{R} \\
G(\pi_0, ..., \pi_T) := \sum_{t=0}^{T} E_0 [S_t] \cdot \pi_t.
\]  

(4.2)

The expectation of \(S_t\) exists in all discussed cases. It is crucial to know the structure of the constrained space over which \(G\) is maximized. For this purpose, we already introduced the market dynamics in Section 3.2. They all have in common to exhibit Gaussian increments. In view of definition (3.3) we obtain \(W_{T+1} \sim \mathcal{N}(\mu_W, \sigma_W^2)\), by the convolution property of the Gaussian distribution. \(\mu_W\) and \(\sigma_W^2\) may depend on \(\pi_0, ..., \pi_T\), and for now the specific form of \(\mu_W\) and \(\sigma_W^2\), respectively, is irrelevant. We write the constraints of the program \((P)\), defined in (3.8) in terms of two separate sets. The set \(\Pi\) specified in (3.5) and (3.7) represents the technical and market constraints that we discussed above. Furthermore, the constraint on the planning result can be rewritten as
4.2. OPTIMAL LIQUIDATION AS A CONVEX OPTIMIZATION PROBLEM

\[ P \left[ W_{T+1} \geq PR \right] = 1 - P \left[ \frac{W_{T+1} - \mu_W(\pi)}{\sigma_W(\pi)} < \frac{PR - \mu_W(\pi)}{\sigma_W(\pi)} \right] \]

\[ = 1 - \Phi \left( \frac{PR - \mu_W(\pi)}{\sigma_W(\pi)} \right) \]

\[ = \Phi \left( \frac{\mu_W(\pi) - PR}{\sigma_W(\pi)} \right) \geq 1 - \alpha, \]

where \( \Phi(.) \) denotes the cumulative distribution function of the standard Normal distribution. This inequality is equivalent to

\[ f(\pi) := PR - \mu_W(\pi) - \Phi^{-1}(\alpha)\sigma_W(\pi) \leq 0. \] (4.3)

In other words, we are looking for admissible strategies \( \pi = (\pi_0, \ldots, \pi_T) \) that satisfy inequality (4.3). For this sake, let us define the set

\[ \Gamma := \left\{ \pi \in \mathbb{R}^{T+1} : f(\pi) \leq 0 \right\}. \] (4.4)

With these considerations we can state the optimization problem \((P)\) in shorter form:

\[ (P) \max_{\pi \in \Pi \cap \Gamma} G(\pi). \] (4.5)

Let us establish important structural properties of the constrained space \( \Pi \cap \Gamma \).

**Lemma 4.2.1**

1. For \( \pi \in \mathbb{R}^{T+1} \), \( \Pi \) is a convex, compact, and \( T \)-dimensional polytope.

2. \( \Gamma \) is a second-order cone (and hence convex) if and only if \( \alpha \leq 0.5 \).

3. \( \Pi \cap \Gamma \) is a convex set if and only if \( \alpha \leq 0.5 \).

4. For \( \alpha \to 0 \) (\( PR \) fixed) or \( PR \to \infty \) (\( \alpha \) fixed), the set \( \Gamma \) moves away from the hypercube \([0, Q_{\text{max}}]^{T+1}\). This always leads to \( \Pi \cap \Gamma = \emptyset \).

**Proof:**

1. First consider from (3.5)

\[ (Q_t - (T-t)Q_{\text{max}})^{+} \leq \pi_t \leq \min \{Q_t, Q_{\text{max}}\}, t = 0, \ldots, T. \] (4.6)

The key idea is to write these inequalities as a system of linear equations. We rewrite the upper inequality of (4.6) as
\[
\pi_t - \min \left\{ Q_{\text{tot}} - \sum_{n=0}^{t-1} \pi_n, Q_{\text{max}} \right\} \leq 0. \quad (4.7)
\]

Since \( \min\{a, b\} \leq a \) and \( \min\{a, b\} \leq b \) for any real numbers \( a \) and \( b \), we make (4.7) tighter by considering

\[
\sum_{n=0}^{t} \pi_n - Q_{\text{tot}} \leq 0 \\
\pi_t - Q_{\text{max}} \leq 0.
\]

Consider the lower inequality of (4.6) and write

\[
\left( Q_{\text{tot}} - \sum_{n=0}^{t-1} \pi_n - (T-t)Q_{\text{max}} \right)^+ - \pi_t \leq 0. \quad (4.8)
\]

Since \( \max\{a, 0\} \geq a \) and \( \max\{a, 0\} \geq 0 \) for any real number \( a \), the same reasoning as above leads from (4.8) to

\[
Q_{\text{tot}} - \sum_{n=0}^{t} \pi_n - (T-t)Q_{\text{max}} \leq 0 \\
-\pi_t \leq 0.
\]

Now, these inequalities can be rewritten as a linear system \( A\pi \leq b \) with \( A \) being a \((4T+4)\times(T+1)\) matrix and \( b \) being a \((4T+4)\)-dimensional vector:

\[
A = \begin{pmatrix}
1 & 0 & \ldots & 0 \\
1 & 1 & 0 & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
-1 & 0 & \ldots & 0 \\
-1 & -1 & 0 & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
-1 & \ldots & \ldots & -1 \\
1 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & 1 \\
-1 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & -1
\end{pmatrix}, \quad b = \begin{pmatrix}
Q_{\text{tot}} \\
\ldots \\
Q_{\text{tot}} \\
\ldots \\
TQ_{\text{max}} - Q_{\text{tot}} \\
(T-1)Q_{\text{max}} - Q_{\text{tot}} \\
-\pi_{t+1} \\
Q_{\text{max}} \\
\ldots \\
Q_{\text{max}} \\
0 \\
\ldots \\
0
\end{pmatrix} \quad (4.9)
\]
4.2. OPTIMAL LIQUIDATION AS A CONVEX OPTIMIZATION PROBLEM

This is a half-space representation of a closed and convex polyhedron. It is bounded, because the equation $A\pi = 0$ yields only the trivial solution $\pi = 0$. A bounded polyhedron is called polytope. The dimension of this polytope is one less than the rank of $A$, since the matrix inequality contains the conditions $e'\pi \leq Q_{\text{tot}}$ and $e'\pi \geq Q_{\text{tot}}$, where $e$ denotes the $(T + 1)$ dimensional row-vector containing ones on each position. This is intuitively clear, since $\pi_T$ can always be expressed in terms of $\pi_0, ..., \pi_T - 1$: $\pi_T = Q_{\text{tot}} - \sum_{t=0}^{T-1} \pi_t$. The set $\Pi$ can now be restated as

$$\Pi = \left\{ \pi \in \mathbb{R}^{T+1} : A\pi \leq b \right\}.$$  \hspace{1cm} (4.10)

This shows assertion 1.

2. For $\alpha \leq 0.5$, inequality (4.3) can be rewritten as

$$||\Phi^{-1}(1 - \alpha)\Sigma^{1/2}\pi||_2 \leq \pi' E_0[S] - PR,$$  \hspace{1cm} (4.11)

where $\pi'$ is the transpose of the vector $(\pi_0, ..., \pi_T)$ and $S = (S_0, ..., S_T)$. A second-order cone constraint is in general given by

$$||\hat{A}\pi + \hat{b}||_2 \leq \pi' \hat{a} + \hat{d},$$

for some matrix $\hat{A}$ and vectors $\hat{a}$, $\hat{b}$ and $\hat{d}$ (Lobo et al. [91]). $||.||_2$ denotes the Euclidean norm. Hence, set $\hat{A} = \Phi^{-1}(1 - \alpha)\Sigma^{1/2}$, $\hat{b} = 0$, $\pi' \hat{a} = \mu_W$ and $\hat{d} = -PR$.

A cone of second order is always convex, however, for the sake of completeness we argue explicitly: Consider the mapping $\pi \mapsto \sigma_W(\pi)$. Define $\pi := \lambda \tilde{\pi} + (1 - \lambda) \hat{\pi}$ with $\tilde{\pi}, \hat{\pi} \in \mathbb{R}^{T+1}_+$ and let $\lambda \in [0, 1]$. Then

$$\sigma_W(\pi) = \begin{cases} \sqrt{\pi'\Sigma\pi} & \text{Cholesky of } \Sigma \\ = \sqrt{\pi'\Sigma(\pi'\Sigma')'} \\ = ||\pi'\Sigma'||_2 \\ = ||\lambda \tilde{\pi}'C + (1 - \lambda) \hat{\pi}'C||_2 \\ \leq \lambda ||\tilde{\pi}'C||_2 + (1 - \lambda) ||\hat{\pi}'C||_2 \\ = \lambda\sigma_W(\tilde{\pi}) + (1 - \lambda)\sigma_W(\hat{\pi}). \end{cases}$$

Hence, $\pi \mapsto \sigma_W(\pi)$ is a convex function (in general, every norm is a convex function, which follows directly from the triangle inequality). Furthermore, $-\sigma_W$ is obviously concave and the sign of $\sigma_W$ is determined by the sign of $\Phi^{-1}(\alpha)$, which is negative for $\alpha \in [0, 0.5)$ and positive for $\alpha \in (0.5, 1]$. Therefore, $\mu_W + \Phi^{-1}(\alpha)\sigma_W$ is concave if and only if $\alpha \in [0, 0.5)$. Now,

$$\Gamma = \left\{ \pi \in \mathbb{R}^{T+1}_+ : \mu_W + \Phi^{-1}(\alpha)\sigma_W \geq PR \right\}.$$
is the upper contour set of a concave function, and henceforth convex.

3. Follows from 2. and the fact that two intersecting convex sets are again convex.

4. Fix \( PR \). For \( \alpha \to 0 \) we have \( \Phi^{-1}(\alpha) \to -\infty \). From (4.3) it follows that,

\[
\frac{PR - \mu_W(\pi)}{\sigma_W(\pi)} \leq \Phi^{-1}(\alpha) \to -\infty \text{ as } \alpha \to 0.
\]

In other words, for the left-hand-side being fixed, the decreasing upper bound dismisses all \( \pi \in [0,Q_{\max}]^{T+1} \). Furthermore, for \( PR \to \infty \) (\( \alpha \) fixed) we obtain

\[
\mu_W(\pi) + \Phi^{-1}(\alpha)\sigma_W(\pi) \geq PR \to \infty.
\]

This completes the proof.

\[\text{Remark 4.2.2} \quad \text{The last statement in Lemma 4.2.1 has an obvious interpretation. A fairly small }\alpha \text{ is equivalent to a strict management rule (strict in a sense that it is more difficult to meet the planning result at the end of the liquidation period). Hence, it is natural to expect that there exist less admissible strategies and in the extreme situation } \alpha = 0 \text{ there exist no strategies at all, because } \Pi \cap \Gamma = \emptyset. \text{ Furthermore, if the planning result is chosen too large, it is more and more unlikely that there exists a strategy } \pi \text{ under which the accumulated wealth } W_{T+1} \text{ can exceed } PR.\]

\[\text{Corollary 4.2.3} \quad \text{Assume, the management rule } (PR,\alpha) \text{ is chosen such that } \Pi \cap \Gamma \neq \emptyset \text{ and } \pi \in \Pi, \text{ where } \Pi \text{ is defined in (3.5). Assume furthermore that } E_0[S_t] \text{ is well-defined for } t = 0, \ldots, T. \text{ Then there exists an optimal liquidation strategy } \pi^* = (\pi_0^*, \ldots, \pi_T^*) \text{ that solves the maximization problem } (P).\]

\[\text{Proof:} \quad \text{Since } \Pi \cap \Gamma \text{ is non-empty, the admissible set } \Pi \cap \Gamma \text{ is compact. The expectation w.r.t. the spread dynamics is well-defined. Furthermore, the function } G \text{ defined in (4.2) is obviously linear in } \pi, \text{ and hence continuous. Continuous functions always attain their maximum on compact sets. This follows from the Weierstrass theorem.}\]

This Corollary gives us the justification to employ solution methods to calculate optimal liquidation strategies \( \pi^* \). Figure 4.1 shows the admissible set \( \Pi \cap \Gamma \) for the case \( T = 2 \) and varying values of \( PR \in \{500 \text{ EUR}, 1000 \text{ EUR}\} \). The set \( \Gamma \) is indicated by plotting its convex hull. We assumed a simple Brownian motion with drift and representative parameters from Table 3.2. It illustrates the consequence of a too large planning result as it leads to an empty intersection between \( \Pi \) and \( \Gamma \).

In convex optimization theory there are useful tools to characterize optimal solutions \( \pi^* \). Sufficient and necessary conditions for the existence of optimal solutions are given by the Karush-Kuhn-Tucker conditions (KKT), see Kuhn & Tucker [88] and Padberg [104]. We derive these conditions for our specific liquidation problem.
4.2. OPTIMAL LIQUIDATION AS A CONVEX OPTIMIZATION PROBLEM

Figure 4.1: This figure shows the admissible set $\Pi \cap \Gamma$ for different values of $PR \in \{500, 1000\}$ (left to right). Parameters for illustration are $T = 2$, $Q_{\text{tot}} = 70 \text{ MW}$, $Q_{\text{max}} = 45 \text{ MW}$, $\alpha = 0.02$, $\sigma = 2.2$, $\mu = 0.5$, $S_0 = 20 \text{ EUR}$.

4.2.2 Conditions for Optimal Solutions of $(P')$

Let us consider the much simpler and linear program

\[
(P') \quad \begin{cases} 
G(\pi) \to \max_{\pi} \\
A\pi \leq b \\
\pi \in \mathbb{R}^{T+1}
\end{cases}
\]

where $A$ and $b$ are defined (4.9). We rewrite $(P')$ and obtain

\[
(P') \quad \begin{cases} 
-G(\pi) \to \min_{\pi} \\
A\pi - b \leq 0 \\
\pi \in \mathbb{R}^{T+1}
\end{cases}
\]

Next, define the Lagrange function $L : \mathbb{R}^{T+1} \times \mathbb{R}^{4T+4} \to \mathbb{R}$ as it is useful to characterize optimal solutions:

\[
L(\pi, \lambda) := -G(\pi) + \lambda (A\pi - b), \quad (4.12)
\]

where $\lambda \in \mathbb{R}^{4T+4}$ denote the Lagrange multiplier (or dual variables). We have the following connection between the Lagrange function and $(P')$:

\[
\sup_{\lambda \in \mathbb{R}^{4T+4}} L(\pi, \lambda) = \begin{cases} 
-G(\pi), & A\pi - b \leq 0 \\
+\infty, & A\pi - b > 0
\end{cases}
\]

The next proposition characterizes the existence of an optimal solution $\pi^*$. 
Proposition 4.2.4 Consider the linear program \((P')\) as stated above. \(\pi^*\) is an optimal solution of \((P')\) if and only if \(\exists \lambda^* \in \mathbb{R}_{++}^{4T+4}\) such that \((\pi^*, \lambda^*)\) is a saddle point of the corresponding Lagrange function (4.12). More specifically, the saddle point of \(L\) is characterized by the Karush-Kuhn-Tucker (KKT) conditions:

\[\pi^* \text{ is a minimum of } \pi \mapsto L(\pi, \lambda^*) \quad (4.13)\]
\[(\lambda^*)' (A\pi^* - b) = 0 \quad (4.14)\]
\[A\pi^* - b \leq 0 \quad (4.15)\]
\[\lambda^* \geq 0 \quad (4.16)\]

Proof: Let (4.13)-(4.15) be fulfilled. Then

\[\sup_{\lambda \in \mathbb{R}_{++}^{4T+4}} L(\pi^*, \lambda) = L(\pi^*, \lambda^*) = \inf_{\pi \in \mathbb{R}^{T+1}} L(\pi, \lambda^*). \quad (4.17)\]

This is a characterization of a saddle point. On the contrary, let \((\pi^*, \lambda^*)\) be a saddle point. Then

\[-G(\pi) + \lambda'(A\pi^* - b) \leq -G(\pi) + (\lambda^*)'(A\pi^* - b), \]

yields (4.13)-(4.15). The optimality of \(\pi^*\) follows from (4.17).

The KKT conditions give necessary and sufficient conditions for the existence of an optimal solution. (4.13)-(4.16) can be stated in terms of the Lagrange function \(L\) as it is differentiable w.r.t. \(\pi\) and \(\lambda\), respectively.

\[L_{\pi}(\pi^*, \lambda^*) = 0 \quad (4.18)\]
\[(\lambda^*)' L_{\lambda}(\pi^*, \lambda^*) = 0 \quad (4.19)\]
\[L_{\lambda}(\pi^*, \lambda^*) \leq 0 \quad (4.20)\]
\[\lambda^* \geq 0. \quad (4.21)\]

The system of equations resulting from the KKT conditions (4.18)-(4.21) are too complicated to be solved directly. Hence, Proposition 4.2.4 can only be used as a verification theorem for the solution \(\pi^*\). The KKT conditions (4.18)-(4.19) are explicitly given by

\[
\begin{align*}
-\mathbb{E}_0[S_0] + \sum_{n=1}^{T+1} \lambda_n & - \sum_{n=T+2}^{2T+2} \lambda_n + \lambda_{2T+3} - \lambda_{3T+4} = 0 \\
-\mathbb{E}_0[S_1] + \sum_{n=2}^{T+1} \lambda_n & - \sum_{n=T+3}^{2T+2} \lambda_n + \lambda_{2T+4} - \lambda_{3T+5} = 0 \\
& \vdots \\
-\mathbb{E}_0[S_T] + \lambda_{T+1} & - \lambda_{2T+2} + \lambda_{3T+3} - \lambda_{4T+4} = 0
\end{align*}
\]
4.2. OPTIMAL LIQUIDATION AS A CONVEX OPTIMIZATION PROBLEM

(4.19) becomes

$$
\begin{align*}
\lambda_1 (Q_{\text{tot}} - \pi_0) &= 0 \\
\lambda_{T+1} \left( Q_{\text{tot}} - \sum_{n=0}^{T} \pi_n \right) &= 0 \\
\lambda_{T+2} (\pi_0 + TQ_{\text{max}} - Q_{\text{tot}}) &= 0 \\
\lambda_{2T+2} \left( \sum_{n=0}^{T} \pi_n - Q_{\text{tot}} \right) &= 0 \\
\lambda_{2T+3} (Q_{\text{max}} - \pi_0) &= 0 \\
\lambda_{3T+3} (Q_{\text{max}} - \pi_T) &= 0 \\
\lambda_{3T+4} \pi_0 &= 0 \\
\lambda_{4T+4} \pi_T &= 0
\end{align*}
$$

In the following theorem we specify the structure of the optimal solution $\pi^*$, which depends solely on the direction of the drift induced by the underlying dynamics.

**Theorem 4.2.5** Consider again program (P'). Let $Q_t = Q_{\text{tot}} - \sum_{n=0}^{t-1} \pi_n^*$ be the remaining quantity at time $t$, and let $\Pi$ be non-empty.

1. If $t \mapsto E_0 [S_t] \nearrow$, then the optimal policy is given by
   $$\pi_t^* = \min \{Q_t, Q_{\text{max}}\}, t = 0, ..., T.$$

2. If $t \mapsto E_0 [S_t] \searrow$, then the optimal policy is given by
   $$\pi_t^* = (Q_t - (T - t)Q_{\text{max}})^+, t = 0, ..., T.$$

3. If $t \mapsto E_0 [S_t]$ is constant, then all admissible policies $\pi \in \Pi$ are optimal.

**Proof:** A direct verification of the KKT conditions reveals the validity of the optimal solution $\pi^*$. We sketch case 1 as the other cases are similar. Since $\pi_t^* = \min \{Q_t, Q_{\text{max}}\}$, for $t = 0, ..., T$ and $Q_{\text{tot}} \leq (T + 1) \cdot Q_{\text{max}}$ (by assumption that $\Pi \neq \emptyset$), we obtain immediately $A\pi^* - b \leq 0$, which gives us (4.20). This system of equations contains in general strict inequalities. In these cases the dual variables $\lambda$ become zero. The other cases have to be checked separately. Without loss of generality, we can assume there exists a time index $t$, such that $\pi_0^*, ..., \pi_t^* \in (0, Q_{\text{max}}]$ and $\pi_{t+1}^*, ..., \pi_T^* = 0$. Then, using $E_0 [S_0] \leq ... \leq E_0 [S_T]$, we verify (4.18), (4.19), and (4.21). Proposition 4.2.4 ensures the optimality of $\pi^*$. This completes the proof.

The optimal solution $\pi^*$ depends solely on the monotonicity property of the sequence $E_0 [S_0], ..., E_0 [S_T]$. This result illustrates the usefulness of the KKT conditions. Indeed,
Corollary 4.2.3 provides the existence of optimal solutions, however, it does not provide a tool to check for the optimality of a solution. This is provided by Proposition 4.2.4. Note that $\Pi$ is non-empty as long as $Q_{\text{tot}} \leq (T+1) \cdot Q_{\text{max}}$. For $Q_{\text{tot}} = (T+1) \cdot Q_{\text{max}}$ the optimal solutions become $\pi^*_t = Q_{\text{max}}$, for $t = 0, ..., T$.

4.2.3 Conditions for Optimal Solutions of $(P)$

We extend the setting from Section 4.2.2 by augmenting the admissible set by $\Gamma$, as introduced in (4.4). As shown in Lemma 4.2.1, the set of admissible strategies is an intersection of a convex polytope $\Pi$ and a second-order cone $\Gamma$. Unfortunately, the resulting structure of $\Pi \cap \Gamma$ cannot be described in an elegant and concise fashion, i.e. a closed representation of the set does not exist. However, we can extend the KKT conditions by the additional constraint (4.3) and arrive at

$$
(P) \left\{ \begin{array}{l}
-G(\pi) \to \min_{\pi} \\
A\pi - b \leq 0 \\
f(\pi) \leq 0 \\
\pi \in \mathbb{R}^{T+1}
\end{array} \right.
$$

$f(\pi)$ is defined in (4.3). Proposition 4.2.4 cannot be transferred to the extended program $(P)$. We introduce the Slater condition, which is necessary for non-linear constraints. The condition excludes constraints of the form $f(\pi) = 0$.

$$(S) \exists \hat{\pi} \in \mathbb{R}^{T+1} \text{ with } f(\hat{\pi}) < 0.$$ 

The Lagrange function defined in (4.12) has to be extended by an additional dual variable $\xi \in \mathbb{R}_+$. Therefore, we consider $L : \mathbb{R}^{T+1} \times \mathbb{R}_+^{4T+4} \times \mathbb{R}_+$ and

$$L(\pi, \lambda, \xi) := -G(\pi) + \lambda(A\pi - b) + \xi f(\pi). \quad (4.22)$$

Now we can state the main result of this section, which is similar to Proposition 4.2.4.

**Proposition 4.2.6** Consider the linear program $(P)$ and assume that $(S)$ holds. $\pi^*$ is an optimal solution of $(P)$ if and only if $\exists \lambda^* \in \mathbb{R}_+^{4T+4}$, $\xi^* \in \mathbb{R}_+$ such that $(\pi^*, \lambda^*, \xi^*)$ is a saddle point of the corresponding Lagrange function (4.22). More specifically, the saddle point of $L$ is characterized by the Karush-Kuhn-Tucker (KKT) conditions:

$$
\frac{\partial L}{\partial \pi} = 0 \quad (4.23)
$$

$$
(A\pi^* - b) = 0 \quad (4.24)
$$

$$
\xi^* f(\pi^*) = 0 \quad (4.25)
$$

$$
\lambda^* \geq 0 \quad (4.26)
$$

$$
\xi^* \geq 0 \quad (4.27)
$$

$$
(\lambda^*)' (A\pi^* - b) = 0 \quad (4.28)
$$

$$
\xi^* f(\pi^*) = 0 \quad (4.29)
$$
4.3. SOLUTION METHODS AND CASE STUDY

Proof: The necessary condition can be shown by following the same lines as in the proof of Proposition 4.2.4. To show the sufficient condition, one applies the separating hyperplane theorem, see Rockafellar [108]. With the help of the hyperplane one can construct a contradiction to the Slater condition (S), which yields the validity of the KKT conditions.

Condition (4.23) can again be written explicitly, but we do not carry out this calculation here for two reasons. First, it merely extends (4.13) by the derivative of \( f(\pi) \) (see Algorithm 1). This derivative depends on the underlying model and \( \pi \) is the argument of the parameters \( \mu_W \) and \( \sigma_W \), see (4.3). Secondly, the extended program \((P)\) cannot be solved explicitly and we employ an algorithm in the next section. As can be shown, this algorithm provides the optimal solution of \((P)\), i.e. we do not need to check the optimality of \( \pi^* \) with Proposition 4.2.6.

We saw in Theorem 4.2.5 that the optimal strategies are either selling as fast as possible or selling as slowly as possible. The planning result constraint might pull the optimal solution somewhere in between. In Figure 4.1 the optimal strategy lies within the intersection of \( \Pi \) and \( \Gamma \), the planning result does not lead to a suboptimal strategy \( \hat{\pi} \). We define a suboptimal strategy \( \hat{\pi} \) to be a strategy that yields a lower expected wealth compared to the optimal strategy \( \pi^* \), but fulfills the probabilistic constraint. The suboptimal optimal strategy \( \hat{\pi} \) can be divided in two categories: (1) The strategy remains sell in as few blocks as possible, not necessarily as fast as possible (i.e. in the first few weeks) or as slowly as possible (i.e. in the last weeks). In other words, the planning result constraint limits the one-year time frame. (2) The strategy spreads out, i.e. it takes longer to liquidate the total capacity. In this case, the planning result constraint confines the interval \([0, Q_{\text{max}}]\). In Section 4.3 we discuss these issues in more detail.

Solving \((P)\) explicitly is difficult due to the lack of closed-form expressions for the admissible set \( \Pi \cap \Gamma \). Thus, applying some algorithm suited for convex programming is necessary. A simple and intuitive algorithm is the cutting plane method due to Kelley [81]. We show in Section 4.3 how to modify the algorithm to make a suitable for our purposes. Alternative algorithms are e.g. gradient methods, which are also applicable in this setting, see Frank & Wolfe [60].

4.3 Solution Methods and Case Study

For the subsequent analysis we focus on two model dynamics, namely the Brownian motion with drift (model 1) and the arithmetic Gaussian Ornstein-Uhlenbeck process (model 2). They have been introduced in Section 3.2. This suffices to illustrate the different liquidation strategies that can be induced. We examine the sensitivities of the optimal strategies with respect to the model parameters, as well as with respect to the management rule \((\text{PR}, \alpha)\). We showed in Theorem 4.2.5 that the program \((P')\) can be solved by simple block strategies. Hence, no algorithm is necessary. For the program
(P) the situation is more involved. An algorithm is necessary to find optimal strategies and before we proceed with the analysis, we give a brief description of this algorithm.

### 4.3.1 A Cutting Plane Algorithm

We establish a cutting plane algorithm that solves (P). It originates in the work of J.E. Kelley [81]. When considering (P), three different cases can occur:

1. \( \Pi \cap \Gamma = \emptyset \Rightarrow \nexists \pi^* \).
2. \( \Pi \cap \Gamma \neq \emptyset \): Run simplex algorithm on the simplified program \((P')\) to find \( \pi^* \). If \( \pi^* \in \Pi \cap \Gamma \Rightarrow \text{Stop} \).
3. \( \Pi \cap \Gamma \neq \emptyset \): Run simplex algorithm on \((P')\) to find \( \pi^* \). If \( \pi^* \notin \Pi \cap \Gamma \Rightarrow \text{find a suboptimal solution } \hat{\pi} \in \Pi \cap \Gamma \).

Case 3 is the interesting one, because it requires to trim the admissible set \( \Pi \cap \Gamma \) until we find a suboptimal strategy \( \hat{\pi} \). This procedure is described in Algorithm 1.

The algorithm converges within a few seconds to a suboptimal solution \( \hat{\pi} \) and it can conveniently be implemented in Matlab. The following proposition summarizes some properties of the algorithm.

**Proposition 4.3.1**

1. \( \pi^k \in P_k \) and \( P_k \supset P_{k+1} \supset \Pi \cap \Gamma, \forall k \in \mathbb{N}_0 \).
2. \( f(\pi^k) > 0 \Rightarrow \pi^k \notin P_{k+1} \).
3. If Algorithm 1 aborts for some \( \pi^k, k \in \mathbb{N}_0 \), then \( \pi^k \) is an optimal solution of \((P)\). Otherwise, the sequence \( \{\pi^k\}_{k \geq 0} \) exhibits accumulation points \( \bar{\pi} \) and each \( \bar{\pi} \) is an optimal solution of \((P)\).

**Proof:** Kelley [81].

### 4.3.2 Further Properties of the Management Rule \((\text{PR}, \alpha)\)

Lemma 4.2.1 showed that a too strict management rule leads to \( \Pi \cap \Gamma = \emptyset \). We are interested in the combinations \((\text{PR}, \alpha)\) such that \( \Pi \cap \Gamma = \{\pi\} \). In other words, we are looking for the osculation points of both sets, \( \Pi \) and \( \Gamma \). This touching point exists, since both sets are convex and closed. \( \Gamma \) is not bounded. This provides the information about admissible management rules. Naturally, the resulting boundary in the \( \text{PR} - \alpha \)-plane ultimately depends on the underlying model and its parameters. For this sake it is necessary to parameterize the sets (more specifically, the boundaries of the sets) \( \Pi \) and \( \Gamma \), respectively. The set \( \Pi \) is independent of \((\text{PR}, \alpha)\). Therefore, we need to consider
Algorithm 1: Cutting Plane Algorithm

**Step 0:** If $\Pi \cap \Gamma = \emptyset$, then there does not exist an optimal solution $\pi^*$ of the program $(P)$. Otherwise, go to step 1.

**Step 1:** Assume $\Pi \cap \Gamma \neq \emptyset$. Find a polytope $P_0$ with the property $P_0 \supset \Pi \cap \Gamma$. This polytope exists, because $\Pi \cap \Gamma$ is compact and non-empty. Consequently, we take $P_0 = \Pi$. Set $k := 0$ and go to step 2.

**Step 2:** Find a solution $\pi^k$ of the auxiliary program

$$\begin{cases} G(\pi) \to \max \\ \pi \in P_k \end{cases}$$

This can be performed by a simple simplex algorithm. We look for strategies $\pi^k$ that fulfill the only additional constraint

$$f(\pi^k) = PR - \mu_w(\pi^k) - \Phi^{-1}(\alpha)\sigma_w(\pi^k) \leq 0.$$  

If $f(\pi^k) > 0$, then $\pi^k \notin \Pi \cap \Gamma$ and we go to step 3. Otherwise, $\pi^* = \pi^k$ is the optimal strategy and we stop.

**Step 3:** Since $\pi^k$ is not optimal (because it is not an element of the admissible set), we cut a slice off of $P_k$ by another hyperplane. This hyperplane is given by

$$\left\{ \pi \in \mathbb{R}^{T+1} : \nabla f(\pi^k)' \cdot \pi \leq \nabla f(\pi^k)' \cdot \pi^k - f(\pi^k) \right\},$$

with

$$\nabla f(\pi^k) = -\frac{\Phi^{-1}(\alpha)}{\sqrt{(\pi^k)'}\Sigma\pi^k} (\pi^k)' \Sigma - \begin{pmatrix} E_0 [S_0] \\ \vdots \\ E_0 [S_T] \end{pmatrix},$$

and $\Sigma$ is the covariance matrix which obtains its specific form from the underlying model: $\pi'\Sigma\pi = \text{Var} \left[ \sum_{t=1}^T S_t \pi_t \right]$. The new set $P_{k+1}$ can be defined as

$$P_{k+1} = P_k \cap \left\{ \pi \in \mathbb{R}^{T+1} : \nabla f(\pi^k)' \cdot \pi \leq \nabla f(\pi^k)' \cdot \pi^k - f(\pi^k) \right\}.$$  

Set $k := k + 1$ and go back to step 2.
\( \left( \text{PR}, \alpha \right) \mapsto \text{PR} - \mu W(\pi) - \Phi^{-1}(\alpha) \sigma W(\pi) \).

The first argument contributes linearly. \( \alpha \) is the argument of the inverted cumulative Gaussian distribution function, which is not given explicitly. For \( \alpha < 0.5 \), \( \Phi^{-1}(\alpha) \) is a concave function. This structure is ultimately reflected in the resulting boundary of the admissible pairs \( \left( \text{PR}, \alpha \right) \) in the \( \text{PR} - \alpha \)-plane, see Figure 4.2. Desirable was a continuous function \( g : [0, 1] \to \mathbb{R}_+ \), however, the sets \( \Pi \) and \( \Gamma \) are too complicated. They do not allow for an easy derivation of an explicit form for \( g \). Though, Algorithm 1 can be used to determine \( g \) numerically. We determine all pairs \( \left( \text{PR}, \alpha \right) \) for which \( \Pi \cap \Gamma \) contains approximately only one element.

For the subsequent analysis we use the parameters listed in Table 3.2 scaled to a monthly granularity. For the random walk we obtain \( \mu = -2, 0, 2 \), \( \sigma = 2 \), and for the autoregressive process we obtain \( \sigma = 2.2 \) and \( \kappa = 0.003, 0.04, 0.36 \). The power plant has a total capacity of \( Q_{\text{tot}} = 500 \text{ MW} \) and we assume that \( Q_{\text{max}} = 80 \text{ MW} \) per month.

### 4.3.3 Negative Market Drift

We discussed in Section 3.2 the different approaches to model the underlying market and the resulting drift. Both models have in common that there is a unique (negative) drift direction determined by \( \mu < 0 \) (model 1) and by \( \bar{S} - S_0 < 0 \) (model 2). Furthermore, there is no changing drift direction that could possibly be implied by, e.g. some deterministic seasonality. The optimal block strategy for both models is given by

\[ \pi^* = (80, 80, 80, 80, 80, 80, 80, 20, 0, 0, 0, 0, 0). \]

This is obvious since the drift has the same direction for both models. More interestingly, there is no sensitivity w.r.t. the management rule \( \left( \text{PR}, \alpha \right) \). However, there are pairs for which \( \Pi \cap \Gamma = \emptyset \). This is where the Slater condition from Proposition 4.2.6 comes into play. Assume model 1 (Brownian motion with negative drift). Then for \( S_0 < 0 \)

\[ f(\pi) = \text{PR} - S_0 \sum_{t=0}^{T} \pi_t - \mu \sum_{t=1}^{T} t \pi_t - \sigma \Phi^{-1}(\alpha) \sqrt{\sum_{t=1}^{T} t^2 \pi_t^2 + 2 \sum_{i=1}^{T-1} \sum_{j=i+1}^{T} \pi_i \pi_j \min\{i, j\}} > 0, \]

for all \( \pi \in \Pi \). That means, the Slater condition is not fulfilled and we cannot find an optimal solution of \((P)\). For \( S_0 > 0 \) the issue is slightly more involved. For simplicity, we set \( T = 1 \). Then

\[ f(\pi) = \text{PR} - S_0 \pi_0 - \pi_1 \left( S_0 + \mu + \sigma \Phi^{-1}(\alpha) \right)^{1/2} \leq 0, \]

which is equivalent to
\[ \pi_1 \begin{cases} \frac{PR - S_0\pi_0}{S_0 + \mu + \sigma\Phi^{-1}(\alpha)}, & \text{if } S_0 + \mu + \sigma\Phi^{-1}(\alpha) > 0 \\ \frac{PR - S_0\pi_0}{S_0 + \mu + \sigma\Phi^{-1}(\alpha)}, & \text{if } S_0 + \mu + \sigma\Phi^{-1}(\alpha) < 0. \end{cases} \] (4.31)

\[ \mu + \sigma\Phi^{-1}(\alpha) < 0 \text{ in any case. If } S_0 \text{ is not large enough, we have the second case in (4.31). Then, the simple block trading strategy is not being trimmed. However, for } S_0 \text{ being large enough, the first case in (4.31) can occur and the trader might indeed be induced to deviate from the block strategy. This is an astonishing fact, but can be explained as follows: } S_0 + \mu + \sigma\Phi^{-1}(\alpha) \text{ has to be positive to trim the block strategy. Since, } \sigma\Phi^{-1}(\alpha) < 0, \text{ either } S_0 \text{ has to be quite large or } \alpha \text{ has to be close to 0.5. The latter corresponds to a trader who is more willing to take a risk. In other words, there are situations where a trader might exploit the variability in the market (reflected in } \sigma) \text{ even though the drift is negative. For this action there must be a reasonable tradeoff between } S_0 \text{ and } \mu + \sigma\Phi^{-1}(\alpha). \text{ The same argumentation holds for } T > 1 \text{ and model 2.} \]

For the realistic parameter set taken from Table 3.2, we want to know the pairs \((\text{PR}, \alpha)\) that yield the osculation point of } \Pi \text{ and } \Gamma. \text{ This is depicted in Figure 4.2.}

![Figure 4.2: Left panel: Model 1 and the blue area depicts the admissible pairs \((\text{PR}, \alpha)\), which yield } \Pi \cap \Gamma \neq 0. \text{ Right panel: Model 2 and the blue area depicts the admissible pairs } \((\text{PR}, \alpha)\) which yield } \Pi \cap \Gamma \neq 0.

The graphs for both models are quite intuitive. A higher planning result has to be accompanied by a lower probability level to achieve this target. The block strategy under the random walk has an expected revenue of 7360 EUR. For a reasonable risk level of } \alpha = 0.02 \text{ the utility should not fix more than 4450 EUR for the planning result (model 1, left panel). Otherwise, there are no strategies that could ever fulfill the management rule.
4.3.4 Positive Market Drift

Now let $\mu = 2$ (model 1) and $S_0 - \bar{S} = 5$ (model 2). Figure 4.3 shows the liquidation strategies for model 1 and model 2, respectively. The left panel fixes $\alpha$ and the right panel fixes $\text{PR}$.

The difference to the negative market drift is obvious. First, consider model 1. Making the management rule $\left(\text{PR}, \alpha\right)$ stricter induces the investor to deviate from the optimal block strategy $\pi^* = (0, 0, 0, 0, 20, 80, 80, 80, 80, 80, 80)$.

Increasing $\text{PR}$ or decreasing $\alpha$ has basically the same effect, compare the left with the right panel in Figure 4.3. The suboptimal strategies show a pronounced deviation from the block strategy. For instance, under model 1 and $\alpha = 0.02$ fixed, increasing the planning result up to 14000 EUR yields the strategy $\hat{\pi} = (80, 78, 0, 0, 0, 0, 0, 22, 80, 80, 80, 80)$.

In contrast to a negative market drift, we have a reasonable tradeoff between expected gain and risk. In order to deviate from $\pi^*$ and except a lower expected gain induced by $\hat{\pi}$, the investor is exposed to lower risk, because he liquidates faster. It is astonishing that the investor is willing to waive considerable shares of the expected revenue. Instead of
spreading the liquidation quantities more evenly over the entire time period, the investor sells some at the very beginning and the rest at the very end. In other words, he divides his portfolio into a part that is *home and safe* and a part for speculative purposes. Under model 1, the expected gains from both strategies $\pi^*$ and $\hat{\pi}$ can be calculated by

$$
\mu_W(\pi^*) = \sum_{t=0}^{11} \pi^*_t (S_0 + \mu t) = 18,360 \text{ EUR},
$$

and

$$
\mu_W(\hat{\pi}) = \sum_{t=0}^{11} \hat{\pi}_t (S_0 + \mu t) = 16,544 \text{ EUR}.
$$

These monetary quantities are per MWh. In order to obtain the total amount from selling an entire yearly futures contract, one has to multiply the gain by 8760 (as one year has 8760 hours). $\pi^*$ outperforms $\hat{\pi}$ by about 1800 EUR/MWh. This is quite large but in view of the gap between a low planning result of 10,000 EUR and 14,000 EUR this appears as a reasonable number. If the utility sticks to a planning result of 14,000 EUR (which yields $\hat{\pi}$) and runs the block strategy $\pi^*$, then

$$
P \left[ \sum_{t=0}^{11} S_t \pi^*_t \geq 14,000 \right] = 1 - \Phi \left( \frac{14,000 - \mu_W(\pi^*)}{\sigma_W(\pi^*)} \right) = 67.36%.
$$

In other words, the planning result is only achieved with 67.36% probability.

The deviation from the optimal block strategy $\pi^*$ under model 2 is slightly different compared to the case of model 1. We can still observe a shift towards the other extreme strategy (sell as early as possible), however, there is also more weight in between, i.e. a flattening of the strategy. This is due to the fact that the volatility term structure is flatter, i.e. more well-behaved. Indeed, it is increasing, but converges to the upper bound $\frac{\sigma^2}{1-\kappa^2}$, see Section 3.2. In other words, the tradeoff between lower expected gain and lower volatility is not as large as in model 1, where the volatility is linearly increasing in $t$. The block strategy $\pi^*$ is the same as for model 1. The suboptimal strategy is approximately given by

$$
\hat{\pi} = (80, 80, 56, 8, 8, 8, 8, 8, 76, 80, 80).
$$

The expected gains from both strategies $\pi^*$ and $\hat{\pi}$ can be calculated by

$$
\mu_W(\pi^*) = \sum_{t=0}^{11} \pi^*_t \left( \tilde{S} + \left( S_0 - \tilde{S} \right) e^{-\kappa t} \right) = 12,349 \text{ EUR},
$$

and
\[ \mu_W(\hat{\pi}) = \sum_{t=0}^{11} \hat{\pi}_t \left( \bar{S} + (S_0 - \bar{S}) e^{-\kappa t} \right) = 11,612 \text{ EUR}. \]

The difference is significant with about 740 EUR/MWh. If the utility holds on to \( PR = 9500 \text{ EUR} \) and runs the block strategy \( \pi^* \), then the maximum probability that this planning result will be exceeded is 90%, after all. Note that we can not compare the expected gains between model 1 and model 2, because the drift they induce is fundamentally different. Furthermore, the calculation of the admissible pairs \((PR, \alpha)\) can be performed as in the negative drift case, but is omitted for brevity.

### 4.3.5 No Market Drift

What happens when the utility has no anticipation about future directions of the spread price \( S_t \)? Without the planning result constraint, all admissible strategies are optimal, as shown in Theorem 4.2.5. Taking a risk view into consideration confines the set of admissible strategies in the sense that it rules out slowly selling strategies. Figure 4.4 depicts the liquidation strategies for model 1 and model 2, respectively, under a fixed risk level \( \alpha = 0.02 \). The other case, where the planning result is being fixed yields similar results.

![Liquidation Strategies](image)

Figure 4.4: Liquidation strategies for both models, random walk and mean-reverting process.

For \( PR < 9000 \) there are all strategies optimal, the risk constraint does not have any influence on the strategies as it is too loose. Therefore, we simply take selling evenly over the entire time frame, i.e. roughly 42 MW per month, or \( \pi^*_t = \frac{Q_{\text{tot}}}{T+1} \) for \( t = 0, ..., T = 11 \). It is important to note that this optimal strategy is not unique. Now it is obvious that the planning result constraint pulls the optimal strategy towards a sell as fast as possible block strategy. The same effect can be observed under model 2. The reason for this behavior is even more evident compared to the earlier cases. The investor can choose from all admissible strategies. Under an additional risk constraint he rather chooses the strategy that fulfills this constraint and the risk is smaller the earlier he liquidates.
4.3. Solution Methods and Case Study

4.3.6 Comparison of Different Models for a Fixed Management Rule

It is evident that the management rule \((\text{PR}, \alpha)\) has a significant impact on the optimal strategy. Once the drift of the market is determined, we can enforce suboptimal strategies by making \((\text{PR}, \alpha)\) stricter. In extreme situations this could lead to an opposite liquidation behavior. On the other hand, we are also interested how sensitive the strategies are with respect to varying model parameters (or more generally, to different models). For this purpose we fix \((\text{PR}, \alpha)\) appropriately by setting \(\text{PR} = 9000\) EUR and \(\alpha = 0.02\). Furthermore, we assume a positive market drift. Apparently, the slope of the volatility term structure induces and influences certain liquidation actions. For this reason we look at the volatility term structure for model 1 and model 2, respectively.

Figure 4.5 shows a very different liquidation behavior of the strategies for both models. For model 2, the deviation from the block strategy \(\pi^*\) does not lead to a rapid shifting towards \(t = 0\), but it flattens more slowly throughout the entire liquidation period and approaches a completely flat strategy, i.e. \(\pi_t^* = \frac{Q_{\text{tot}}}{T+1}\) for \(t = 0, \ldots, T = 11\). The is due to the fact that the volatility becomes flatter, i.e. there is not too much of a difference between the volatility in the near future compared to the volatility in the far future. In other words, the investor does not have a strong incentive to shift the liquidation to the beginning of the period, because the tradeoff between expected gain and risk is very flat over the period. The type of behavior is not observable under model 1. If we wanted to flatten the volatility structure of model 1, we should let \(\sigma \to 0\). However, this scotches simultaneously the incentive to deviate from the optimal block strategy \(\pi^*\). To put it in a nutshell, shifting between

steep volatility term structure \(\leftrightarrow\) flat volatility term structure

yields fundamentally different results for model 1 and model 2, respectively.

Figure 4.5: Left panel: For model 1, the varying steepness of the volatility term structure leads to almost opposite strategies. Right panel: For model 2 the varying steepness of the volatility term structure leads to a flattening of the liquidation strategy.
4.3.7 General Remarks

It is decisive, how the management rule is being fixed. In practice, one usually calculates \( PR \) according to

\[
PR = c \cdot S_0 \cdot Q_{\text{tot}},
\]

for some constant \( c > 0 \). This yields sort of a heuristic planning result and the correct value for \( c \) might be more or less determined by the experience of the analysts and traders. Let’s look at some alternatives to determine the planning result. If the strategy applied to the calculation of \( PR \) is already the optimal one, then there is no chance that another strategy can outperform the expected terminal wealth. Consequently, a low \( PR \) should be taken into consideration to find optimal strategies. Let us assume a strategy closing as fast as possible:

\[
PR = \sum_{t=0}^{T} E_0[S_t] \cdot \min\{Q_t, Q_{\text{max}}\}.
\]

In this case, optimal strategies do not exist, unless \( \alpha \geq 0.5 \). Closing as slowly as possible reads as

\[
PR = \sum_{t=0}^{T} E_0[S_t] \cdot (Q_t - (T - t)Q_{\text{max}})^+.
\]

This strategy yields the worst alternative for \( PR \), which makes it very likely that under many parameter constellations the probabilistic constraint is fulfilled. Closing evenly over the entire horizon can be determined by

\[
PR = \frac{Q_{\text{tot}}}{T + 1} \sum_{t=0}^{T} E_0[S_t],
\]

which is in between to two first alternatives. From a risk perspective, the most reasonable strategy is sell as fast as possible: The variance of the terminal wealth \( W_{T+1} \) is minimized by \( \pi^*_t = (Q_t - (T - t)Q_{\text{max}})^+ , t = 0, ..., T \). This holds for model 1 and model 2, respectively.

We can derive an interesting inequality. Assume \( \mu_W > PR \). Then

\[
P\left[W_{T+1} \leq PR\right] = P\left[\mu_W - W_{T+1} \geq \mu_W - PR\right] \\
\leq P\left[\mu_W - W_{T+1} \geq \mu_W - PR\right] \\
\leq \frac{E_0[|W_{T+1} - \mu_W|]}{\mu_W - PR},
\]

by the Markov inequality. Consequently, to ensure \( P\left[W_{T+1} \leq PR\right] \leq \alpha \) the utility has to enforce
This is a helpful information in the planning process.

4.4 Conclusions & Discussion

Optimal liquidation in a static framework has not yet been carried out in this context. We contribute by identifying the heuristic (and praxis-relevant) task as a convex optimization problem. Structural properties are examined and optimality conditions are obtained in a straightforward manner. In the subsequent case study we demonstrate that already a simple model framework gives insight into optimal solutions.

Assuming a negative drift of the underlying market, the block optimal solution of \((P')\) is obviously \textit{sell as fast as possible}. We argued earlier that there is no reasonable tradeoff between the waiver of expected gain and an increase in volatility, because every sub-optimal strategies means, selling a little more slowly, i.e. accepting more exposure to price uncertainty. However, under certain parameter constellations, the trader might be induced to deviate from the simple block strategy. For instance, if the initial spread price \(S_0\) is large enough, he is willing to slow down the liquidation speed in order to profit from the fluctuations in the market, reflected in \(\sigma\). This behavior becomes more pronounced when \(\alpha\) is being reduced as this reflects risk appetite.

On the other hand, if the drift is positive, it implies a block strategy that lies at the end of the liquidation period meaning \textit{sell as slowly as possible}. In this case, the management rule \((\hat{PR}, \alpha)\) can induce a reasonable tradeoff, because suboptimal strategies \(\hat{\pi}\) mean selling earlier and this implies a smaller exposure to price variations.

In case where we assume no drift at all we see a similar behavior, i.e. the investor deviates from \(\pi^*\) by selling faster. As a consequence of the arguments above, it makes perfect sense that an investor who decides upon static information and takes a risk view into consideration always prefers selling faster to selling more slowly. We point out that the drift of the market is not as important. A varying risk appetite can outweigh the drift that an investor might anticipate.

An interesting difference reveals the comparison of the two models. The slope of the volatility term structure yields quite different liquidation strategies for model 1 and model 2, respectively. For model 1 it is impossible to create a flat term structure, unless in the trivial case \(\sigma \to 0\). On the other hand, model 2 can exhibit flat term structures, as long as \(\kappa\) is relatively close to zero. When the variation in the market is fairly small (i.e. \(\sigma \to 0\) for model 1 and \(\kappa \to 1\) for model 2), the management rule \((4.3)\) becomes approximately
\[ f(\pi) \approx PR - \mu_W(\pi). \]

The program \((P')\) determines a \(\pi^*\) with \(\mu_W(\pi) \leq \mu_W(\pi^*)\). Now it is obvious that for very small market fluctuations the optimal strategy \(\pi^*\) is always in \(\Pi \cap \Gamma\). Furthermore, the tradeoff between \(\mu_W(\pi)\) and \(\sigma_W(\pi)\) (more specifically, \(\Phi^{-1}(\alpha)\sigma_W(\pi)\)) can have an impact on the admissible set. However, as numerical experiments have shown, this impact has to be significant, which means the volatility has to be fairly big (more specifically, the slope of the volatility term structure has to be steep). This could lead to suboptimal strategies \(\hat{\pi}\), which yield a smaller expected revenue. However, the variation in the market is considered big enough that under favorable market conditions the management rule holds. In other words, \(\Phi^{-1}(\alpha)\sigma_W(\pi)\) has to have a significant impact on \(\mu_W(\pi)\) for varying \(\pi\) in order to make \(\Gamma\) cut off \(\pi^*\) obtained from solving \((P')\). I.e. if there is enough variability in the market, it could happen that the planning result constraint enforces a suboptimal strategy \(\hat{\pi}\). In the limit this means that if \(\sigma \to 0\) and \(PR = \mu_W(\pi^*) + \varepsilon\), then \(\Pi \cap \Gamma = \{\pi^*\}\), for some small \(\varepsilon > 0\).

In the next chapter, we extend the liquidation problem by incorporating future market information. This requires a dynamic formulation of the problem.
Chapter 5

Optimal Dynamic Liquidation Strategies

In the previous chapter we investigated static strategies. They do not account for future states of the economy such as fluctuating prices, changing inventories, and other uncertain economic factors. It is a natural requirement of a liquidation strategy to react to such changing conditions. Reevaluating the static strategy at each point in time can be performed, however, it is not clear whether the overall strategy is optimal. This chapter works out a recursive method, known as *dynamic programming* that is able to model this type of economy.

The starting point is again the two programs \((P')\) and \((P)\), defined in (3.2) and (3.8), respectively. The treatment of these objects in a dynamic context requires much more effort, e.g. when it comes to the definition of the important variables. This is carried out in Section 5.1.2. It is our main contribution to define an adequate stochastic dynamic program, to examine the existence of optimal solutions, derive explicit formulae for optimal strategies, and to scrutinize practical examples.

5.1 Introduction

5.1.1 Related Literature & Contribution

The available literature on optimal dynamic liquidation problems is commendably concise and most researchers focus on finding the optimal deterministic or static liquidation strategies. It is assumed that at the beginning of the trading period a fixed trading schedule is already determined. This schedule is then carried out irrespectively of the size and direction of the commodity price movements. However, unwinding large portfolio positions is a problem that builds on future states of the economy. This requires a problem formulation that reacts *dynamically* to these uncertain conditions. A natural candidate for this approach is the concept of stochastic dynamic programming (SDP). Its underlying idea, the *Bellman principle of optimality* gives a simple, yet powerful...
method to translate economic problems into the notion of mathematics. Dynamic programming (DP) is a recursive method for solving sequential decision problems. The term *Dynamic Programming* was introduced by Richard Bellman [12]. He was the first to recognize the common structure underlying most sequential decision problems.

There are basically two small branches on liquidation that have evolved in recent years. These are order book models, introduced by Obizhaeva & Wang [101] and further examined by Alfonsi et al. [4] and the extended Bachelier model, introduced by Almgren & Chriss (see comprehensive paper series, e.g. [5],[6], and [7]) in the context of asset allocation in a mean-variance framework. The execution strategies of Almgren & Chriss [6] are path-independent (also called static): They do not modify the execution speed in response to price movements during trading. They argue that reevaluating their optimization criterion at each intermediate time point yields a trading schedule that coincides with the trading schedule specified at the initial time. Bertsimas & Lo [17] discuss optimal liquidation in a single-asset case and derive explicit formulae. Other work on optimal execution of portfolio transactions was done by Konishi & Makimoto [86], Huberman & Stanzl [69], Engle & Ferstenberg [56], Krokhmal & Uryasev [87], Butenko et al. [29], He & Mamaysky [66]. Schöneborn [111] considers optimal execution for an investor with exponential utility function (CARA utility). For this specific utility function, optimal dynamic strategies are indeed the path-independent static trajectories of Almgren & Chriss [6].

Our contribution of this current chapter is manifold. A sound definition of a stochastic dynamic program (SDP) is the foundation for all further calculations and numerical schemes. We find frameworks that allow for explicit formulae for the the optimal liquidation strategies. These formulae are easy to analyze and provide insight into optimal liquidation behavior. Extending the settings by a target wealth is straightforward from an SDP point of view. All variables remain the same and existence results are readily applicable. However, explicit formulae are much more difficult to obtain. This motivates the development of an efficient grid search algorithm to examine optimal liquidation under the target wealth constraint.

The rest of this chapter is organized as follows:

In Section 5.1.2 we translate the economic problem \((P')\) into the notion of an SDP. We discuss the structure and properties of the necessary quantities and sets. This framework is also applicable to a slightly different form of \((P)\), which is discussed in Section 5.3. Furthermore, the imposed constraints and underlying spread models are a novelty to the literature. Section 5.2 derives explicit optimal strategies under different framework specifications. In Section 5.3 we extend our setting by introducing the target revenue constraint that we already discussed in Chapter 4. Since value functions and optimal strategies become quite involved, we employ numerical schemes to solve the optimization problem. A subsequent case study completes and illustrates the optimal behavior of a trader and draws conclusions about the differences between static and
5.1. INTRODUCTION

dynamic approaches and their implications for the expected wealth at the end of the trading period.

5.1.2 General Formulation as a Stochastic Dynamic Program

In contrast to the static liquidation case it is important to define the information set generated by the underlying stochastic processes. Let's return to the general liquidation problem \((P')\). The following discussion is readily applicable to a slimmed down version of \((P)\), see Section 5.3. The objective is

\[
\max_{\pi \in \Pi} E_0 \left[ \sum_{t=0}^{T} S_t \pi_t \right],
\]

where \(\Pi\) is defined in (3.5) and (3.7), respectively. For the stochastic dynamic program formulation we restrict ourselves to the discrete time case with a finite time horizon, denoted by \(T\). Other frameworks are conceivable, see Section 3.3.2. For a comprehensive treatment on (stochastic) dynamic programming see Bertsekas [15] and Bertsekas & Shreve [16]. We have an underlying dynamic system that evolves through time according to a (sometimes unknown) law of motion, which can be either deterministic or stochastic. At each point in time \(t = 0, \ldots, T\), the system is in a certain state. One can influence and alter the future states of the system by a control variable. The goal is to maximize the resulting revenue from taking certain actions (or controls). With this little heuristic, we already introduced some of the basic ingredients for any stochastic dynamic program, which we specify for our liquidation problem. Consider the tuple

\[
(E, A, \Pi, r, g),
\]

with the following meanings.

1. \(E\) denotes the state space. Here we let \(E = \mathbb{R}^3\). Elements of the state are the spread price \(s_t\) at time \(t\), the remaining quantity \(q_t\) at time \(t\), and the accumulated wealth \(w_t\) at time \(t\). Sometimes, we write \(x_t = (s_t, q_t, w_t) \in E \subset \mathbb{R}^3\). It turns out that without the target wealth constraint in program \((P)\), the state space is reduced by one dimension - the accumulated wealth \(w_t\). In this case we write \(x_t = (s_t, q_t) \in E \subset \mathbb{R}^2\). The distinction between a random variable and a realized value is important in this current section. For this purpose we use the following notation for \(t = 0, \ldots, T\) (to be precise, at \(t = 0\) all variables are deterministic):

<table>
<thead>
<tr>
<th>Description</th>
<th>Random Variable</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>(X_t)</td>
<td>(x_t)</td>
</tr>
<tr>
<td>Spread Price</td>
<td>(S_t)</td>
<td>(s_t)</td>
</tr>
<tr>
<td>Quantity</td>
<td>(Q_t)</td>
<td>(q_t)</td>
</tr>
<tr>
<td>Wealth</td>
<td>(W_t)</td>
<td>(w_t)</td>
</tr>
</tbody>
</table>
2. $A$ denotes the *action space*. Here we let $A = \mathbb{R}$. An action $a \in A$ in our setting means that the trader chooses on how much electricity to liquidate at some point in time.

3. $\Pi = (\Pi_0, ..., \Pi_T)$ denotes the admissible set and we already defined it in (3.5) and (3.7), respectively. As it turns out, $\Pi_t$ depends only on $q_t$: $\Pi_t(x_t) = \Pi_t(q_t)$.

4. $r : E \times A \rightarrow \mathbb{R}$ denotes the one-period reward function. We implement several forms of this function. They all have in common that they are continuous and differentiable. Here the reward function does not depend on $t$.

5. $g : E \rightarrow \mathbb{R}$ denotes the terminal reward function, which is independent of an action $a \in A$ at time $T$. This function is not necessarily continuous, see Section 5.3.

The *law of motion* describes how the system evolves from state $x_{t-1}$ to state $x_t$, given a certain action $a_{t-1} \in \Pi_{t-1}(q_{t-1})$:

$$
\begin{align*}
    s_t &= f(s_{t-1}, \varepsilon_t), \\
    q_t &= q_{t-1} - a_{t-1}, \\
    w_t &= w_{t-1} + s_{t-1} a_{t-1},
\end{align*}
$$

The function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is linear in both variables and the stochastic shocks stem from the Gaussian random variable $\varepsilon_t : \Omega \rightarrow \mathbb{R}$. We need some further notions, which we summarize in the following definition.

**Definition 5.1.1**

1. A *decision rule* (or control variable) at time $t$ is a measurable function $\pi_t : E \rightarrow A$ with $\pi_t(x_t) \in \Pi_t(x_t)$, $\forall x_t \in X$. $\pi_t(x_t)$ corresponds to the quantity of electricity that is to be sold at time $t$. In other words, for each possible state $x_t$ at time $t$, the decision rule assigns an action to this state. A sequence of decision rules $\pi := (\pi_0, ..., \pi_T)$ is referred to as policy. Sometimes we omit the argument $x_t$ in $\pi_t$ to ease notation.

2. The expected total wealth is defined as

$$
    V_0(x_0) := E_0 \left[ \sum_{t=0}^{T-1} r(X_t, \pi_t(X_t)) + g(X_T) \right], \forall x_0 \in X. \tag{5.4}
$$

3. The maximal expected total wealth is defined as

$$
    V_0(x_0) := \sup_{\pi \in \Pi} V_0(x_0), \forall x_0 \in X. \tag{5.5}
$$

4. A policy $\pi^*$ is optimal if

$$
    V_0(x_0) = V_0^{\pi^*}(x_0), \forall x_0 \in X.
$$
The initial state is always \( x_0 = (s_0, q_0, w_0) = (s_0, Q_{\text{tot}}, 0) \) and \( q_{T+1} = 0 \). It is important to guarantee the existence of the expectation in (5.4) and (5.5), respectively. This justifies deriving the Bellman equation in the next section. The following integrability condition \((I)\) must hold throughout the rest of this chapter:

\[
(I) \quad \sup_{\pi} E_0 \left[ \sum_{t=0}^{T-1} r^+ (X_t, \pi_t (X_t)) + g^+ (X_T) \right] < \infty.
\]

In the next section we derive the Bellman equation for this setup and prove an existence theorem.

5.2 Solving \((P')\): Explicit Optimal Liquidation Strategies

5.2.1 The Bellman Equation and Existence of Optimal Solutions

As in Chapter 4, we start the discussion by considering the simpler program \((P')\), i.e. we omit the target wealth constraint. In this case, the state variable \( x_t \) is downsized by the accumulated wealth \( w_t \), thus \( x_t = (s_t, q_t) \). This follows directly from the specific form of the Bellman equation derived below. Furthermore, we consider definition (3.5) for \( \Pi \). The one-period reward function depends only on \( s_t \) and \( a_t \) at time \( t \). At the beginning of the liquidation period we look at

\[
V_0(s_0, q_0) = \sup_{\pi \in \Pi} E_0 \left[ \sum_{t=0}^{T-1} r(S_t, \pi_t (S_t)) + g(S_T, Q_T) \right].
\]

The expectation is again taken with respect to the Gaussian shocks \( \varepsilon_1, ..., \varepsilon_T \). The Bellman equation of dynamic programming with a finite horizon is given by:

\[
V_T(s_t, q_t) = g(s_t, q_t)
\]

\[
V_t(s_t, q_t) = \max_{a_t \in \Pi_t(q_t)} \left\{ r(s_t, a_t) + E_t [V_{t+1}(f(s_{t+1}, \varepsilon_{t+1}, q_{t+1} - a_t))] \right\},
\]

for any \((s_t, q_t) \in E\). It can be shown that the recursion (5.7)-(5.8) is valid as long as \((I)\) holds, see Bertsekas & Shreve [16]. Now it becomes clear, why Bellman’s equation is so useful. It reduces the choice of a sequence of decision rules to a sequence of choices for the control variable. It is sufficient to solve the problem sequentially \( T + 1 \) times. Hence, a dynamic problem is reduced to a sequence of static problems. A consequence of this result is the so-called Bellman’s principle of optimality: If the sequence of functions \( \pi^* = (\pi_0^*, ..., \pi_T^*) \) is the optimal policy that maximizes \( V_0(s_0, q_0) \), then, if we consider the remainder of the objective function after \( t \) periods \( V_{t \pi_t}(s_t, q_t) \), the functions \( (\pi_t^*, ..., \pi_T^*) \), which were optimal for the original problem are still the optimal ones. Thus, as time advances there is no incentive to depart from the original plan. Policies with this property are called time-consistent. Time consistency depends on the recursive structure of the problem and does not apply to more general settings.
We want to know when the dynamic program (5.7)-(5.8) has a solution. To formulate an existence result, it is crucial require the admissible set $\Pi$ to be compact, which is guaranteed by the definition (3.5). The largest possible expansion of the interval (3.5) is $[0, Q_{\text{max}}]$ and the smallest is $\{0\}$. Moreover, $\Pi_T$ is always a single element set with values between 0 and $Q_{\text{max}}$: $\Pi_T = \{q_T\}$. This ensures that the boundary condition $q_{T+1} = 0$ is met. For our purposes it is sufficient to know that $q_T$ varies continuously in $\Pi_t(q_t)$. The following theorem gives us sufficient conditions for the existence of at least one optimal solution $\pi^*$ of $(P')$.

**Theorem 5.2.1** Assume that the admissible sets $\Pi_t(q_t)$, $t = 0, \ldots, T$ defined in (3.5) are non-empty, which implies that they are compact and real-valued intervals. The dynamics of $S$ are described by the function $f$ in (5.3). Then there exist measurable functions $\pi^*_t : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ as maximizers of (5.8) and the policy $\pi^* = (\pi^*_0, \ldots, \pi^*_T)$ solves $(P')$. Furthermore, the value functions $V_t(s_t, q_t)$, $t = 0, \ldots, T$, $(s_t, q_t) \in E$ are continuous.

**Proof:** We show via induction that the continuity property of the terminal value function is being passed on to all other value functions for $t = T - 1, \ldots, 0$. At maturity $T$ we have the terminal value function $V_T(s_T, q_T) = r(s_T, q_T)$. $r$ is a continuous function and $(s_T, q_T) \mapsto V_T(s_T, q_T)$ is also continuous as $q_T$ varies continuously in $\Pi_T(q_T)$. At time $T - 1$ we look at

$$(s_{T-1}, a_{T-1}) \mapsto r(s_{T-1}, a_{T-1}) + \mathbb{E}_{T-1} [V_T(f(s_{T-1}, \varepsilon_T), q_{T-1} - a_{T-1})].$$

Since $V_T$ is continuous and $\varepsilon_T$ are Gaussian i.i.d. random variables, the expectation is also continuous. Consequently, the entire expression is continuous and we maximize it over the compact interval $\Pi_{T-1}(q_{T-1})$. Thus, $\pi^*_{T-1}$ exists (being the maximizer) by the Weierstrass theorem. The maximized value function $(s_{T-1}, q_{T-1}) \mapsto V_T(s_{T-1}, q_{T-1})$ is also continuous in both variables. Now assume, $(s_t, q_t) \mapsto V_t(s_t, q_t)$ is continuous for arbitrary $t$. Then

$$(s_{t-1}, a_{t-1}) \mapsto r(s_{t-1}, a_{t-1}) + \mathbb{E}_{t-1} [V_t(f(s_{t-1}, \varepsilon_t), q_{t-1} - a_{t-1})]$$

is again continuous by the same argument as in $T - 1$. Thus, the maximum is attained at $\pi^*_{t-1}$ and $(s_{t-1}, q_{t-1}) \mapsto V_{t-1}(s_{t-1}, q_{t-1})$ is also continuous.

**Remark 5.2.2** In Section 5.2.3 we relax the condition on $\Pi$ to be compact, (i.e. we merely require (3.7)) and give examples where optimal solutions still exist. The optimality of these solutions is guaranteed by the sufficient optimality condition in form of the DP principle, stated below in Proposition 5.2.3. We also give examples which do neither satisfy Theorem 5.2.1, nor Proposition 5.2.3.

Starting at time $T$, going backwards in time and solving each single stage (time-$t$) maximization problem yields an optimal policy $\pi^* = (\pi^*_0, \ldots, \pi^*_T)$. This procedure is known as backward induction. The optimal strategy at time $T$ can always be found
5.2. SOLVING \((P')\): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES

easily by setting \(\pi_T^* = q_T\) (since \(\Pi_T(q_T)\) contains only one element). However, as we
go back in time, the value functions might become too complicated to be maximized
effectively. In the subsequent sections, we give examples where we can calculate \(V_t\)
and \(\pi_t^*\) explicitly. Then, Bertsekas [15] shows that the dynamic programming principle
\((5.7)-(5.8)\) leads indeed to the optimal policy \(\pi^*\). This gives us the justification to
derive optimal policies in Section 5.2: If we can find the maximum value function and
its corresponding maximizing control \(\pi_t^*\) at each stage \(t = 0,\ldots,T\), then the sequence
\(\pi_0^*,\ldots,\pi_T^*\) is optimal for the program \((P')\). This is stated in the next proposition.

Proposition 5.2.3 If an optimal control \(\pi^*_t\) exists for \((5.7)-(5.8)\) and \(t = 0,\ldots,T\), then
the policy \(\pi^* = (\pi_0^*,\ldots,\pi_T^*)\) is optimal for the program \((P')\).

Proof: The proof can be found in Bertsekas [15], p.23. It uses basically induction
arguments.

Remark 5.2.4 Proposition 5.2.3 is more general than Theorem 5.2.1. However, a re-
sult similar to Theorem 5.2.1 can be proved for the general case where \(\Pi_t\) is not neces-
sarily compact.

Let us introduce some variables to ease notation in the sequel. Define \(\pi_{t}^{\min}(q_t) := \min \Pi_t(q_t)\) and \(\pi_{t}^{\max}(q_t) := \max \Pi_t(q_t)\) as the minimal value in \(\Pi_t\) and maximal value
in \(\Pi_t\), respectively. These variables exist as long as \(\Pi_t\) is non-empty and definition
\((3.5)\) applies, and they depend on \(q_t\). Moreover, we define (if existent) the liquidation
enforcement time \(t'\) by

\[
t' := \min \left\{ t \in \{0,\ldots,T\} : \pi_t^{\min} > 0 \right\}.
\]  

(5.9)

This is the earliest time where the trader is forced to speed up liquidation due to the
boundary condition \(q_{T+1} = 0\). Furthermore, we define

\[
t'' := \min \left\{ t \in \{0,\ldots,T\} : \pi_t^{\max} = 0 \right\}
\]  

(5.10)
as the earliest point in time, where the portfolio can be completely liquidated. In contrast
to \(t'\), \(t''\) always exists. For the forthcoming analysis it is important to understand
the structure of the admissible set \(\Pi_t(q_t)\):

\[
\Pi_t(q_t) = \left[ (q_t - (T-t)Q_{\max})^+, \min \left\{ q_t, Q_{\max} \right\} \right].
\]

We intentionally write \(\pi_{t}^{\min}\) and \(\pi_{t}^{\max}\) as functions of \(a_{t-1}\), because we can see that the
upper and lower bound of \(\Pi_t(q_t)\) is linear and monotonically decreasing in \(a_{t-1}\) with
slope either \(-1\) or \(0\). Note that this holds under the assumption that \(q_t\) is fixed. Figure
5.1 depicts the upper, resp. lower bound of \(\Pi_t(q_t)\) for a specific parameter constellation.
There are three points in time left to liquidate the remaining quantity of \(q_{49} = 120\) MW.
If \(a_{49} < 40\) MW, then the trader is being enforced in \(t = 50\) to sell at least \((40 - a_{49})\)
MW to meet \(q_{52} = 0\) MW. On the other hand, if \(a_{49} > 40\) MW, then the trader cannot
sell the maximum possible quantity \(Q_{\max}\) in the last periods.
5.2.2 Frictionless Markets

We consider the common conception of a frictionless market, which is regarded a trading environment where all costs and restraints associated with transactions are non-existent (Merton [99], p.17). In this section, the one-period reward function and the terminal reward function become

\[
\begin{align*}
    r(s_t, a_t) &= s_t \cdot a_t, \\
    g(s_T, q_T) &= s_T \cdot q_T.
\end{align*}
\]

We need to check whether the integrability condition \((I)\) holds. Indeed, we verify

\[
\begin{align*}
    E_0 \left[ r^+ (S_t, \pi_t (S_t, Q_t)) \right] &= E_0 \left[ (S_t \pi_t)^+ \right] \\
    &\leq Q_{\text{max}} \cdot E_0 \left[ S_t^+ \right],
\end{align*}
\]

and

\[
\begin{align*}
    E_0 \left[ g^+ (S_T, Q_t) \right] &= E_0 \left[ (S_T Q_T)^+ \right] \\
    &\leq Q_{\text{max}} \cdot E_0 \left[ S_T^+ \right].
\end{align*}
\]
5.2. SOLVING \((P')\): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES

The expectations are finite. Therefore, \((I)\) holds and the Bellman equation \((5.7)-(5.8)\) is well-defined. Remember that the spread price \(S_t\) reflects the revenues from selling the electricity reduced by the fuel costs, since the quantities of necessary coal and CO\(_2\) emissions are tied deterministically through the efficiency rates, see Table 1.5. If the spread \(S_t\) is described by a simple random walk with drift, the optimal trading strategy \(\pi_t^*\) depends merely on the drift induced by the random walk and the initial portfolio position \(Q_{tot}\), as stated in the next proposition.

**Proposition 5.2.5** Consider the objective \((3.2)\) with \(0 \leq \pi_t \leq Q_{max}\) and assume \(Q_{max} \leq Q_{tot} \leq (T + 1) \cdot Q_{max}\). The spread is modeled by a random walk with drift, i.e. \(S_t = \mu + S_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)\). Then, the optimal liquidation strategies and value functions have the following form:

- If \(\mu \leq 0\), then for \(t = 0, ..., T\):
  \[
  \pi_t^* = \min \{ q_t, Q_{max} \}
  \]
  \[
  V_t(s_t, q_t) = -\mu \left( (T - t) \pi_t^{max} + ... + \pi_{T-1}^{max} \right) + [(T - t) \mu + s_t] \cdot q_t.
  \]

- If \(\mu > 0\), then for \(t = 0, ..., T\):
  \[
  \pi_t^* = (q_t - (T - t)Q_{max})^+
  \]
  \[
  V_t(s_t, q_t) = -\mu \left( (T - t) \pi_t^{min} + ... + \pi_{T-1}^{min} \right) + [(T - t) \mu + s_t] \cdot q_t.
  \]

**Proof:** It suffices to prove the case \(\mu > 0\) as \(\mu < 0\) works similarly. The condition \(Q_{max} \leq Q_{tot} \leq (T + 1) \cdot Q_{max}\) ensures that \(\Pi_t \neq \emptyset\) for \(t = 0, ..., T\). The assumptions of Theorem 5.2.1 are fulfilled, hence optimal solutions exist. In order to show the validity of the recursive formulae we proceed as follows. At each point in time \(t\) we consider \(s_t\) and \(q_t\) to be fixed. This allows us to determine a maximizer \(a_t^*\) for fixed \(s_t, q_t\). The optimal control is then obtained by setting \(\pi_t^* := a_t^* (S_t, Q_t)\). The terminal value function is

\[
V_T(s_T, q_T) = g(s_T, q_T) = s_T \cdot q_T.
\]

In order to meet the boundary condition \(q_{T+1} = 0\), the optimal strategy is simply \(\pi_T^* = q_T\), i.e. selling what is left in the portfolio. Going one step back in time yields

\[
V_{T-1}(s_{T-1}, q_{T-1}) = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ s_{T-1} a_{T-1} + V_T (\mu + s_{T-1}, q_{T-1} - a_{T-1}) \right\}
\]

\[
= \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ -\mu \pi_{T-1} + (\mu + s_{T-1}) q_{T-1} \right\}.
\]

The maximum is attained for the smallest value in \(\Pi_{T-1}(q_{T-1})\), which we denoted by \(\pi_{T-1}^{min}\). At time \(T - 2\), we obtain

\[
V_{T-2}(s_{T-2}, q_{T-2}) = \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ s_{T-2} a_{T-2} + V_{T-1} (\mu + s_{T-2}, q_{T-2} - a_{T-2}) \right\}
\]

\[
= \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ -\mu \left( 2a_{T-2} + \pi_{T-1}^{min} \right) \right\} + (2\mu + s_{T-2}) q_{T-2}.
\]
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES

Now, $\pi_{T-1}^{\text{min}}$ is a linear and decreasing function in $a_{T-2}$ (the more the trader sells at time $T-2$, the rather there is more freedom to sell less at time $T-1$) with slope either -1 or 0, compare Figure 5.1. Consequently the maximum is attained for the smallest value in $\Pi_{T-2}(q_{T-2})$, again denoted by $\pi_{T-2}^{\text{min}}$. Proceeding in this fashion yields for arbitrary $t$

$$V_t(s_t, q_t) = \max_{a_t \in \Pi_t(q_t)} \left\{ -\mu \left( (T-t)a_t + \ldots + \pi_{T-1}^{\text{min}} \right) + ((T-t)\mu + s_t)q_t \right\}$$

(5.11)

$$= -\mu \left( (T-t)\pi_{t-1}^{\text{min}} + \ldots + \pi_{T-1}^{\text{min}} \right) + ((T-t)\mu + s_t)q_t.$$  

(5.12)

The same reasoning as above suggests that the minimum value in $\Pi_t(q_t)$ attains the maximum. To complete the proof, we need to show that the recursive formula for $V_t(s_t, q_t)$ is valid for arbitrary $t$. Assume, (5.12) holds for $t$. We show that this formula holds also for $t-1$. Indeed, we verify

$$V_{t-1}(s_{t-1}, q_{t-1}) = \max_{a_{t-1} \in \Pi_{t-1}(q_{t-1})} \left\{ s_{t-1}a_{t-1} + V_t(\mu + s_{t-1}, q_{t-1} - a_{t-1}) \right\}$$

$$= \max_{a_{t-1} \in \Pi_{t-1}(q_{t-1})} \left\{ -\mu \left( (T-t+1)a_{t-1} + \ldots + \pi_{t-1}^{\text{min}} \right) + ((T-t+1)\mu + s_{t-1})q_{t-1} \right\}$$

$$= -\mu \left( (T-t+1)\pi_{t-1}^{\text{min}} + \ldots + \pi_{t-1}^{\text{min}} \right) + ((T-t+1)\mu + s_{t-1})q_{t-1}.$$ 

According to the definition of the sets $\Pi_0, \ldots, \Pi_T$ in (3.5) we obtain $\pi_t^* = \min \Pi_t(q_t) = (q_t - (T-t)Q_{\text{max}})^+$, $t = 0, \ldots, T$. For $\mu = 0$ all admissible strategies are optimal which follows directly from the specific form of $V_t$.

Remark 5.2.6 The optimal strategies in the last proposition are unique as long as $\mu \neq 0$. Furthermore, it shows that the result from Theorem 5.2.1 is carried over to this setting and we applied the DP algorithm (Proposition 5.2.3) to the explicitly calculated value functions and optimal policies.

The proof showed that the maximum is always attained on the boundary of $\Pi_t(q_t)$. In other words, the solution of (3.2) is a simple block trading strategy (sometimes referred to as bang-bang strategy) under a random walk with drift. If $Q_{\text{max}} = Q_{\text{tot}}$, we end up with a trivial strategy, i.e. either sell everything at time 0 ($\mu < 0$) or at time $T$ ($\mu > 0$). The liquidation enforcement time $t'$ defined in (5.9) does not exist for $\mu < 0$ as long as we employ the optimal strategy. In case of $\mu > 0$ the optimal liquidation strategy can be written in terms of $t'$:
5.2. SOLVING \((P')\): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES

\[
\begin{align*}
\pi^*_0 &= 0 \\
\ldots \\
\pi^*_{t-1} &= 0 \\
\pi^*_t &= (Q_{\text{tot}} \mod Q_{\text{max}}) I(Q_{\text{tot}} \mod Q_{\text{max}}) > 0 + Q_{\text{max}} I(Q_{\text{tot}} \mod Q_{\text{max}}) = 0 \\
\pi^*_{t+1} &= Q_{\text{max}} \\
\ldots \\
\pi^*_T &= Q_{\text{max}}.
\end{align*}
\]

\(Q_{\text{tot}} \mod Q_{\text{max}}\) denotes the remainder of the division of \(Q_{\text{tot}}\) by \(Q_{\text{max}}\). When imposing mean-reverting dynamics, the trader has one more piece of information - the mean-reversion level, denoted by \(\bar{S}\). While in practice it is questionable how to measure and determine this quantity, within the model framework it allows us to draw conclusion about optimal behavior. More specifically, we have the following result under a mean-reverting process.

**Proposition 5.2.7** Consider again the objective (3.2) with \(0 \leq \pi_t \leq Q_{\text{max}}\). Assume \(Q_{\text{max}} \leq Q_{\text{tot}} \leq (T + 1) \cdot Q_{\text{max}}\). The spread is modeled by the mean-reverting dynamics \(S_t = (1 - \kappa)\bar{S} + \kappa S_{t-1} + \varepsilon_t\) with \(\varepsilon_t \sim \mathcal{N}(0, \sigma^2)\) and \(0 < \kappa < 1\). The optimal liquidation strategy and value function for \(t = 0, \ldots, T\) are given by

\[
\begin{align*}
\pi^*_t &= (q_t - (T - t)Q_{\text{max}})^+ \mathbf{1}_{s_t \leq \bar{S}} + \min \{q_t, Q_{\text{max}}\} \mathbf{1}_{s_t > \bar{S}} \\
V_t(s_t, q_t) &= \pi^* \cdot \left(1 - \kappa^{T-t}\right) \left(s_t - \bar{S}\right) + \sum_{n=1}^{T-t-1} \left(1 - \kappa^n\right) \\
&\quad \cdot \left\{\pi^*_{T-n} \cdot \left(-\frac{\nu_n}{\sqrt{2\pi}} e^{-\frac{m_n^2}{2\nu_n^2}} + m_n \Phi \left(-\frac{m_n}{\nu_n}\right)\right) + \pi^*_{T-n} \cdot \left(\frac{\nu_n}{\sqrt{2\pi}} e^{-\frac{m_n^2}{2\nu_n^2}} + m_n \Phi \left(\frac{m_n}{\nu_n}\right)\right)\right\} \\
&\quad + q_t \left(1 - \kappa^{T-t}\right) \bar{S} + \kappa^{T-t} s_t,\end{align*}
\]

where \(m_n = \kappa^{T-t-n} (s_t - \bar{S})\), \(\nu_n^2 = \sigma^2 \sum_{j=0}^{n-1} \kappa^{2j}\) and \(\Phi\) denotes the cumulative Gaussian distribution function. To prevent confusion, we denote the circle number by \(\bar{n}\).

**Proof:** Again, the condition \(Q_{\text{max}} \leq Q_{\text{tot}} \leq (T + 1) \cdot Q_{\text{max}}\) ensures that \(\Pi_t \neq \emptyset\), for \(t = 0, \ldots, T\), and the assumptions of Theorem 5.2.1 are fulfilled. Hence, optimal solutions exist. We show the validity of the recursive formulas for \(\pi^*_t\) and \(V_t\), respectively. As in the proof of Proposition 5.2.5, we consider \(s_t\) and \(q_t\) to be fixed in each maximization step. The value function at time \(T\) is as in the proof of Proposition 5.2.5. At time \(T-1\) we have
\[ V_{T-1}(s_{T-1}, q_{T-1}) = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \{ s_{T-1} a_{T-1} + E_{T-1}\left[(1 - \kappa)\bar{S} + \kappa s_{T-1} + \varepsilon_{T-1}, q_{T-1} - a_{T-1}\right]\} \]

\[ = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \{ a_{T-1}(1 - \kappa)\left(s_{T-1} - \bar{S}\right)\} \]
\[ + q_{T-1}\left((1 - \kappa)\bar{S} + \kappa s_{T-1}\right). \]

The optimal strategy is consequently \( \pi^*_T = \pi^\text{min}_{T-1}1_{s_{T-1} \leq \bar{S}} + \pi^\text{max}_{T-1}1_{s_{T-1} > \bar{S}} \). At time \( T - 2 \) we have

\[ V_{T-2}(s_{T-2}, q_{T-2}) = \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \{ s_{T-2} a_{T-2} + E_{T-2}\left[(1 - \kappa)\bar{S} + \kappa s_{T-2} + \varepsilon_{T-2} - a_{T-2}\right]\} \]

\[ = \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ a_{T-2}(1 - \kappa)\left(1 + \kappa\right)\left(s_{T-2} - \bar{S}\right) \right. \]
\[ + \left. (1 - \kappa)\pi^\text{min}_{T-1} E_{T-2}\left[\left(\kappa (s_{T-2} - \bar{S}) + \varepsilon_{T-1}\right)1_{\kappa (s_{T-2} - \bar{S}) + \varepsilon_{T-1} \leq 0}\right]\right\} \]
\[ + \left. (1 - \kappa)\pi^\text{max}_{T-1} E_{T-2}\left[\left(\kappa (s_{T-2} - \bar{S}) + \varepsilon_{T-1}\right)1_{\kappa (s_{T-2} - \bar{S}) + \varepsilon_{T-1} > 0}\right]\right\} \]
\[ + q_{T-2}\left((1 - \kappa^2)\bar{S} + \kappa^2 s_{T-1}\right). \]

Now, we can factor out \((1 - \kappa)\) and consider the remaining expression as a linear function in \( a_{T-2} \): \( h(a_{T-2}) := c_1 a_{T-2} + c_2 \pi^\text{min}_{T-1} + c_3 \pi^\text{max}_{T-1} \). \( \pi^\text{min}_{T-1} \) and \( \pi^\text{max}_{T-1} \) are linear and decreasing in \( a_{T-2} \) with slope \( -1 \) or \( 0 \), compare Figure 5.1. Furthermore, \( c_2 < 0 \) and \( c_3 > 0 \). For \( s_{T-2} > \bar{S} \) we obtain \( 0 < c_2 < c_3 < \kappa (s_{T-2} - \bar{S}) < s_{T-2} - \bar{S} \). Consequently, \( h \) is increasing in \( a_{T-2} \) and we choose the maximum possible value in \( \Pi_{T-2}(q_{T-2}) \) as the optimal control at time \( T - 2 \): \( \pi^*_T = \pi^\text{max}_{T-1} \). Similarly, for \( s_{T-2} < \bar{S} \) we obtain \( 0 < c_3 < \kappa |s_{T-2} - \bar{S}| < |c_2| \) and the linear function \( h \) is decreasing in \( a_{T-2} \), i.e. we choose the minimum value in \( \Pi_{T-2}(q_{T-2}) \) as the optimal control at time \( T - 2 \): \( \pi^*_T = \pi^\text{min}_{T-1} \). Assume the validity of the form of the optimal strategy and value function a time \( t \). By the same reasoning as at time \( T - 2 \) it follows directly that for \( s_{t-1} < \bar{S}, \pi^\text{min}_{t-1} \) and for \( s_{t-1} > \bar{S}, \pi^\text{max}_{t-1} \) maximize the value function \( V_{t-1}\pi_{t-1}(s_{t-1}, q_{t-1}) \).

The explicit form of \( V_t \) follows from the fact that for any Gaussian random variable \( X \sim \mathcal{N}(\mu, \sigma^2) \) we can calculate
\[ \mathbb{E}[X \cdot 1_{X \leq 0}] = -\frac{\sigma}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}} + \mu \Phi \left( -\frac{\mu}{\sigma} \right), \]

and

\[ \mathbb{E}[X \cdot 1_{X > 0}] = \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}} + \mu \Phi \left( \frac{\mu}{\sigma} \right). \]

This yields the specific form of \( V_t(s_t, q_t) \).

Writing the optimal strategy in terms of \( t' \) (defined in (5.9)) is more involved than in the random walk case. When the spread price is quite in favor of the trader (i.e. \( s_t > \bar{S} \)), already in the beginning of the liquidation period it could happen that \( t' \) does not exist since the open position is already small enough to meet \( q_{T+1} = 0 \). In this case, the optimal strategy simplifies to

\[ \pi^*_t = \min \{ q_t, Q_{\max} \} 1_{s_t > \bar{S}}. \]

When it comes to a situation where \( \exists t' \) with \( t' \leq T \), we can write

\[ \pi^*_0 = \min \{ q_0, Q_{\max} \} 1_{s_0 > \bar{S}} \]

\[ \ldots \]

\[ \pi^*_{t'-1} = \min \{ q_{t'-1}, Q_{\max} \} 1_{s_{t'-1} > \bar{S}} \]

\[ \pi^*_t = (q_{t'} - (T - t')Q_{\max})^+ 1_{s_{t'} \leq \bar{S}} + \min \{ q_{t'}, Q_{\max} \} 1_{s_{t'} > \bar{S}} \]

\[ \ldots \]

\[ \pi^*_T = q_T. \]

In other words, the trader sells at little as possible as long as the spread price is below the mean-reversion level. Whenever the spread price raises above the mean-reversion level, he speeds up liquidating the maximum possible quantity \( Q_{\max} \). Note that the last proposition again relates nicely the results from Theorem 5.2.1 and Proposition 5.2.3. Another astonishing aspect is in contrast to Bertsimas’ [17] comments on the difficulty of imposing constraints in dynamic programs. We can find settings in which explicit optimal solutions \( \pi^*_t \) as well as explicit expressions for the value function \( V_t \) can be derived. The explicit form of the optimal strategies in Proposition 5.2.5 allows us to calculate the expected terminal wealth \( \mathbb{E}_0 [W_{T+1}] \) as stated in the following corollary.

**Corollary 5.2.8** Under the settings of Proposition 5.2.5, \( W_{T+1} \) is normally distributed and we obtain for the expected terminal wealth at time 0:

- If \( \mu \leq 0 \), then

\[ \mathbb{E}_0 [W_{T+1}] = (Q_{\text{tot mod } Q_{\max}}) \cdot (s_0 + t'' \mu) + Q_{\max} \left( (t'' + 1)s_0 + \mu \frac{t''(t'' + 1)}{2} \right). \]
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES

- If \( \mu > 0 \), then
  \[
  \mathbb{E}_0 [W_{T+1}] = (Q_{\text{tot}} \mod Q_{\text{max}}) \cdot (s_0 + t'\mu)
  + Q_{\text{max}} \left( (T - t')s_0 + \mu \frac{T(T + 1) - t'(t' + 1)}{2} \right).
  \]
  
  \( t' \) and \( t'' \) are defined in (5.9) and (5.10), respectively.

**Proof:** Under the random walk specifications, \( W_{T+1} \) is normally distributed since \( \pi_t^* \) is deterministic and the accumulated wealth is merely the sum of Gaussian random variables:

\[
W_{T+1} = \sum_{t=0}^{T} S_t \pi_t^*
\]

\[
= \sum_{t=0}^{T} \left( (1 - \kappa^t) \bar{S} + \kappa^t s_0 + \sum_{n=0}^{t-1} \kappa^n \varepsilon_{t-n} \right) \pi_t^*
\]

\[
= \sum_{t=0}^{T} \left( (1 - \kappa^t) \bar{S} + \kappa^t s_0 \right) \pi_t^* + \sum_{t=1}^{T} \sum_{n=0}^{t-1} \pi_t^* \kappa^n \varepsilon_{t-n}.
\]

The convolution property of the Normal distribution and the uncorrelated shocks \( \varepsilon_t \) make \( W_{T+1} \) Gaussian. The expected accumulated wealth for the random walk is easy to comprehend. Let \( \mu > 0 \) and assume for simplicity \( Q_{\text{tot}} \mod Q_{\text{max}} > 0 \):

\[
\mathbb{E}_0 [W_{T+1}] = \mathbb{E}_0 \left[ \sum_{t=0}^{T} S_t (Q_t - (T - t)Q_{\text{max}})^+ \right]
\]

\[
= \sum_{t=0}^{T} (s_0 + t\mu)(q_t - (T - t)Q_{\text{max}})^+
\]

\[
= \sum_{t=t'}^{T} (s_0 + t\mu)Q_{\text{max}} + (S_0 + t\mu')(Q_{\text{tot}} \mod Q_{\text{max}})
\]

\[
= (Q_{\text{tot}} \mod Q_{\text{max}}) \cdot (s_0 + t'\mu) + Q_{\text{max}} \left( (T - t')s_0 + \mu \frac{T(T + 1) - t'(t' + 1)}{2} \right).
\]

Unfortunately, the distribution and expectation of the terminal wealth \( W_{T+1} \) under mean-reverting dynamics is quite involved, since the decision at some future point in time depends on the future spread price through indicator functions. This makes the calculation of basic moments infeasible. The simple results in Propositions 5.2.5 and 5.2.7 already indicate that a dynamic strategy might induce a completely different liquidation behavior in comparison to a static setting. While under a simple random walk
there is no different behavior (i.e., the dynamic optimal strategy coincides with the static strategy, compare Section 4.3), the mean-reverting process takes advantage of the future information whether or not the current price is above or below the mean-reversion level $\bar{S}$. The true reason for this behavior lies even deeper. A mean-reverting process induces a local drift $\mu_t = \kappa (\bar{S} - s_t)$ (see Appendix, Section A.1 for more details) which depends on future prices and requires the knowledge of the mean-reversion level $\bar{S}$.

5.2.3 Markets with Liquidity Risk

In the previous section we demonstrated that even in simple settings, liquidation strategies and optimal value functions can become quite involved. We continue with this discussion and introduce the concept of illiquidity as it is a highly relevant factor when it comes to portfolio liquidation in energy markets.

In basically all of the academic literature, liquidity is defined in terms of the bid-ask spread or in terms of transaction costs, which come up when trading an asset. For instance, this notion of liquidity is discussed in Glosten & Milgrom [64] and Amihud & Mendelson [8]. Illiquidity is the situation where investors face higher trading and execution costs than at other times or in other markets. In this view, an investor can trade whenever he likes, but he might has to accept high costs. From a practitioners point of view, however, a somewhat different meaning is often attached to the term illiquidity. Utilities view illiquidity as the situation where their ability to trade commodities is limited or restricted. In extreme situations, illiquidity may be so severe that commodity markets temporarily disappear. For instance, forward markets encounter this phenomenon often in the beginning of a trading period of a specific contract. This type of illiquidity has more to do with the quantity of trades that can be executed. This notion of thinly traded commodities is difficult to reconcile with the standard economic view that there should be an equilibrium market price at which any desired quantity can be traded. Whatever the underlying reason for thin markets is, there are markets in which investors can find themselves in the situation where they are not able to trade as much as they like and cannot unwind large positions instantly. Although this type of illiquidity appears to be an important factor in financial market, it has not yet received much attention in the academic literature.

The quantity $Q_{max}$ can be interpreted as the maximum quantity of electricity that can be liquidated in the market without having any impact on the market price or without changing the temporary bid-ask spread noticeably. This interpretation is in line with the competitive market paradigm. The underlying assumption is that traders are price takers, i.e. regardless of how many assets they buy or sell, they do not move the price by any quantity impact. In other words, the markets are assumed to be perfectly elastic. This assumption is, even though not measurable properly, not satisfied even in the most liquid markets. The absence of this condition is sometimes labeled liquidity risk, see e.g. Jarrow [73]. To examine the impact of liquidity risk, we relax the constraint on $\pi_t$ being bounded and allow $\pi_t \in \mathbb{R}$, for $t = 0, \ldots, T$. This yields admissible sets $\Pi_t(q_t)$ as defined
in (3.7), which are clearly not compact. We examine a temporary market impact effect by assuming that the trader is not able to obtain the current market price $S_t$ at time $t$, but a slightly smaller price $\tilde{S}_t$. The obvious interpretation is that there is not enough market liquidity, or too few counterparts, a realistic scenario in energy markets. Thus, the trader has to lower his ask price (i.e. we have a quantity effect on the market price). More specifically, we assume the following relation:

$$\tilde{S}_t := S_t - \theta \pi_t,$$

for some constant $\theta > 0$, referred to as the price impact factor. This factor can also be time-dependent or even dependent on the current spread $S_t$. In some sense, we replaced the constraint $0 \leq \pi_t \leq Q_{\text{max}}$ by the impact specification (5.13) and let $\pi_t \in \mathbb{R}$, for $t = 0, ..., T$. Even though, the latter approach is conceptually different, the interpretation is the same. Other authors also introduced the notion of permanent price impact (see Almgren [5], [6], for a motivation and further discussion) where the quantity moves indeed the entire market price sustainably. This setup induces a tradeoff between selling fast and selling slowly. Selling fast reduces the price risk, because the trader locks in the prices at the earliest possible stages and is therefore no longer exposed to changing market conditions. On the other hand, liquidity risk is induced by selling/purchasing large asset positions too quickly which could either drive the prices into unfavorable directions or leads to too big bid-ask spreads, so the trader is forced to accept a lower price than noted in the market. Many researchers argue, see e.g. Chan & Lakonishok [34],[35], and Holthausen et al. [67], that temporary and permanent market impact occur simultaneously, thereby should be taken into account in one model.

Another aspect is important to notice. Liquidating the capacity of a coal-fired power plant involves basically three different counterparts. One that buys the electricity, one that sells the coal, and one that sells the emissions certificates. However, this does not change the setup as is easy to see: If there is an impact in all three markets (represented by the non-negative impact factors $\rho$, $\varrho$ and $\tau$), we can write

$$\tilde{F}_1(t, T_0) = F_1(t, T_0) - \rho \pi_t$$
$$\tilde{F}_2(t, T_0) = F_2(t, T_0) + \varrho \pi_t$$
$$\tilde{F}_3(t, T_0) = F_3(t, T_0) + \tau \pi_t.$$ 

Then, the spread becomes

$$\tilde{S}_t = \tilde{F}_1(t, T_0) - h_1 \tilde{F}_2(t, T_0) - h_2 \tilde{F}_3(t, T_0)$$
$$= F_1(t, T_0) - \rho \pi_t - h_1 (F_2(t, T_0) + \varrho \pi_t) - h_2 (F_3(t, T_0) + \tau \pi_t)$$
$$= S_t - (\rho + h_1 \varrho + b_2 \tau) \pi_t.$$
Consequently, we only require one of the parameters \( \rho \), \( \varrho \), or \( \tau \) to be non-negative. Two out of three can be zero. Consider again the problem (3.2) with \( \pi_t \in \mathbb{R}, t = 0, \ldots, T \). In this section, the one-period reward function and the terminal reward function become

\[
    r(s_t, a_t) = (s_t - \theta a_t) a_t, \\
    g(s_T, q_T) = (s_T - \theta q_T) q_T.
\]

We need to check whether the integrability condition \((I)\) holds. It suffices to show:

\[
    \mathbb{E}_0 [r(S_t, \pi_t(S_t, Q_t))] = \mathbb{E}_0 \left[ S_t \pi_t - \theta \pi_t^2 \right] \\
    = \mathbb{E}_0 \left[ -\theta \left( \pi_t^2 - \frac{S_t}{\theta} \pi_t + \left( \frac{S_t}{2\theta} \right)^2 - \left( \frac{S_t}{2\theta} \right)^2 \right) \right] \\
    = \mathbb{E}_0 \left[ -\theta \left( \pi_t - \frac{S_t}{2\theta} \right)^2 + \frac{S_t^2}{4\theta} \right] \\
    \leq \frac{1}{4\theta} \mathbb{E}_0 \left[ S_t^2 \right],
\]

and similarly

\[
    \mathbb{E}_0 [g(S_T, Q_T)] \leq \frac{1}{4\theta} \mathbb{E}_0 \left[ S_T^2 \right].
\]

The underlying dynamics have finite second moments. Since both functions, \( r \) and \( g \) are concave, quadratic functions in \( a_t \) it is not necessary to check the positive part of these functions. This implies that \((I)\) is fulfilled. Let \( S_t \) again be given by the random walk specification:

\[
    S_t = \mu + S_{t-1} + \varepsilon_t.
\]

\( \tilde{S}_t \) is then calculated from (5.13). \( \mu \) is the real-valued drift parameter and \( \varepsilon_t \) is a random variable following a Normal distribution with zero mean and constant volatility \( \sigma \). In the spirit of Theorem 5.2.1, we do not know about the existence of an optimal solution, since the compactness condition on \( \Pi_t \) is not fulfilled. However, the DP algorithm (Proposition 5.2.3) guarantees again the optimality of the resulting strategies. These are stated in the next proposition.

**Proposition 5.2.9** Consider the dynamics (5.13) and (5.14), respectively. Furthermore, assume that \( \theta > 0 \). The optimal strategy of the maximization problem (3.2) is unique and given explicitly by

\[
    \pi_t^* = \frac{Q_{t+1}}{T + 1} + \frac{\mu}{4\theta} \left( 2t - T \right),
\]

for \( t = 0, \ldots, T \). Furthermore, the optimal value function is given by
The value function at time \( T \) considers the recursive form for arbitrary predominant guess about their structure, we apply an induction argument to validate the recursive form for arbitrary \( t = 0, \ldots, T \). As in the proof of Proposition 5.2.5, we consider \( s_t \) and \( q_t \) to be fixed in each maximization step. The terminal value function is

\[
V_T(s_T, q_T) = g(s_T, q_T) = (s_T - \theta q_T) q_T.
\]

The value function at time \( T - 1 \) is given by

\[
V_{T-1}(s_{T-1}, q_{T-1}) = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ s_{T-1} a_{T-1} + E_{T-1} \left[ V_T(\mu + s_{T-1} + \varepsilon_{T-1}, q_{T-1} - a_{T-1}) \right] \right\}
\]

\[
= \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ \left. -2\theta a_{T-1}^2 + 2\theta a_{T-1} (2\theta q_{T-1} - \mu) \right| a_{T-1} = 1 \right\}.
\]

To maximize the last expression we take the first derivative of w.r.t. \( a_{T-1} \), which yields

\[-4\theta a_{T-1} + 2\theta q_{T-1} - \mu = 0.\]

The optimal strategy at time \( T - 1 \) becomes:

\[\pi^*_T = \frac{q_{T-1}}{2} - \frac{\mu}{4\theta}.\]

The second derivative is \(-4\theta < 0\), hence \(\pi^*_{T-1}\) is a maximizer. Now, plugging \(\pi^*_{T-1}\) into (5.17) yields

\[V_{T-1}(s_{T-1}, q_{T-1}) = q_{T-1} \left( s_{T-1} - \frac{\theta q_{T-1}}{2} + \frac{\mu}{2} \right) + \frac{\mu^2}{8\theta}.\]

Going one step back in time we arrive at the \( T - 2 \) value function

\[
V_{T-2}(s_{T-2}, q_{T-2}) = \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ s_{T-2} a_{T-2} + E_{T-2} \left[ V_{T-1}(\mu + s_{T-2} + \varepsilon_{T-1}, q_{T-2} - a_{T-2}) \right] \right\}
\]

\[
= \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ \left. \left. \frac{3}{2} \theta a_{T-2}^2 + 2\theta a_{T-2} \left( \theta q_{T-2} - \frac{3}{2}\mu \right) \right| a_{T-2} = 1 \right\}.
\]

Thus, the value function for the optimal strategy is given by

\[
\pi^*(s, q) = \begin{cases} s - \theta q & \text{if } s > \theta q + \frac{\mu}{2} \theta^2 - \frac{\mu^2}{8\theta}, \\ \theta q + \frac{\mu}{2} \theta^2 - \frac{\mu^2}{8\theta} & \text{if } s = \theta q + \frac{\mu}{2} \theta^2 - \frac{\mu^2}{8\theta}, \\ \theta q & \text{if } s < \theta q + \frac{\mu}{2} \theta^2 - \frac{\mu^2}{8\theta}. \end{cases}
\]
5.2. **SOLVING (P')**: **EXPLICIT OPTIMAL LIQUIDATION STRATEGIES**

The quantity that maximizes (5.18) is given by

\[ \pi^*_T = \frac{q_T - 2}{3} - \frac{\mu}{2\theta}, \]

and

\[ V_T (s_T, q_T) = q_T \left( s_T - \frac{\theta q_T}{3} + \mu \right) + \frac{\mu^2}{2\theta}. \]

Proceeding in this fashion for the value function at time \( T-3 \), we find

\[ V_{T-3} (s_{T-3}, q_{T-3}) = q_{T-3} \left( s_{T-3} - \frac{\theta q_{T-3}}{4} + \frac{3}{2} \mu \right) + \frac{5}{4} \frac{\mu^2}{\theta}, \]

and for the optimal strategy

\[ \pi^*_{T-3} = \frac{q_{T-3} - 2}{4} - \frac{3\mu}{4\theta}. \]

In order to find the general recursive equation for \( \pi^*_t \) and \( V_t \), we have to make a guess on how the fractions \((i)-(v)\) evolve as we go further back in time. For \( (i) \) we clearly have \( \frac{\theta q}{T-t+1} \). Similarly, for \( (iv) \) we have \( \frac{\theta q}{T-t+1} \). \( (ii) \) might evolve according to \( \frac{T}{2} \). \( (iii) \) is slightly more involved. We calculate the first four value functions for \( T, ..., T-3 \) and observe a sequence \( \frac{1}{8}, \frac{1}{8}, \frac{3}{8}, \frac{10}{8} \). A predominant guess on the numerator might be \( \frac{1}{6} (T-t)(T-t+1)(T-t+2) \). Finally, \( (v) \) might evolve according to \( \frac{(T-t)\mu}{4\theta} \). These guesses give us the form

\[ \pi^*_t = \frac{qt}{T-t+1} - \frac{(T-t)\mu}{4\theta}, \]

for the optimal strategy and (5.16) for the corresponding value function, respectively.

To prove their validity, we need to show that this recursion does also hold for time \( t-1 \). We obtain immediately

\[ V_{t-1} (s_{t-1}, q_{t-1}) = \max_{a_{t-1} \in \Pi_{t-1}(q_{t-1})} \left\{ s_{t-1} - \frac{\theta q_{t-1}}{T-t+1} + \frac{\mu}{2} \right\} + \frac{\mu^2}{8\theta} \delta_{t-1}. \]

The formula for \( \pi^*_t \) can further be simplified by using the equation \( q_t = q_0 - \pi^*_0 - ... - \pi^*_{t-1} \).

This gives us (5.15). Last but not least, we verify

\[ \sum_{t=0}^{T} \pi^*_t = Q_{tot}. \]
This completes the proof.

The optimal strategy is deterministic and for \( \mu = 0 \) we obtain \( \pi_t^* = \frac{Q_{tot}}{T+1} \), for \( t = 0, ..., T \). In other words, the strategy becomes stationary and independent of the current state \((s_t, q_t)\) at time \( t \). Moreover, for arbitrary \( \mu \neq 0 \) the liquidation behavior is symmetric with respect to the half time \( t = \frac{T}{2} \) (assuming \( T \) is even). For instance, when \( \mu > 0 \), then \( \pi_t^* < \frac{Q_{tot}}{T+1} \) for \( t = 0, ..., \frac{T}{2} \), i.e. the trader sells less than the average quantity per week. For \( t > \frac{T}{2} \) he compensates the little quantities by speeding up the liquidation. This is due to the fact that the expected spread price is higher in the far future and the linear and temporary price impact induces an exactly symmetric behavior. Furthermore, \( \pi_t^* \) can become negative, meaning that short-selling is allowed. That means, the trader has to buy back some of the contracts before maturity in order to meet \( q_{T+1} = 0 \). However, in the spirit of (5.13) the trader can charge a higher price than the market price, which in turn requires counterparts who face a similar situation as the trader himself does (under the condition that \( \tau, \rho > 0 \)). Because, he can claim a higher bid price and the model suggests that there is indeed someone who pays this price. Figure 5.2 shows how a trader would liquidate the portfolio for different values of the drift parameter \( \mu \).

Figure 5.2: Portfolio position \( q_t \) in dependence of the drift parameter \( \mu \) for the random walk model. The other parameters are \( Q_{tot} = 500 \text{ MW} \) and \( \theta = 0.05 \).

Comparing this static strategy with the static strategy in Chapter 4 leads to the insight that a trader deviates from, e.g. selling as slowly as possible (i.e. \( \mu > 0 \) in the static case) by choosing the optimal tradeoff between the price impact represented by the parameter \( \theta \) and the waiving of potentially higher revenues from selling as slowly as
5.2. SOLVING \((P')\): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES

possible.

We are always interested in the terminal wealth distribution under the optimal strategies. In this case, \(W_{T+1}\) is a Gaussian random variable (by the same arguments as in the proof of Corollary 5.2.8) and we readily obtain

\[
E_0 [W_{T+1}] = \underbrace{q_0 s_0} + T \mu \frac{q_0}{2} + \frac{\mu^2}{24 \theta} T (T + 1) (T + 2).
\]

In other words, the expected terminal wealth is composed of three parts. The initial portfolio, an additional factor that can be either negative \((\mu < 0)\) or positive \((\mu > 0)\), and a third contributing factor which is always positive. This last factor is higher when \(\theta\) is smaller, i.e. when the cost of trading is lower. Furthermore, the second factor reduces the overall expected terminal wealth when the drift is negative. This interpretation is very intuitive and highlights the usefulness of closed-form formulae.

The second process class of interest is the autoregressive process. Let \(S_t\) again be given by

\[
S_t = (1 - \kappa) \bar{S} + \kappa S_{t-1} + \varepsilon_t.
\]

(5.19)

\(\tilde{S}_t\) is then again calculated from (5.13). As long as \(\kappa \in (0, 1)\) we have the discrete-time version of an arithmetic Ornstein-Uhlenbeck process (see Appendix A.1). For \(\kappa = 1\) we obtain the random walk specification (5.14) with \(\mu = 0\). \(\varepsilon_t\) is again a random variable following a Gaussian distribution with zero mean and constant volatility \(\sigma\). And again, in the spirit of Theorem 5.2.1 we do not know about the existence of an optimal solution, since \(\Pi_t\) is not compact. However, the DP algorithm (Proposition 5.2.3) guarantees the optimality of the resulting strategies. These are stated in the next proposition.

**Proposition 5.2.10** Consider the dynamics (5.19) and (5.13), respectively. Furthermore, assume that \(\theta > 0\). The optimal strategy of the maximization problem (3.2) is unique and given explicitly by

\[
\pi_t^* = \delta_t^{(1)} \left( s_t - \bar{S} \right) + \frac{q_t}{T - t + 1},
\]

(5.20)

where

\[
\delta_t^{(1)} = \frac{5 \cdot 2^{T-t-2} - 1}{3 \cdot 2^{T-t} \theta} \left( 1 - \frac{\kappa}{T - t} \left( \frac{\kappa^{T-t} - 1}{\kappa - 1} \right) \right),
\]

for \(t = 0, ..., T\). Furthermore, the optimal value function is given by
\[ V_t(s_t, q_t) = \delta^{(2)}_{T-t} \left( s_t - \bar{S} \right)^2 + \sigma^2 \sum_{n=1}^{T-t-1} \delta^{(2)}_n + \frac{q_t}{T-t+1} \left[ s_t \left( \frac{k^{T-t+1} - 1}{k-1} \right) \right. \\
\left. + \bar{S} \left( T - t - \kappa \left( \frac{k^{T-t} - 1}{k-1} \right) \right) - \theta q_t \right], \tag{5.21} \]

where

\[ \delta^{(2)}_0 = 0, \]
\[ \delta^{(2)}_1 = \frac{(1 - \kappa)^2}{8\theta}, \]
\[ \delta^{(2)}_{T-t} = \kappa^2 \delta^{(2)}_{T-t-1} + \frac{T-t}{4\theta(T-t+1)} \left( 1 - \kappa \left( \frac{k^{T-t} - 1}{k-1} \right) \right). \]

**Proof:** As in the proof of Proposition 5.2.5, we consider \( s_t \) and \( q_t \) to be fixed in each maximization step. The terminal value function is the same as in the proof of Proposition 5.2.9:

\[ V_T(s_T, q_T) = (s_T - \theta q_T) q_T. \]

The value function at time \( T - 1 \) is given by

\[ V_{T-1}(s_{T-1}, q_{T-1}) = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ (s_{T-1} - \theta a_{T-1}) a_{T-1} \right\} + \mathbb{E}_{T-1} \left[ V_T \left( (1 - \kappa)\bar{S} + \kappa s_{T-1} + \varepsilon_T, q_{T-1} - \pi_{T-1} \right) \right]. \tag{5.22} \]

\[ = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ (s_{T-1} - \theta a_{T-1}) a_{T-1} \right\} + \left( (1 - \kappa)\bar{S} + \kappa s_{T-1} - \theta (q_{T-1} - a_{T-1}) \right) (q_{T-1} - a_{T-1}) \tag{5.23} \]

\[ = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ -2\theta a_{T-1}^2 + a_{T-1} \left( (1 - \kappa) \left( s_{T-1} - \bar{S} \right) + 2\theta q_{T-1} \right) \right\} \]

\[ q_{T-1} \left( \kappa s_{T-1} + (1 - \kappa)\bar{S} - \theta q_{T-1} \right). \tag{5.24} \]

Taking the first derivative w.r.t. \( a_{T-1} \) yields

\[ -4\theta a_{T-1} + (1 - \kappa) \left( s_{T-1} - \bar{S} \right) + 2\theta q_{T-1} = 0. \]

Solving this last equation gives the optimal strategy at time \( T - 1 \):

\[ \pi^*_T = \frac{1}{4\theta} \left( s_{T-1} - \bar{S} \right) + \frac{q_{T-1}}{2}. \]

The second derivative is \( -4\theta < 0 \), i.e. \( \pi^*_T \) is a maximizer. Now, plugging \( \pi^*_T \) into (5.24) yields
5.2. SOLVING ($P'$): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES

\[ V_{T-1}(s_{T-1}, q_{T-1}) = \frac{(1 - \kappa)^2}{8\theta} \left( s_{T-1} - \bar{S} \right)^2 + \frac{q_{T-1}}{2} \left( s_{T-1}(1 + \kappa) + \bar{S}(1 - \kappa) - \theta q_{T-1} \right). \]

Going one step back in time we arrive at the $T - 2$ value function

\[ V_{T-2}(s_{T-2}, q_{T-2}) = \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ (s_{T-2} - \theta a_{T-2}) a_{T-2} \right. \]
\[ + E_{T-2} \left[ V_{T-1} \left( (1 - \kappa)\bar{S} + \kappa s_{T-2} + \varepsilon_{T-1}, q_{T-2} - a_{T-2} \right) \right] \}
\[ = \max_{a_{T-2} \in \Pi_{T-2}(q_{T-2})} \left\{ -\frac{3}{2} \theta a_{T-2}^2 \right. \]
\[ + a_{T-2} \left( 1 - \frac{\kappa}{2} - \frac{\kappa^2}{2} \right) \left( s_{T-2} - \bar{S} \right) + \theta q_{T-2} \left\} \right. \]
\[ + q_{T-2} \left( \frac{\kappa}{2} + \frac{\kappa^2}{2} \right) s_{T-2} + \left( 1 - \frac{\kappa}{2} - \frac{\kappa^2}{2} \right) \bar{S} - \theta q_{T-2} \left. \right) \]
\[ + \frac{\kappa^2(1 - \kappa)^2}{8\theta} \left( s_{T-2} - \bar{S} \right)^2 + \delta_{1(2)}^2 \sigma^2. \]

The quantity that maximizes the last expression is given by

\[ \pi_{T-2}^* = \frac{1 - \frac{\kappa}{2} - \frac{\kappa^2}{2}}{3\theta} \left( s_{T-2} - \bar{S} \right) + \frac{q_{T-2}}{3}, \]

and

\[ V_{T-2}(s_{T-2}, q_{T-2}) = \frac{(1 - \kappa)^2(\kappa^2 + \kappa + 1)}{6\theta} \left( s_{T-2} - \bar{S} \right)^2 + \delta_{1(2)}^2 \sigma^2 \]
\[ = \delta_{1(2)}^2 \]
\[ + \frac{q_{T-2}}{3} \left( s_{T-2}(1 + \kappa + \kappa^2) + \bar{S}(2 - \kappa - \kappa^2) - \theta q_{T-2} \right). \]

Proceeding in this fashion for the value function at time $T - 3$, we find
\[ V_{T-3} (s_{T-3}, q_{T-3}) = \max_{a_{T-3} \in \Pi_{T-3}(q_{T-3})} \left\{ -\frac{4}{3} \theta a_{T-3}^2 + a_{T-3} \left( 1 - \frac{\kappa}{3} - \frac{\kappa^2}{3} - \frac{\kappa^3}{3} \right) (s_{T-3} - \bar{S}) + \frac{2\theta}{3} q_{T-3} \right\} \]

\[ + q_{T-3} \left( \left( \frac{\kappa}{3} + \frac{\kappa^2}{3} + \frac{\kappa^3}{3} \right) s_{T-3} \right) \]

\[ + \left( 1 - \frac{\kappa}{3} - \frac{\kappa^2}{3} - \frac{\kappa^3}{3} \right) \bar{S} - \frac{\theta q_{T-3}}{3} \]

\[ + \kappa^2 \delta_2^{(2)} \left( s_{T-3} - \bar{S} \right)^2 + \left( \delta_1^{(2)} + \delta_2^{(2)} \right) \sigma^2, \]

and for the optimal strategy

\[ \pi_{T-3} = \frac{1}{\theta} \left( \frac{3}{8} \left( 1 - \frac{\kappa}{3} - \frac{\kappa^2}{3} - \frac{\kappa^3}{3} \right) (s_{T-3} - \bar{S}) + \frac{q_{T-3}}{4} \right). \]

Consequently,

\[ V_{T-3} (s_{T-3}, q_{T-3}) = \left( \kappa^2 \delta_2^{(2)} + \frac{16}{3\theta} \left( 1 - \frac{\kappa}{3} - \frac{\kappa^2}{3} - \frac{\kappa^3}{3} \right)^2 \right) \left( s_{T-3} - \bar{S} \right)^2 \]

\[ + \frac{q_{T-3}}{4} \left( s_{T-3} (1 + \kappa + \kappa^2 + \kappa^3) + \bar{S} (3 - \kappa - \kappa^2 - \kappa^3) - \theta q_{T-3} \right) \]

\[ + \left( \delta_1^{(2)} + \delta_2^{(2)} \right) \sigma^2. \]

In order to find the general recursive equation for \( \pi_t \) and \( V_t \) we have to make guess about the coefficients \( \delta_k^{(2)} \), (i) and (ii). Clearly, the denominator of \( q_{T-k} \) evolves according to \( k + 1 \). Moreover, (ii) has the form

\[ 1 - \frac{\kappa}{k} - ... - \frac{\kappa^k}{k} = 1 - \frac{1}{k} \sum_{i=1}^{k} \kappa^i \]

\[ = 1 - \frac{1}{k} \left( \frac{\kappa^{k+1} - \kappa}{\kappa - 1} \right) \]

(i) can be recovered as follows. It evolves according to
5.2. SOLVING \((P')\): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES

\[
\begin{align*}
T - 1 : & \quad \frac{1.5}{6} \\
T - 2 : & \quad \frac{4}{12} \\
T - 3 : & \quad \frac{9}{24}
\end{align*}
\]

In other words, the nominator has the recursive form \(c_k = 2c_{k-1} + 1, \ c_0 = \frac{1}{4}\) and the denominator has the form \(d_k = 2d_{k-1}, \ d_0 = 3\). This fraction can be further simplified by recognizing its geometric behavior. Consequently, for \(k \geq 0\) we obtain \(c_k = 5 \cdot 2^{k-2} - 1\) and \(d_k = 3 \cdot 2^k\). Writing these quantities in forward time, e.g. \(\pi^*_t\) for \(t = 0, \ldots, T\) and not \(\pi^*_{T-k}\) for \(k = T, \ldots, 0\) yields

\[
\pi^*_t = \frac{ct}{dt} \left( 1 - \frac{\kappa}{T - t} \left( \frac{\kappa^{T-t} - 1}{\kappa - 1} \right) \right) (s_t - \bar{S}) + \frac{qt}{T - t + 1},
\]

and (5.21). To complete the proof we need to show first that (5.20) and (5.21) hold for arbitrary \(t = 0, \ldots, T\). This is a straightforward induction proof, which we omit here. And secondly, the DP algorithm yields the optimal strategy and value function as mentioned in Proposition 5.2.3. This completes the proof.

**Remark 5.2.11** The expression for \(\pi^*_t\) can be stated in a different form. Using the relation \(q_t = q_{t-1} - \pi^*_{t-1}\) yields

\[
\pi^*_t = \delta^{(1)}_t (s_t - \bar{S}) - \sum_{n=0}^{t-1} \frac{\delta^{(1)}_n}{T - n} (s_n - \bar{S}) + \frac{Q_{tot}}{T + 1}
\]

This control is obviously not Markovian as it depends on all past realized spread prices.

The fraction \(\frac{5 \cdot 2^{T-t-2} - 1}{3 \cdot 2^T} \) is for \(t = 0, \ldots, T\) a monotonically decreasing sequence. For \(T \to \infty\) we can see where it starts (set \(t = 0\)):

\[
\frac{5 \cdot 2^{T-2} - 1}{3 \cdot 2^T} = \frac{\frac{5}{3} - \frac{2^{-T}}{3}}{3} \to \frac{5}{12}, \quad \text{as } T \to \infty.
\]

For \(t \to T\) and \(T\) fixed, we obtain

\[
\frac{5 \cdot 2^{-2} - 1}{3} = \frac{1}{12}.
\]

Furthermore, \(1 - \frac{\kappa}{T-t} \left( \frac{\kappa^{T-t} - 1}{\kappa - 1} \right)\) is for \(t = 0, \ldots, T\) also a monotonically decreasing sequence. For \(T \to \infty\) we can see where it starts (set \(t = 0\)):

\[
1 - \frac{\kappa}{T} \left( \frac{\kappa^{T} - 1}{\kappa - 1} \right) \to 1, \quad \text{as } T \to \infty.
\]
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES

For $t \to T$ and $T$ fixed, we obtain with l’Hospital zero in the limit. To cut a long story short, the coefficient $\delta_i^{(1)}$ of $s_t - \bar{S}$ in (5.20) is monotonically decreasing as $t \to T$ and evolves roughly between $\frac{5}{12}$ and 0. In other words, this coefficient is well-behaved in some sense and always positive. The optimal liquidation strategy (5.20) consists of two parts, the first one can be positive or negative, depending on if the current price is above or below the mean-reversion level $\bar{S}$. For instance, at time 0 we liquidate more than $\frac{Q_{ot}}{T+1}$ if the current price $s_0$ is above the mean-reversion level $\bar{S}$ and we liquidate less than $\frac{Q_{ot}}{T+1}$ if the current price $s_0$ is below the mean-reversion level $\bar{S}$. This argument does also hold for an arbitrary point in time $t$ and shows that the mean-reverting character of the underlying dynamics induces the utilization of expected favorable situations: When prices are below $\bar{S}$, then, knowing the current price one can expect the prices to raise, so the trader prefers to hold back some capacity that he can liquidate at a later point in time when the prices have reverted. When prices are above $\bar{S}$, he speeds up trading to secure the earnings. This behavior is similar to an aggressive-in-the-money behavior, see Kissel [83].

For $\kappa \to 1$, we have a model similar to that introduced by Bertsimas & Lo [17] for the single-asset case and the resulting optimal strategy is simple and stationary, as stated in the next corollary.

**Corollary 5.2.12** For $\kappa = 1$, we obtain

$$\pi_t^* = \frac{Q_{ot}}{T+1}$$

$$V_t(s_t, q_t) = q_t \left( s_t - \frac{\theta q_t}{T-t+1} \right),$$

for $t = 0, \ldots, T$.

**Proof:** Applying l’Hospital’s rule to (5.20) and (5.21) yields the desired result. ■

In other words, the linear price impact (independent of $\theta$, as long as $\theta > 0$) forces the trader not to sell everything right in the beginning in one block, but to distribute the entire position evenly over the entire liquidation period. This stems from the fact that the price impact is independent of the actual spread price $s_t$ and of the remaining quantity $q_t$. However, since each period’s reward $r(s_t, a_t) = (s_t - \theta a_t) a_t$ is a concave and quadratic function of the action $a_t$, the sum of these single-period execution returns is maximized at the point where the marginal execution returns are equated across all periods. There is no advantage to shifting trades to one period or another. They all offer the same trade-offs to the objective function. Hence, the trade sizes should be set equal across all periods. The same result can be obtained if we set $\mu = 0$ in Proposition 5.2.9.

For $\kappa < 1$ and when $\theta$ becomes very large the trader is forced not to deviate from the static strategy (5.25), because speeding up (i.e. when $s_t > \bar{S}$) is too costly and slowing
down (i.e. when \( s_t < \bar{S} \)) will be costly in the future. Because, in order to meet \( q_{T+1} \), the trader has to sell more than in (5.25) at least at one future point in time. As \( \theta \to 0 \), the trader is more and more tempted to exploit favorable situations, because the price impact vanishes. This leads to a very aggressive and active trading scheme consisting of liquidating and even repurchasing electricity in large amounts. The trader has always the information of being either below or above the mean-reversion level and could therefore repurchase lots of electricity when the current price is far below the mean-reverting level and sell lots of electricity when the current price is above the mean-reverting level.

Since the spread \( S_t \) is modeled in a simple fashion, basically being normally distributed (see (5.19)), we can determine the distribution of the optimal strategy to gain further insight. Recall that \( \varepsilon_t \sim N(0, \sigma^2) \). The conditional distribution of the market price is given by

\[
S_t|s_0 \sim N \left( (1 - \kappa^t) \bar{S} + \kappa^t s_0, \sigma^2 \kappa^{2t} - 1 \right).
\]

Now going forward in time, we can deduce the distribution of the optimal strategy using (5.20). At time \( t = 0 \), we have

\[
\pi_0^* = \frac{Q_{T_0}}{T + 1} + (s_0 - \bar{S}) \delta_0^{(1)},
\]

i.e. \( \pi_0^* \) is always deterministic. For \( t > 0 \), \( \pi_t^* \) is random and follows a Gaussian distribution. The expected optimal strategy is given by

\[
E_0[\pi_t^*] = \delta_t^{(1)} (s_0 - \bar{S}) - \sum_{n=0}^{t-1} \frac{\delta_n^{(1)}}{T-n} E_0 \left[ S_n - \bar{S} \right] + \frac{Q_{T_0}}{T + 1} + (s_0 - \bar{S}) \left( \delta_t^{(1)} \kappa^t - \sum_{n=0}^{t-1} \frac{\delta_n^{(1)} \kappa^n}{T-n} \right).
\]

Since \( \pi_t^* \) are Gaussian random variables, they can take negative values with positive probability, i.e. buying back is absolutely possible for every reasonable parameter setting. Unfortunately, the distribution of \( W_{T+1} \) is highly complicated and unknown in closed form (see Figure 5.3), but we can calculate at least the first moment. We merely state the result, since the derivation is quite tedious, see the Appendix A.3.

\[
E_0[W_{T+1}] = \bar{S} \sum_{t=0}^{T} E_0[\pi_t^*] + \sum_{t=0}^{T} E_0[\pi_t^*] \cdot \kappa^t \cdot (s_0 - \bar{S})
+ \frac{\sigma^2}{\kappa^{2t} - 1} \sum_{t=1}^{T} \left( \delta_t^{(1)} (\kappa^{2t} - 1) - \sum_{n=1}^{t-1} \frac{\delta_n^{(1)} \kappa^{2n} - 1}{T-n} \right). \tag{5.26}
\]

In order to gain further insight we conduct a little numerical study.
Numerical Example

Consider a coal-fired power plant with a total capacity of $Q_{\text{tot}} = 500$ MW. We intend to sell this capacity in yearly forward contracts (to possibly several counterparts). These contracts are assumed to mature at the same time $T_0$. Recall that we have to buy the corresponding quantities of coal in order to be able to produce the electricity. The liquidation period is assumed to be one year long and we liquidate on a weekly basis, i.e. $t = 0, ..., 51 = T < T_0$. The initial spread price is $s_0 = 20$ EUR. The aim is to maximize the expected revenue over the entire liquidation horizon, given model parameters in Table 5.1.

<table>
<thead>
<tr>
<th>$\hat{S}$</th>
<th>$\sigma$</th>
<th>$\kappa$</th>
<th>$\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 EUR</td>
<td>1.0</td>
<td>0.25, 0.50, 0.75</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Table 5.1: Model and input parameters for the numerical example, granularity is weekly.

The parameters are the weekly-scaled quantities of Table 3.2. The initial portfolio position has a value of $Q_{\text{tot}} \times 8760 \text{ h} \times s_0 = 87,600,000$ EUR. The full-impact cost is

$$Q_{\text{tot}} \times Q_{\text{tot}} \times \theta \times 8760 \text{ h} = 10,950,000 \text{ EUR},$$

which is after all about 13% of the current portfolio value. This calculation requires some explanation. The electricity is sold in yearly forward contracts. Since the spread price is noted in EUR per MWh, we have to multiply the ‘per MWh’ revenue by MW and by 8760 hours (since $365 \times 24 = 8760$). The price impact can be seen as a unitless quantity (actually it is ‘MW$^2$h’) that reduces the EUR per MWh spread price. Thus, the realizable (assuming that complete liquidation is possible) total revenue is

$$(s_0 - \theta Q_{\text{tot}}) \times Q_{\text{tot}} \times 8760 \text{ h} = 76,650,000 \text{ EUR}.$$  

For a specific trade, say $a_t = 50$ MW on week $t$ at a market-quoted spread price of $s_t = 20$ EUR, the trader has to accept $\tilde{s}_t = s_t - \theta a_t = 19.75$ EUR per MWh. Therefore, the revenue is $19.75 \cdot 50 = 987.50$ EUR per hour or $987.50 \cdot 8760 = 8,650,500$ EUR for the entire year. The impact cost amount to $12.50$ EUR per MWh or $109,500$ EUR for the entire year. To introduce a benchmark for the optimal strategy $\pi^*$ we consider a suboptimal, naive strategy $\hat{\pi}_t = \frac{Q_{\text{tot}}}{T+1}, t = 0, ..., T$. Note that $\hat{\pi} = \pi^*$ for $\kappa = 1$ (cp. Corollary 5.2.12). We simulate 50,000 paths according to the spread price dynamics (5.19) and calculate the terminal wealth for each path under $\pi^*$ and $\hat{\pi}$, respectively. Averaging over all wealths yields the expected wealths for both strategies. Actually, the distribution of the terminal wealth under $\hat{\pi}$ is Gaussian with parameters
5.2. **SOLVING (P'): EXPLICIT OPTIMAL LIQUIDATION STRATEGIES**

\[
\mu_{\hat{\pi}} = \frac{Q_{\text{tot}}}{T+1} \left( \bar{S} \cdot (T + 1) + \frac{\kappa^{T+1} - 1}{\kappa - 1} (s_0 - \bar{S}) - \theta Q_{\text{tot}} \right) \tag{5.27}
\]

\[
\sigma^2_{\hat{\pi}} = \left( \frac{Q_{\text{tot}} \sigma}{T+1} \right)^2 \left( \frac{T (\kappa^2 - 1) + \kappa (\kappa^T - 1) (\kappa^{T+1} - \kappa - 2)}{(\kappa - 1)(\kappa^2 - 1)} \right)^2 \tag{5.28}
\]

The derivation can be found in the Appendix A.3.2. For 50,000 Monte Carlo simulations, the explicit and the approximate expected value of \(W_{T+1}\) under \(\hat{\pi}\) have a relative error of about 0.03%. The error for the volatility is about 0.01%. Figure 5.3 shows the histograms (using a Normal kernel smoother) of the terminal wealth under \(\pi^*\) and \(\hat{\pi}\), respectively.

![Histograms of terminal wealth](image)

**Figure 5.3:** Distribution of terminal wealth \(W_{T+1}\) under the optimal strategy \(\pi^*\) (blue area) and naive strategy \(\hat{\pi}\) (green area). Model parameters are \(s_0 = 20\) EUR, \(\bar{S} = 20\) EUR, \(\theta = 0.005\), \(Q_{\text{tot}} = 500\) MW, \(T = 51\), Monte Carlo iterations: 50,000. Left graph: \(\kappa = 0.25\), middle graph: \(\kappa = 0.5\), right graph: \(\kappa = 0.75\).

They indicate that \(\hat{\pi}\) is a suboptimal strategy. Changing the speed of mean-reversion \(\kappa \in \{0.25, 0.50, 0.75\}\) does not yield a different expected terminal wealth, always about 87,400,000 EUR. However, the volatility increases significantly as \(\kappa\) increases (it triples for \(\kappa = 0.25\) to \(\kappa = 0.75\)). This is due to the fact that the volatility of the underlying spread increases as \(\kappa \to 1\), since it reduces the stabilizing mean-reversion property. Since \(\hat{\pi}\) is a deterministic and constant strategy, the increasing volatility is ultimately reflected in the terminal wealth distribution. The same effect can be observed for the optimal strategy \(\pi^*\), however, its dynamic structure allows for punctual reaction to changing market situation, and therefore for the exploitation of favorable situations (i.e. \(s_t > \bar{S}\)). This results in a strong improvement of the expected terminal wealth. The higher \(\kappa\), the more volatile the underlying spread. Consequently, the optimal strategy reacts more pronounced to falling prices below \(\bar{S}\) or raising prices above \(\bar{S}\). The difference in the expected terminal wealth nearly 11,000,000 EUR for \(\kappa\) ranging from 0.25 to 0.75.
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES

Table 5.2: Summary of the results of the numerical example. $C_{T+1} = \theta \sum_{t=0}^{T} (\pi_t)^2$ denotes the trading costs incurred by the impact parameter $\theta$. The volatility is calculated from (5.28). All numbers are in EUR, except $\kappa$.

<table>
<thead>
<tr>
<th>$\kappa$</th>
<th>$E_0[W_{T+1}]$</th>
<th>$E_0[C_{T+1}]$</th>
<th>Volatility</th>
<th>$E_0[W_{T+1}]$</th>
<th>$E_0[C_{T+1}]$</th>
<th>Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>107,760,028</td>
<td>16,262,712</td>
<td>4,734,334</td>
<td>87,394,215</td>
<td>210,577</td>
<td>796,061</td>
</tr>
<tr>
<td>0.50</td>
<td>110,595,353</td>
<td>18,354,110</td>
<td>6,839,406</td>
<td>87,401,936</td>
<td>210,577</td>
<td>1,181,935</td>
</tr>
<tr>
<td>0.75</td>
<td>118,706,615</td>
<td>24,302,049</td>
<td>14,461,392</td>
<td>87,350,795</td>
<td>210,577</td>
<td>2,255,450</td>
</tr>
</tbody>
</table>

Table 5.2 reveals the significant dominance of the optimal strategy $\pi^*$ over $\hat{\pi}$. However, the expected cost of trading are pretty much constant at 210,577 EUR for the naive strategy, but enormous and increasing for the optimal strategy: 16,262,712 EUR ($\kappa = 0.25$), 18,354,110 EUR ($\kappa = 0.50$), 24,302,049 EUR ($\kappa = 0.75$).

Figure 5.4 shows three different price paths and the corresponding liquidation behavior. The black line is simply the deterministic price path $s_t = \bar{S}$, $\forall t = 0, ..., T$. This yields $\pi^*_t = \hat{\pi}_t = \frac{Q_{tot}}{T+1}$ from (5.20). In other words, the suboptimal strategy becomes optimal and selling occurs at a constant and deterministic rate. The red price path that is most of the time below the mean-reversion level $\bar{S} = 20$ EUR generates a typical liquidation strategy: In the beginning, the trader observes a price above the mean-reversion level, thus speeds up trading. However, after week 19 the price path remains below the mean-reversion level and the trader expects a temporary increase. This induces him to buy back lots of forward contracts and to wait for the more favorable situations. As time passes by and the price path does not return as expected, the trader has to start selling to meet the boundary condition $q_{T+1} = 0$. The closer maturity moves, regardless of the price, the trader has to sell, which incurs large losses in revenues. In the case of the blue price path, the realization is most of the time fairly favorable which yields a high liquidation tempo. This even leads to the extreme case of short-selling, starting in week 30. In week 44 the trader has sold about 1100 MW in yearly forward contracts, however, he needs to buy back 600 MW in order to meet the boundary condition, which is an unfortunate situation, because the price raises even more in the very last trading weeks. But still, the optimal trading is based on the temporary expectation of the trader and in this case it is clearly an expected negative drift in the last couple of weeks. The terminal wealth is the highest in this case with 10,004 EUR/MWh. Interestingly, the short-selling leads to company’s bursting coffers with a temporary wealth of about 23,000 EUR/MWh in week 45. However, in order to meet the boundary condition, the trader has to buy back the excess contracts. The red price path is simply the $\bar{S}$-mirrored lower blue price path and it results in the lowest terminal wealth with 9,535 EUR/MWh.

This example shows an interesting aspect. When a trader models a market according to some assumptions, he might get some useful results. Even closed-form formulae are conceivable as in our case. However, the trader does never know about the true nature
of the underlying markets. This is reflected in the sample paths in Figure 5.4 as they are probably not mean-reverting. Consequently, the trader is being fooled by the market, because it does not behave as expected.

In Theorem 5.2.1 we proved that as long as the admissible set is compact, there exists at least one optimal solution for \((P')\). The previous example shows that even for \(\pi_t \in \mathbb{R}\), we can find optimal solutions, which are even unique (as long as \(\theta > 0\)). This failure of Theorem 5.2.1 is absorbed by Proposition 5.2.3. However, changing the setting slightly does not necessarily yield optimal strategies (i.e. even Proposition 5.2.3 fails). Consider the percentage impact relation \(\tilde{S}_t = S_t - S_t \theta \pi_t\). In other words, the quantity effect is tied to the current spread price, which might be regarded as a more realistic scenario. The integrability condition \((I)\) is not fulfilled in this case: Assume \(s_t < 0\) fixed. Then \(r(s_t, a_t) = s_t a_t - \theta s_t a_t^2\) is a convex function in \(a_t\) and cannot be bounded above by a function that depends exclusively on \(s_t\). In other words, we cannot find a maximizer.

Another extension of our simple program is a more realistic liquidity impact function. For instance, we would like to have negligible quantity impact for \(\pi_t \leq Q_{\text{max}}\) and increasing quantity impact for \(\pi_t > Q_{\text{max}}\):

\[
h(\pi_t) = \theta (\pi_t + Q_{\text{max}}) \mathbf{1}_{-\pi_t > Q_{\text{max}}} + \theta (\pi_t - Q_{\text{max}}) \mathbf{1}_{\pi_t > Q_{\text{max}}},
\]

and consequently,

\[
\tilde{S}_t = S_t - h(\pi_t).
\]

Condition \((I)\) is fulfilled in this case, but the derivation of the value functions for arbitrary \(t\) is too involved due to the presence of the indicator function. A simpler approach might be

\[
h(\pi_t) = \theta \pi_t^m,
\]
for \( m = 3, 5, \ldots, 2n + 1 \). Under simple driftless random walk dynamics we obtain exactly (5.25) for any \( n \in \mathbb{N} \). For more complicated dynamics the explicit forms of \( \pi^*_t \) are again involved and tedious to be calculated explicitly. We leave these issues for future work.

Thus far, we have presented and discussed dynamic liquidation programs that can be solved explicitly. In many other practical application, however, it is impossible to obtain closed-form solutions for the functions \( \pi_t \) and \( V_t \), respectively. A vast literature has evolved in the recent decades that deal with this obstacle. There are basically two different approaches that can handle more involved programs. (1) One simplifies the state and control space by discretizing them, which circumvents the curse of dimension for many problems. A more recent and more powerful approach is (2) via approximate dynamic programming, see Powell [106]. In the next section we develop an extended grid search procedure to solve an optimization program similar to \((P)\). We do not intend to employ highly sophisticated approximation methods, but our core purpose is to show how different model parameters and specifications influence the optimal liquidation strategy. For this sake it is sufficient to consider simple tree approximations of diffusion processes.

### 5.3 Solving \((P)\): Extended Settings and Numerical Schemes

Formulating \((P')\) in a dynamic fashion allows for many more ramifications and leads to a very different liquidation behavior compared to the static case, discussed in Chapter 4. We now answer the last question in our optimal liquidation framework. How to incorporate a target wealth constraint into a dynamic program. The above discussed cases help to bridge rather simple frameworks and more sophisticated frameworks. Furthermore, they can be used as sanity checks to validate a correct extended setup.

#### Literature on Risk Constraints and Dynamic Programming

The main difficulty of merging risk constraints and dynamic programming lies in the fact that most risk measures do not obey the dynamic programming principle, i.e. they are not of recursive nature. This leads to time-inconsistent solutions. Boda & Filar [20] introduce a class of time-consistent dynamic risk measures. Krokhmal & Uryasev [87] implements risk in their model, which is a Monte-Carlo sample path approach. More specifically, Bertsimas & Lo [17] note:

*Using the first moment as the objective function implicitly assumes that the investor is risk neutral with respect to execution costs, and while this may be a plausible assumption for broker/dealers and other institutional block-equity traders, we might expect such traders to charge some sort of premium above minimum expected execution costs due to adverse selection and other strategic considerations.*

Furthermore, Geman et al. [62] argue:
5.3. SOLVING \((P)\): EXTENDED SETTINGS AND NUMERICAL SCHEMES

In finance, dynamic risk measures were recently introduced to account for the occurrence of a stream of random cash-flows over time. A general requirement for these risk measures is their time-consistency because, as emphasized by Wang [121], multi-period risks are reevaluated as new information becomes available, which raises the issue of the compatibility between consecutive decisions implied by the risk measure.

5.3.1 The Bellman Equation and Existence of Optimal Solutions under a Target Wealth

Recall that the original program \((P)\) is given by:

\[
(P) \quad \begin{cases} 
\max_{\pi \in \Pi} \mathbb{E}_0 \left[ \sum_{t=0}^{T} S_t \pi_t \right] \\
\text{s.t. } \mathbb{P} \left[ \sum_{t=0}^{T} S_t \pi_t \geq PR \right] \geq 1 - \alpha
\end{cases}
\]

In other words, risk attitude is not described by a utility function, as is often the case, but by a specific risk constraint. A common procedure to solve this program is to augment the objective by the constraint with the help of a Lagrange multiplier \(\lambda \geq 0\):

\[
L(\pi, \lambda) := \mathbb{E}_0 \left[ \sum_{t=0}^{T} S_t \pi_t \right] + \lambda \left( \alpha - 1 + \mathbb{P} \left[ \sum_{t=0}^{T} S_t \pi_t \geq PR \right] \right)
\]

As in the convex optimization problem in Chapter 4, the Lagrange function is useful to describe optimal solutions. This leads to a classical \(\text{minmax}\) problem:

\[
\min_{\lambda \geq 0} \max_{\pi \in \Pi} L(\pi, \lambda).
\]

In Chapter 4 we derived the KKT conditions to characterize optimal solution pairs \((\pi^*, \lambda^*)\). In the current case we have to solve all equations simultaneously, which is a daunting task and not feasible in many settings. Applying again the Bellman principle reduces the dimension but requires the target function to be separable in some sense. In general, one might fix \(\lambda\) and seek an optimal solution \(\pi^*(\lambda)\). In a second step, one might determine \(\lambda^*\) such that

\[
L(\pi(\lambda^*), \lambda^*) \leq L(\pi^*(\lambda^*), \lambda^*) \leq L(\pi^*(\lambda), \lambda).
\]

In other words, \((\pi^*(\lambda^*), \lambda^*)\) is a saddle point of \((\pi, \lambda) \mapsto L(\pi, \lambda)\), and hence an optimal solution of \((P)\). However, this requires the knowledge of the value functions at each
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES

time step $t = 0, \ldots, T$.

Showing existence of optimal solutions in this extended setting is ultimately tied to the risk constraint and its recursive structure. There are cases where one can find the recursive structure of the Bellman equation, and thereby prove the existence of optimal solutions. One of the earliest papers dealing with probabilistic constraints was published by White [122]. In general, $W_{T+1}$ and $P \left[ \sum_{t=0}^{T} S_t \pi_t (X_t) \geq PR \right]$ are not concave functions in $\pi_0, \ldots, \pi_T$, and the Lagrange multiplier approach does not necessarily solve the problem for any $\lambda$. The following considerations lead us to a slightly different definition of the program ($P$). Let $\lambda$ be a fixed parameter, then the minmax problem in (5.29) becomes a maximization problem of the form

$$\pi^* = \arg\max_{\pi \in \Pi} E_0 \left[ \sum_{t=0}^{T} S_t \pi_t (X_t) + \lambda \left( \alpha - 1 + 1 \sum_{t=0}^{T} S_t \pi_t (X_t) \geq PR \right) \right]$$

$$= \arg\max_{\pi \in \Pi} E_0 \left[ \sum_{t=0}^{T} S_t \pi_t (X_t) + \lambda \sum_{t=0}^{T} S_t \pi_t (X_t) \geq PR \right].$$

(5.30)

The indicator function has a discontinuity at $w = PR$. This yields optimal strategies depending on $\lambda$: $\pi^*(\lambda)$. The best approximation to the minmax problem is then obtained by choosing $\lambda$, which solves

$$\lambda^* = \arg\min_{\lambda \geq 0} \left\{ \sum_{t=0}^{T} S_t \pi_t^* (\lambda) + \lambda \left( \alpha - 1 + P \left[ \sum_{t=0}^{T} S_t \pi_t^* (\lambda) \geq PR \right] \right) \right\}.$$  

It is important to notice that the constraint means we seek to reach the planning result with the expected terminal wealth seen at time 0. It is a difference to consider the problem at each point in time. The latter case is more difficult and does only apply to very few utility functions. We have to take into consideration the time-$t$ accumulated wealth $w_t$, i.e. we would look at $\sum_{n=t+1}^{T} S_n \pi_n (X_n) \geq PR - \sum_{n=0}^{t} s_n a_n$. This constraint requires to be time-separable or at least of recursive nature. This is not applicable to our setting, i.e. we can not solve the time-$t$ problem and apply it recursively to the problem at time $t-1$.

The planning result $PR$ might be deferred from potential changes in a utility rating. On the other hand, the penalty $\lambda$ can be interpreted as the maximum cost associated with not meeting the desired level of wealth $PR$. Consequently, we leave $\lambda$ as a free parameter, which gives us a degree of freedom. For a fixed $\lambda$ we obtain optimal strategies in the $\alpha - PR$ plane. While it would be more desirable to obtain an optimal strategy without any degree of freedom, this approach still gives insight into optimal behavior of traders faced by liquidation tasks. We note that for many decision makers the values of the parameters $PR$, $\alpha$, and $\lambda$ may be unknown or available only as approximate estimates. However, by optimizing the system over ranges of values of these parameters, utilities can gain insight into their personal assessments of the risk/return trade-offs. Different values of $PR$ result in different ranges of probabilities $P \left[ \sum_{t=0}^{T} S_t \pi_t (X_t) \geq PR \right]$ that can
be attained. Typically, a decision maker has some idea of the range of probabilities he
is interested in. In a risk-neutral setting, the state space can be represented by only
two variables: \( s_t \) and \( q_t \). Due to the presence of the wealth target, the state needs to be
augmented by an additional state variable \( w_t \), which describes the accumulated wealth
up to a time \( t \) before the time-\( t \) action has occurred. The definition of the stochastic
dynamic program in Section 5.1.2 is readily transferable to this current setting. The
only slight difference is in the definition of the terminal value function:

\[
\begin{align*}

r(s_t, a_t) &= s_t \cdot a_t \\
g(s_T, q_T, w_T) &= s_T \cdot q_T + \lambda \mathbb{I}_{w_T + s_T q_T \geq PR}.
\end{align*}
\]

Condition (I) is fulfilled, because the one-period reward function is as before and the
indicator function in the terminal reward function is bounded:

\[
\begin{align*}

E_0 \left[ g^+ (s_T, q_T, w_T) \right] &\leq E_0 \left[ g (s_T, q_T, w_T) \right] \\
&\leq E_0 \left[ |s_T q_T| \right] + \lambda E_0 \left[ \mathbb{I}_{w_T + s_T q_T \geq PR} \right] \\
&\leq Q_{\max} E_0 \left[ |s_T| \right] + \lambda E_0 \left[ \mathbb{I}_{w_T + s_T q_T \geq PR} \right].
\end{align*}
\]

This justifies that we state the Bellman equation for the problem (5.30):

\[
\begin{align*}

V_T(s_t, q_t, w_t) &= g(s_T, q_T, w_T) \\
V_t(s_t, q_t, w_t) &= \max_{a_t \in \Pi_t(q_t)} \{ s_t a_t + E_t \left[ V_{t+1}(f(s_t, \varepsilon_{t+1}), q_t - a_t, w_t + s_t a_t) \right] \}.
\end{align*}
\]

For \( \lambda = 0 \) we obtain (5.7)-(5.8). The existence result stated in Theorem 5.2.1 can be
transferred to this setting without much effort. There are merely two small extensions.
(1) The state space is augmented by the accumulated wealth \( w_t \) and (2) the terminal
value function \( V_T \) has a different form.

**Theorem 5.3.1** Assume that the admissible sets \( \Pi_t(q_t), t = 0, \ldots, T \) defined in (3.5)
are non-empty, which implies that they are compact and real-valued intervals. The dynamics
of \( S \) are described by the function \( f \) in (5.3). Furthermore, consider the program

\[
\max_{\pi \in \Pi} E_0 \left[ \sum_{t=0}^{T} S_t \pi_t(X_t) + \lambda \mathbb{I}_{\sum_{t=0}^{T} S_t \pi_t(X_t) \geq PR} \right].
\]

Then there exist measurable functions \( \pi^*_t : \mathbb{R}^3 \rightarrow \mathbb{R}_+ \) as maximizers of (5.32) and
the policy \( \pi^* = (\pi^*_0, \ldots, \pi^*_T) \) solves (5.33). Furthermore, the value functions \( V_t(x_t),
t = 0, \ldots, T - 1, x_t \in E \), are continuous.

**Proof:** The proof follows the same lines as the proof of Theorem 5.2.1.
To develop some intuition for optimal solutions of (5.31)-(5.32), let us try to calculate the value functions (5.31)-(5.32) explicitly. As an example, we impose a simple white noise dynamic: \( S_t = \sigma \varepsilon_t \), for \( t = 0, \ldots, T \) and \( \varepsilon_t \sim N(0,1) \). Starting at time \( T \) yields

\[
V_T(s_T, q_T, w_T) = s_T q_T + \lambda \mathbb{I}_{w_T + s_T q_T \geq \mathbb{PR}}.
\]

Going one step back in time and using (5.32) allows us to calculate for fixed \( s_{T-1}, q_{T-1}, \) and \( w_{T-1} \):

\[
V_{T-1}(s_{T-1}, q_{T-1}, w_{T-1}) = \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ s_{T-1} a_{T-1} \right. \\
+ \mathbb{E}_{T-1} \left[ V_T(\sigma \varepsilon_{T-1}, q_{T-1} - a_{T-1}, w_{T-1} + s_{T-1} a_{T-1}) \right] \bigg| s_{T-1} a_{T-1} \right\} \\
= \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ s_{T-1} a_{T-1} + \frac{\lambda}{\sqrt{2\pi}} \int_0^1 e^{-\frac{1}{2} h(a_{T-1})} \right. \\
\left. \left( \begin{array}{c}
\Phi(a_{T-1}) \\
\Phi(a_{T-1}) - \Phi(a_{T-1} - a_{T-1})
\end{array} \right) \right\} \\
= \max_{a_{T-1} \in \Pi_{T-1}(q_{T-1})} \left\{ s_{T-1} a_{T-1} + \lambda (1 - \Phi(a_{T-1})) \right\},
\]

with \( \Phi(a_{T-1}) := \frac{\mathbb{PR} - w_{T-1} - s_{T-1} a_{T-1}}{\sigma(q_{T-1} - a_{T-1})} \). Recall that \( a_{T-1} \in [0, Q_{\max}] \). The function \( 1 - \Phi(a_{T-1}) \) is monotonically increasing for \( \mathbb{PR} < w_{T-1} + s_{T-1} q_{T-1} \) and decreasing for \( \mathbb{PR} > w_{T-1} + s_{T-1} q_{T-1} \). This yields four cases that have to be scrutinized (see Figure 5.5):

1. \( \mathbb{PR} > w_{T-1} + s_{T-1} q_{T-1} \) & \( s_{T-1} < 0 \). Then \( a_{T-1} \mapsto s_{T-1} a_{T-1} + \lambda (1 - \Phi(a_{T-1})) \) is monotonically decreasing and the maximum is attained at \( \pi^*_{T-1} = \pi^\text{max}_{T-1} \).

2. \( \mathbb{PR} > w_{T-1} + s_{T-1} q_{T-1} \) & \( s_{T-1} > 0 \). Then the maximum is attained at either bounds 0 or \( Q_{\max} \).

3. \( \mathbb{PR} < w_{T-1} + s_{T-1} q_{T-1} \) & \( s_{T-1} > 0 \). Then \( a_{T-1} \mapsto s_{T-1} a_{T-1} + \lambda (1 - \Phi(a_{T-1})) \) is monotonically increasing and the maximum is attained at \( \pi^\ast_{T-1} = \pi^\text{max}_{T-1} \).

4. \( \mathbb{PR} < w_{T-1} + s_{T-1} q_{T-1} \) & \( s_{T-1} < 0 \). Then it is unclear where the maximum is attained.

In two out of four cases we can determine explicit optimal strategies due to a unique monotonicity property of the value function. However, the value function can not be maximized explicitly, which can be seen from the first derivative w.r.t. \( a_{T-1} \):

\[
s_{T-1} - \frac{\lambda}{\sqrt{2\pi}} h'(a_{T-1}) e^{-\frac{1}{2} h^2(a_{T-1})} \neq 0,
\]

with \( h'(a_{T-1}) = \frac{\mathbb{PR} - w_{T-1} - s_{T-1} q_{T-1}}{\sigma(q_{T-1} - a_{T-1})^2} \). This equation can only be solved numerically for \( a_{T-1} \), but for the consecutive calculation of the value functions we need to know the explicit form of \( \pi^\ast_{T-1} \). In other words, we cannot find an explicit optimal solution \( \pi^\ast_{T-1} \), but we showed that the optimal strategy does indeed depend on \( \lambda \) and all other model parameters. Furthermore, the introduction of the target wealth induces the trader to deviate
from the block strategy. To see this, set $\lambda = 0$, which yields $\pi_t^* = \pi_t^{\min} 1_{s_t < 0} + \pi_t^{\max} 1_{s_t \geq 0}$.

Figure 5.5: The monotonicity of the function $a_{T-1} \mapsto s_{T-1} a_{T-1} + \lambda (1 - \Phi(h(a_{T-1})))$ can be categorized. In two out of four cases the maximum is uniquely determined (left panel).

Unlike in Section 5.2, we cannot calculate the value functions (5.31)-(5.32) in each stage explicitly. Furthermore, useful structural properties are also not readily available from our setup. This motivates us to employ numerical methods to examine optimal liquidation behavior. We construct a simple tree approximation method as we merely intend to highlight the qualitative results. In a second step one might consider more sophisticated numerical schemes, such as approximative dynamic programming. But we leave these discussions, amongst others for future research (see Section 5.5 for further comments).

5.3.2 An Extended Grid Algorithm

In this current section, we work out a grid algorithm based on a very simple idea. We discretize the state space and control space to perform the backward induction via a grid search procedure. The advantage of this method over approximation schemes is the fact that it always leads to the optimal liquidation strategy, since the optimization is performed over a finite set. Approximation schemes yield usually sub-optimal strategies. More specifically, we set the variation trade size to 1 MWh, which is a natural choice, since this is the minimum quantity that is traded in the market. From a practical point of view, continuous trading is not applicable, and hence the discrete and finite time grid with intervals of length one week are not a simplification of reality. The only unnatural
and simplistic discretization regards the price process. We approximate the underlying linear diffusions by a binomial tree with tick size one Euro. One could also impose multinomial trees with the natural tick size 0.01 EUR, but this makes the problem computationally infeasible, especially for longer horizons. Define $h := \Delta t$ as the time step and let $s_0$ be given. Then

\begin{align}
  s^u_t &= s_{t-1} + \sigma \sqrt{h} \\
  s^d_t &= s_{t-1} - \sigma \sqrt{h},
\end{align}

where the probability of an upward jump is given by $p_u := 0.5 + \frac{\mu}{2\sigma}$ and similarly, $p_d := 1 - p_u$. For the approximation of the autoregressive process we follow mainly Nelson [100]. Consider the dynamics (5.19). The upward and downward movement is as in (5.34)-(5.35). The probability for an upward move at time $t$ is given by

$$p_u(t) = \frac{1}{2} + \frac{\sqrt{h} \kappa (\bar{S} - s_t)}{2\sigma}.$$ 

Consequently, the probability for a downward move is simply $p_d(t) = 1 - p_u(t)$. The local drift is given by

$$\mu(s_t) = \kappa \left( \bar{S} - s_t \right).$$

If $p_u(t) < 0$, we set $p_u(t) := 0$ and $\mu(s_t) = \frac{\sigma}{\sqrt{h}}$. If $p_u(t) > 1$, we set $p_u(t) := 1$ and $\mu(s_t) = -\frac{\sigma}{\sqrt{h}}$. For the rest of this chapter we consider $h = 1$, since all parameters are scaled to a weekly basis. The simple idea behind a grid search procedure is to discretize the state space (i.e. make it finite) and to apply brute force. In other words, $V_t(s_t,q_t,w_t)$ is stored for all possible value combinations of $s_t$, $q_t$, and $w_t$. This requires to store a vast amount of data, and thus the task is to find a reasonable discretization scheme that on the one hand limits the computational effort and at the same time still allows for realistic scenarios. The smallest tradable unit in the electricity market in one MW. That means, $\pi_t$ is integer-valued and (3.5) becomes

$$\hat{\Pi}_t(q_t) = \left\{ \pi_t \in \mathbb{N}_0 : (q_t - (T - t) \cdot Q_{\max})^+ \leq \pi_t \leq \min\{q_t, Q_{\max}\} \right\}.$$ 

$\hat{\Pi}_t$ contains maximal $Q_{\max} + 1$ elements that have to be stored at a time. Furthermore, the price scenario tree has $t + 1$ elements $s_0^t, ..., s_1^t$ at time $t$. Consequently, at time zero there are $Q_{\max} + 1$ possible values. At time 1, there are $2 \cdot (Q_{\max} + 1)$ possible values. Up to time $T$ there are in total $(Q_{\max} + 1)^{(T+1)(T+2)/2}$ values to be stored. For, e.g. $T = 51$ and $Q_{\max} = 80$ MW, we have to store approximately 112,000 values. With the today’s computer capacities this task is easy and fast. Even though, $q_t$ is a path-dependent state variable, storing it is still manageable, because it only depends on the past realized controls $\pi_0, ..., \pi_{t-1}$. However, things become more involved when we consider the other path-dependent state variable $w_t$. This variable depends on the past actions $a_0, ..., a_{t-1}$ and the spread prices $s_0, ..., s_{t-1}$: $w_t = \sum_{n=0}^{t-1} s_n a_n$. Consequently, $w_1$
can take $Q_{\text{max}} + 1$ values. At time 1, there are two possible realizations of $s_1$. Thus, $w_1$ can take $2(Q_{\text{max}} + 1)^2$ different values. It is easy to see that the capacity needed to store the different realizations of $w_1$ grows exponentially, because there are $2^t$ different paths to get to a specific node $s_i$ at time $t$. Also, the construction of a non-degenerating tree is not feasible (this is, e.g., common practice in pricing path-dependent options). The argumentation leads to a work-around that has already been treated in the case of pricing path-dependent options (Hull & White [70]). We transfer this concept to our setting.

**Constructing a Grid for $w_t$**

At time $t$ there is a maximum value and a minimum value for $w_t$, since the spread price and the controls are bounded. This constitutes a finite set of values $w_t$. The key idea is to cover this quite dense set with a coarse grid. Assume that this grid exhibits $l$ different, equally spaced values $\min < g^1_t < \ldots < g^l_t < \max$, $t = 0, \ldots, T$. Since we imposed a tree for the spread price process, there is a maximum spread price $S_{\text{max}} := s_0 + T\sigma$ and a minimum spread price $S_{\text{min}} := s_0 - T\sigma$. Consequently, the maximum and minimum value for the wealth variable is determined by these bounds. For simplicity, we set $g^1_t := S_{\text{min}}Q_{\text{tot}}$ and $g^l_t := S_{\text{max}}Q_{\text{tot}}$, for $t = 0, \ldots, T$. The state of the system at time $t$ (formerly described by $x_t = (s_t, q_t, w_t)$) can now be specified by a triple in the finite set

$$\mathcal{E}_t = \left\{ (s^i_t, q^j_t, g^k_t) : s^i_t \in \{s_0 - t\sigma, \ldots, s_0 + t\sigma\}, q^j_t \in \{0, \ldots, Q_{\text{max}}\}, g^k_t \in \{g^1_t, \ldots, g^l_t\} \right\},$$

which is the discretized state space. The optimal value function $V_t$ becomes a three dimensional matrix, which we denote by $\hat{V}_t$. The size of this matrix is determined by the state space dimensions of $\mathcal{E}_t$. Since the imposed grid is coarse, we can have the situation where $w_t + s_t a_t$ does not fall onto the grid points $g^1_t, \ldots, g^l_t$, but somewhere in between. In these cases we employ a simple interpolation scheme to calculate the missing value. The procedure works as follows: Consider time $T$. The matrix $\hat{V}_T$ has dimensions $T + 1 \times Q_{\text{max}} + 1 \times l$. For fixed $s^j_T$, $q^j_T$, and a grid point $g^k_T$ we determine another grid point $g'^k_T$ with $l'$ being the index with $g'^k_T < g^k_T + s^j_T q^j_T < g^{l'+1}_T$ and store

$$\hat{V}_T \left(s^j_T, q^j_T, g^k_T \right) = \frac{g^k_T + s^j_T q^j_T - g'^k_T}{g^{l'+1}_T - g^k_T} \left( g^{l'+1}_T + \lambda I_{g'^k_T < \text{PR}} \right) + \frac{g^{l'+1}_T - g^k_T}{g^{l'+1}_T - g'_T} \left( g'_T + \lambda I_{g'_T \geq \text{PR}} \right). \tag{5.37}$$

This is performed for all $s^0_T, \ldots, s^j_T$, $q_T = 0, \ldots, Q_{\text{max}}$, and $g^1_T, \ldots, g^l_T$. Consider now time $t = T - 1, \ldots, 0$. Fix again a node $s^i_t$, a remaining quantity $q^j_t$, and a grid point $g^k_T$. Then,
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES

\[ V_t(s^t_i, q^t_i, g^k_t) = \max_{a_t \in \Pi_t(q^t_i)} \left\{ s^t_i a_t + \mathbb{E}_t \left[ V_{t+1}(s^t_i \pm \sigma, q^t_i - a_t, g^k_t + s^t_i a_t) \right] \right\} \]
\[ = \max_{a_t \in \Pi_t(q^t_i)} \left\{ s^t_i a_t + p_a(t) \cdot V_{t+1}(s^t_i + \sigma, q^t_i - a_t, g^k_t + s^t_i a_t) + p_d(t) \cdot V_{t+1}(s^t_i - \sigma, q^t_i - a_t, g^k_t + s^t_i a_t) \right\} \quad (5.38) \]

Being in state \((s^t_i, q^t_i, g^k_t) \in \mathcal{E}_t\) triggers the following: Choose any \(\bar{a}_t \in \bar{\Pi}(q^t_i)\). This yields a hypothetical wealth of \(w_{t+1} = g^k_t + s^t_i \bar{a}_t\), but \(w_{t+1}\) might fall between two grid points: \(g^k_t < w_{t+1} < g^k_{t+1}\). In other words, \(V_{t+1}(\ldots, w_{t+1})\) is not necessarily stored, but only \(V_{t+1}(\ldots, g^k_t)\) and \(V_{t+1}(\ldots, g^k_{t+1})\). In this case, we interpolate linearly:

\[ V_{t+1} \left( s^t_i \pm \sigma, q^t_i - \bar{a}_t, w_{t+1} \right) = \frac{w_{t+1} - g^k_{t+1}}{g^k_{t+1} - g^k_t} V_{t+1} \left( s^t_i \pm \sigma, q^t_i - \bar{a}_t, g^k_{t+1} \right) + \frac{g^k_{t+1} - w_{t+1}}{g^k_{t+1} - g^k_t} V_{t+1} \left( s^t_i \pm \sigma, q^t_i - \bar{a}_t, g^k_t \right). \quad (5.39) \]

Performing this for all elements \(a_t \in \bar{\Pi}(q^t_i)\) and choosing the strategy \(\pi^*_t\) that attains the maximum in (5.38) completes the optimization for state \((s^t_i, q^t_i, g^k_t)\) at time \(t\). Storing \(\bar{V}_t(s^t_i, q^t_i, g^k_t)\) and \(\pi^*_t\), and running through all states back to time \(t = 0\) completes the algorithm. This procedure is summarized in Algorithm 2.

**Algorithm 2: Extended Grid Algorithm**

**Step 1:** Calculate the terminal value function \(\bar{V}_T(s^T_j, q^T_j, g^T_l)\) according to (5.37) for \(s^T_j \in \{s^T_0, \ldots, s^T_T\}\), \(q^T_j \in \{0, \ldots, Q_{\text{max}}\}\), and \(g^T_l \in \{g^T_1, \ldots, g^T_T\}\).

**Step 2:**

for \(t = T - 1 : -1 : 0\) do

for \(i = 1 : t\) do

for \(j = 0 : Q_{\text{max}}\) do

for \(k = 1 : t\) do

Calculate \(\bar{V}_t(s^t_i, q^t_i, g^k_t)\) according to (5.38), and store \(\pi^*_t(s^t_i, q^t_i, g^k_t)\) and \(\bar{V}_t(s^t_i, q^t_i, g^k_t)\).

End do

End do

End do

End do

**Result:** For all combinations \((s^t_i, q^t_i, g^k_t) \in \mathcal{E}_t, t = 0, \ldots, T, \bar{V}_t\) and the corresponding optimal control \(\pi^*_t\) is stored.
In order to determine the optimal strategy \( \pi^* \) and value functions \( \bar{V}_t \) we simulate a path according to the tree dynamics (5.34)-(5.35), plug in the initial values \( s_0, q_0 = Q_{tot} \), and \( w_0 = 0 \), and run forward through the value functions \( \bar{V}_t \) until maturity \( T \). This gives us \( \pi^*_t \) and \( \bar{V}_t \), for \( t = 0, \ldots, T \) for the specific path. This procedure is summarized in Algorithm 3.

Algorithm 3: Determination of Optimal Strategies \( \pi^* \)

**Step 1:** Initialize parameters and setup matrices according to Algorithm 2.

**Step 2:** Simulate a path \((s_0, ..., s_T)\) according to the tree dynamics (5.34)-(5.35).

**Step 3:**

\[
\text{for } t = 0 : T \text{ do}
\]

Determine \( \bar{V}_t(s_t, q_t, w_t) \) according to (5.39) and the corresponding optimal strategy \( \pi^*_t(s_t, q_t, g_t) \) and set

\[
q_{t+1} = q_t - \pi^*_t(s_t, q_t, g_t),
\]

\[
g_{t+1} = g_t + s_t \pi^*_t(s_t, q_t, g_t).
\]

Finding an optimal strategy \( \pi^*(\lambda) \) eventually depends on the choice of the Lagrange parameter \( \lambda \). For a fixed \( \lambda \) and specific parameter set, we are interested in the probability that the optimal strategy \( \pi^*(\lambda) \) yields a wealth \( w_{T+1} \) that is larger than the planning result \( PR \). Assume we simulate \( N \) paths according to the tree dynamics (5.34)-(5.35) and calculate the terminal wealth \( w_{T+1}^{i} \) for each path \( i = 1, \ldots, N \) under the optimal strategy \( \pi^*_i \). The probability of reaching the planning result can approximately be deduced by

\[
\frac{1}{N} \sum_{i=1}^{N} 1_{w_{T+1}^{i} \geq PR} \approx P \left[ W_{T+1} \geq PR \right].
\]

This is an obvious result. For varying \( \lambda \in (0, \infty) \), we can examine the optimal \( E_0 [W_{T+1}] - P \left[ W_{T+1} \geq PR \right] \) pairs. This gives us more information about an optimal \( \lambda^* \), since we can fix the point on the frontier where \( P \left[ W_{T+1} \geq PR \right] = \alpha \) and read off \( \lambda^* \) and the corresponding expected terminal wealth.

The computational effort of Algorithm 2 and 3, respectively, is surprisingly small. There are two parameters, which influence the computational time most: The time horizon \( T \) and the refinement of the grid, i.e. the size of \( l \). The computational time for the time horizon is \( \sim T^2 \) and for the grid size \( \sim l \). For instance, choosing \( T = 51 \) and \( l = 400 \),
Algorithm 2 takes a few seconds to provide the optimal strategies and the optimal value functions and Algorithm 3 takes with \( N = 50,000 \) simulations about the same time, i.e. only a few seconds.

### 5.3.3 Case Study

In this section we conduct numerical experiments for the extended program using the algorithms introduced in the previous section. The representative portfolio is again taken from Section 1.3. The developed framework allows for plenty of settings to be scrutinized. For demonstration purposes it suffices to focus on the random walk model with positive drift. In order to make things comparable, let us fix three realistic spread price paths and examine the optimal liquidation strategies under different model specifications. Figure 5.6 depicts the representative paths. In general, the market participants do not really know the true nature of the underlying dynamics. The paths in Figure 5.6 could have positive as well as negative drift. The decisive question is whether the resulting liquidation differs significantly.

![Three representative spread price paths with \( s_0 = 20 \) EUR.](image)

**Figure 5.6:** Three representative spread price paths with \( s_0 = 20 \) EUR.

### Sanity Checks for the Frictionless Market

In a first step let us apply the grid algorithm for the case \( \lambda = 0 \) and compare the results with Proposition 5.2.5. In this case, closed-form formulae are available. This serves as a *double check* to verify that Algorithms 2 and 3 are implemented correctly and that
the coarse grid is constructed reasonably. We see in Figure 5.7 that the optimal strategy is to sell everything at the end regardless of the planning result since there is no penalty, i.e. we have a simple wealth maximization. This is in line with our results in Proposition 5.2.5. Note that there are still paths and corresponding optimal strategies that fail to reach the planning result $PR = 22,000$ EUR/MWh. More specifically, we have $\Pr[W_{T+1} \geq PR] = 75.56\%$ (this probability could basically be calculated explicitly, since the distribution of $W_{T+1}$ is known in this case). The histogram is mainly shaped by the terminal distribution of the random walk with positive drift, which is Gaussian. Comparing the mean and volatility of the simulation with the result in Corollary 5.2.8 reveals an astonishing consistence. The Monte-Carlo error for the mean is roughly 0.2% and for the volatility is roughly 0.5%. This sanity check validates the usefulness of the grid algorithm.

Figure 5.7: Sanity check for the random walk: $\sigma = 1$, $\mu = 0.5$, $s_0 = 20$ EUR, $PR = 20,000$ EUR (indicated by the shorter bar in the lower, right figure), $\lambda = 0$ EUR. The longer bar in the lower, right figure corresponds to the mean of the terminal wealth, which is equal to 22,118 EUR (this number can also be recovered by using Corollary 5.2.8: $E_0[W_{T+1}] = 22,090$ EUR, i.e. a Monte-Carlo error of less than 0.2%).

What Does a Trader Induce to Change Strategies?

Let’s look Figure 5.8.

In a static setting, the management rule induces an acceleration of the liquidation,
Figure 5.8: Optimal liquidation under a target wealth of 12,000 EUR and a penalty of $\lambda = 2500$ EUR.

as shown in Section 4.3.4. In a dynamic setting, the situation is different. Assume an arbitrary time $\hat{t} \in \{0, ..., T\}$. The trader has the following information (before the action is taken): The state variables $s_{\hat{t}}, q_{\hat{t}}, w_{\hat{t}}$, the market parameters $\mu > 0$ and $\sigma$, the penalty $\lambda$, and the residual revenues that he has to achieve:

$$s_{\hat{t}} \cdot \pi_{\hat{t}}(x_{\hat{t}}) \sum_{t=\hat{t}+1}^{T} S_t \pi_t(X_t) \geq \text{PR} - w_{\hat{t}}$$

These information are incorporated into the decision at time $\hat{t}$. The liquidation behavior gives a surprising insight. The blue path in Figure 5.6 is quite in favor of the trader. He observes a high and persisting spread price and knows that the achievement of the planning result occurs with high probability. Thus, there is no incentive to deviate from the block strategy, since the volatility $\sigma$ and the penalty $\lambda$ are moderate. The same behavior is observable for the red path. The liquidation behavior in Figure 5.8 is the same block strategy, because the trader knows that the planning result will probably not be achieved. Thus, he optimizes his strategy on the drift (which is positive), neglecting other state variables. The penalty has to be paid in any case. A different conclusion can be drawn for the black price path. In this case, the trader anticipates the chance to achieve the planning result. To secure some of the profits he starts liquidating at some earlier point in time. And indeed, at the end of the liquidation period he achieves the planning result.

This argumentation can also be conducted by calculating the value function at time $T - 1$ according to (5.38). Arranging the terms appropriately yields
5.3. SOLVING \((P)\): EXTENDED SETTINGS AND NUMERICAL SCHEMES

\[
V_{T-1}(s_{T-1}, q_{T-1}, w_{T-1}) = \max_{a_{T-1} \in \Pi(q_t)} \left\{ -\mu a_{T-1} + \frac{\lambda}{2} \left( 1 + \frac{\mu}{\sigma} \right) \mathbb{I}_{c+q_{T-1} \geq a_{T-1}} \right. \\
+ \frac{\lambda}{2} \left( 1 - \frac{\mu}{\sigma} \right) \mathbb{I}_{c+q_{T-1} \leq a_{T-1}} \right\} + q_{T-1}(s_{T-1} + \mu),
\]

with \(c := \frac{1}{\sigma}(w_{T-1} + s_{T-1}q_{T-1} - \text{PR})\). If the trader has not made much money during the liquidation period, then \(c < 0\). Let us assume that there is no \(a_{T-1}\) such that the second indicator function becomes one. Then, \(\pi_{T-1}^* = \pi_{T-1}^{\text{min}}\). On the other hand, assume the trader has already a significant amount of wealth secured, then \(c > 0\) and \(\pi_{T-1}^* = \pi_{T-1}^{\text{min}}\). This argumentation supports the liquidation in Figure 5.8. For \(\lambda\) being quite large we can observe different liquidation behavior, compare Figure 5.9.

![Figure 5.9: Optimal liquidation under a target wealth of 12,000 EUR and a penalty of \(\lambda = 7500\) EUR (upper panel) and \(\lambda = 100,000\) EUR (lower panel).](image)

The favorable blue price path induces an acceleration of the trading tempo, because the penalty appears to large. Consequently, the trader locks in prices at an early stage and at the end he achieves the planning result. The red path is still too poor and the trader almost neglects the high penalty and considers merely the market drift. This is a very intuitive explanation.

The above cases show that the trader incorporates the penalty \(\lambda\) into his decision and deviates from the simple block trading strategy. In other words, \(\lambda\) is some sort of risk-
aversion parameter. The same analysis can be performed for different values of \( \mu \) and \( \sigma \), as well as for the autoregressive process. The results are as intuitive as in the presented case, but omitted for brevity.

As stated in earlier, we can also calculate the probability of reaching the planning result for a fixed penalty \( \lambda \) and plot this probability for varying values of \( \text{PR} \). Even though, we do not intend to find an optimal \( \lambda \) in the optimization problem, this information gives the trader an idea of the success of an optimal strategy under a specific framework. For illustration, we consider a random walk specification with \( \mu = -0.5, 0, 0.5, \sigma = 1 \) and \( s_0 = 20 \text{ EUR} \). Figure 5.10 depicts some typical distribution functions which can be seen as an approximation to \( \text{PR} \mapsto \mathbb{P} \left[ W_{T+1} \geq \text{PR} \right] \). For a given risk level \( \alpha \), the trader can read off the planning result that should be fixed in order to meet the risk level under the optimal strategy. This information might give policy makers an indication for setting up the management rule \( (\text{PR}, \alpha) \).

![Figure 5.10: Probability of reaching the planning result \( \text{PR} \) under a random walk specification.](image)

The higher the planning result, the rather the trader fails to reach it. This is an obvious consequence, however, the insight has also a quantitative feature. Utility companies usually have a compulsory risk level \( \alpha \) and the penalty \( \lambda \) is, e.g. specified by some control mechanism or inspecting authority. Consequently, Figure 5.10 provides a reasonable target wealth, which simply has to be read off this graph. Another aspect that is important from a practical point of view is the fact that for varying \( \lambda \) the probability of reaching the planning result does not differ. Even though, the resulting optimal strategies might do (as a matter of fact, the strategies can be very different for different values of \( \lambda \)), the distribution of the terminal wealth is independent of \( \lambda \). Furthermore, in the view of the second optimization step, where we choose the optimal strategy \( \pi^*(\lambda^*) \) from the set \( \{\pi^*(\lambda)\}_{\lambda \geq 0} \), it gives us an idea that the final optimal strategy \( \pi^*(\lambda^*) \) yields terminal wealths that are close to the ones reached by the other optimal strategies \( \pi^*(\lambda) \).
5.4. **CONCLUSIONS & DISCUSSION**

In Figure 5.10 we can read off that for $\alpha = 5\%$, i.e. the policy makers look for a strategy that reaches a planning result with 95% probability. The corresponding target wealth should be set to $\overline{PR} = 6000$ EUR/MWh. This is a fairly small number, considering that the initial portfolio value is 12,000 EUR/MWh and we expect a positive drift in the market. Nevertheless, 95% probability is quite high and we are convinced that the numbers presented here are realistic.

5.4 **Conclusions & Discussion**

In contrast to the static liquidation framework discussed in Chapter 4, a dynamic approach allows for flexible and precise reaction to changing market situations. However, this comes along with a price of higher complexity of the underlying mathematical concepts. Even though, we discussed quite basic frameworks, they help to understand optimal liquidation behavior in illiquid markets.

One of the main differences between a random walk and an autoregressive process is the drift. A random walk induces a global drift, which remains constant irrespectively of future prices. An autoregressive process induces a local drift, because the trader has one more piece of information - the mean-reversion level $\overline{S}$. This advantage is reflected in the optimal dynamic liquidation strategies. Already in a frictionless environment (Section 5.2.2) the trader deviates from the optimal static strategy by incorporating future information whether or not the prices are above or below the mean-reversion level $\overline{S}$. This is in contrast to a simple random walk model where the optimal dynamic strategy coincides with the optimal static strategy. Extending the setting by introducing the notion of liquidity risk yields a different strategy, however, it is still a deterministic one, not regarding future prices or inventory levels. This is a direct consequence of the simple random walk model. Under an autoregressive process we obtain quite involved formulae for the optimal controls, which depend indeed on future prices $S_t$ and inventory levels $Q_t$. The recursive structure of the strategy (5.20) can further be simplified such that it does not depend on the inventory level any longer and for $\kappa \to 1$ we can show that it converges to a stationary policy.

The advantage of closed-form formulae becomes apparent in these settings as they allow for a detailed analysis of optimal behavior. Incorporating a target wealth into the framework raises many questions. In a first step we have to augment the target function by this constraint, which brings along a dual variable $\lambda$. This complicates the whole issue considerably. We argue how to circumvent the problem of finding optimal pairs $(\pi^*, \lambda^*)$ and show that existence of optimal solutions $\pi^*$ under $\lambda$ being a degree of freedom are straightforward to obtain. Having derived the Bellman equation for this problem setup leads ultimately to the attempt to solve the sequential maximization problems. We show in Section 5.5 that the explicit calculation of optimal strategies becomes quite involved due to the presence of indicator functions. However, already the calculation of the value function $V_{T-1, \pi_{T-1}}$ at time $T-1$ provides valuable insight and contributes
to the explanations of the subsequent numerical experiments. We develop a numerical scheme that is computationally tractable and allows at the same time for reasonable tests on dynamic liquidation behavior under a target wealth. The astonishing clue is the separation of the trading behavior. Imposing a random walk with a positive drift yields a slow liquidation tempo for a very favorable and unfavorable path, as long as the penalty $\lambda$ is moderate. This can be explained by the insignificance of the penalty as either (1) the trader observes a high enough price such that it is very unlikely to ever fall below the planning result $\overline{PR}$, or (2) the trader observes very low prices such that he has to pay the penalty in any case, thus he only optimizes on the positive drift. For a path in between both extremes he locks prices at an early stage, since there is a chance to reach the planning result as well as to stay below.

After all, the dynamic framework requires much more effort but still yields valuable formulae and numerical experiments. Even though the setup is fairly simple, it establishes a golden thread that could be adhered to when imposing extended settings such as modeling of the margins. To complete this comprehensive section on optimal liquidation we give some suggestions on future work.

5.5 Extensions: Margin Modeling

The framework developed in this current chapter is wide open to be extended in several directions. First of all, since the electricity is sold in the forward/futures market, there is no need to buy the fuels (such as coal) at the same time. This gives room for optimization. Usually, utility companies optimize their portfolios considering the spread and by establishing positions that minimize the risk, e.g. the variance. This is basically a Markowitz approach, which is quite robust, but static. Considering a dynamic setup first of all leads to more instability and a higher dependence on the underlying model. However, from a theoretical standpoint it is interesting to examine a possible framework for the valuation and determination of such an optimization problem. When the quantity of coal necessary for production is not ultimately tied to the quantity of electricity sold at some point in time, it is required that the margins are modeled separately, i.e. electricity and coal (and possibly further portfolio elements) have their own dynamics. Consequently, in a first step one could model the margins separately, but still require to buy the necessary coal at the same time the electricity is sold. In a second step one could consider two separate optimization problems, i.e. liquidate the quantity of electricity $Q_{e,\text{tot}}$ optimally and purchase the needed quantity of coal $Q_{c,\text{tot}}$ optimally within $t = 0, \ldots, T$. We saw that a major impact on the liquidation behavior is induced by the drift exhibited by the underlying model. Thus, modeling the margins by some simple dynamics which are tied via a correlation coefficient $\rho$ would probably not have much influence on the resulting strategies. In particular, when we consider Gaussian dynamics, a linear correlation coefficient $\rho$ contributes merely to the volatility of the resulting spread dynamics. Consequently, one might impose models that have an impact on the resulting spread drift component. This leads ultimately to cointegrated systems,
5.5. EXTENSIONS: MARGIN MODELING

which are of increasing interest when it comes to modeling energy markets.

Cointegrated Systems

Cointegrated systems are natural extensions of one-dimensional econometric models, see e.g. Comte [40]. Alexander [2] introduced the concept of cointegration to energy markets. The simplest cointegrated system is a straightforward extension of an autoregressive process of order 1 (AR(1)). This model is referred to as a vector-autoregressive process (VAR(1)) and the dynamics are given by

\[
\begin{align*}
\frac{dF_1(t,T_0)}{dt} &= (\mu_1 + \delta_1 S_t)dt + \sigma_1 dW_1(t) \\
\frac{dF_2(t,T_0)}{dt} &= (\mu_2 + \delta_2 S_t)dt + \sigma_2 dW_2(t) \\
S_t &= a + bt + cF_2(t,T_0) + F_1(t,T_0),
\end{align*}
\]

where \(\{W_1(t)\}_{t \geq 0}\) and \(\{W_2(t)\}_{t \geq 0}\) are two correlated Brownian motions with correlation coefficient \(\rho\). For simplicity we omit the emissions prices. Writing \(S_t\) in differential form yields

\[
\begin{align*}
\frac{dS_t}{dt} &= b \cdot dt + c \cdot \frac{dF_1(t,T_0)}{dt} + \frac{dF_2(t,T_0)}{dt} \\
&= (b + c\mu_1 + \mu_2 + (c\delta_1 + \delta_2)S_t) dt + c \cdot \sigma_1 dW_1(t) + c \cdot \sigma_2 dW_2(t) \\
&= -c\delta_1 - c\delta_2 - b + c\mu_1 + \mu_2 - c\delta_1 - c\delta_2 - S_t dt + c \cdot \sigma_1 dW_1(t) + c \cdot \sigma_2 dW_2(t).
\end{align*}
\]

If \(-(c\delta_1 + \delta_2)\) is a positive constant, then \(S_t\) is a stationary Ornstein-Uhlenbeck process, which is of the same form as (3.16) with \(\kappa = -(c\delta_1 + \delta_2)\) and \(\bar{S} = -\frac{b + c\mu_1 + \mu_2}{c\delta_1 + \delta_2}\). \(W_1(t)\) and \(W_2(t)\) are standard Brownian motions, and hence \(W(t) := c\sigma_1 W_1(t) + \sigma_2 W_2(t)\) is again a Brownian motion with zero drift and variance \(\sigma^2 = (c^2 \sigma_1^2 + \sigma_2^2) t\). An important feature, in contrast to the previous models, of the cointegrated diffusion system is that the margins must not be mean-reverting (or stationary in some sense). This is a desirable property, since forward markets of electricity or coal do not necessarily exhibit mean-reversion. However, as mentioned earlier, the spread among several commodities is often regarded to be mean-reverting. If we set \(a = b = 0\) and \(c = -h_1\), we obtain the desired form for the spread. We suggest to follow the same lines as we did for the spread models discussed in the previous chapters. In other words, try to find simple settings under which explicit optimal strategies can be computed. This provides some basic insight into optimal liquidation behavior. Ramifications are always applicable, since the numerical methods are feasible for a wide range of frameworks.
CHAPTER 5. OPTIMAL DYNAMIC LIQUIDATION STRATEGIES
Chapter 6

Pricing of Structured Retail Electricity Contracts

In the subsequent text we show how financial models can be used to price structured contracts in the electricity market, in analogy to interest rates products. A structured contract can be understood as a hybrid product of standardized financial instruments available in the markets. Since standard products do not necessarily cover all specific investor needs, more and more specialized products are being developed and issued. In the electricity markets, such contracts are for example delivery contracts with retail or institutional customers. In fact, many features of electricity products exist similarly in structured LIBOR products, and some models have already been adapted to the world of commodities, see e.g. Schwartz & Smith [112]. Yet, we are the first to show how they can be applied to value such structures in the retail electricity market. Though, the market under consideration is far from being driven by rational participants only, we show how to quantify features such as call rights for customers, hedge analysis for utilities, and assign monetary value to laziness, i.e. irrational behavior of customers, based exclusively on publicly available information.

Only recently, power delivery contracts between utilities and retail customers have been endowed with the the one-sided (i.e. for customers) right, but not the obligation to cancel the agreement, for instance, in order to switch to an alternative contributor. These call rights are included naturally in the contracts in contrast to interest rates markets, where they have been added artificially to bonds in order to allow for more interesting payment structures. The main difference is that in interest rates markets the holder of the cancel right is in general the issuing bank while in the case of electricity markets it is the customer. In recent days, utility serving companies advertise such cancel features, since it has an option value which should be appreciated by the holder and accounted for by the seller. A practical challenge is the setup of the framework that makes the different contracts comparable. Until today, there is a fairly broad variety of retail contracts available and we encountered quite some challenges in eliciting all necessary information about the contracts on offer. Apart from the information available through
the internet, where one can find out about demand-dependent rates, contract duration, specifics on the cancel rights, etc., one also needs certain background information such as the pricing date of the contract, the hedge behavior of the specific utility, and many more.

The remainder of this chapter is organized as follows:

In Section 6.1, we give a detailed description of the retail market in Germany and the motivation to price structured retail contracts. Section 6.2 introduces the model used for our analysis, derives properties of the pricing approach via dynamic programming, and describes the valuation methodology based on a least-squares Monte-Carlo method. Section 6.3 provides a comprehensive case study, where we show how to value and examine the chosen contracts. Moreover, we describe a method to price such structured contracts in a sense that makes the contract fair for both, the supplier and the customer. Section 6.4 concludes.

6.1 Introduction

6.1.1 Related Literature & Contribution

It is our goal to develop the entire procedure from modeling the underlying market to pricing and analysis of such retail contracts based on some real-world examples using only publicly available information. Research on electricity markets in general abounds in recent years, though most of it focuses (more or less) on wholesale markets where utilities, banks, and trading firms participate. Literature, related to this current chapter can be divided into two areas: Models for price processes and derivative pricing techniques.

There is a large variety of price processes available and naturally any pricing method depends on them. Here we would like to mention the recent analysis by Fleten & Povh [105], which helps to choose an appropriate forward price process as well as the general overview by Lucia & Schwartz [93] on the same topic. A risk-neutral modeling approach is given by Clewlow & Strickland [39] through a one-factor forward curve model, which we use in a slightly different form of a two-factor dynamics developed by Börger et al. [23]. Research on pricing of derivative products is limited to typical application by utilities. Techniques related to those applied in this chapter can be found in the work by Cartea & Williams [33], who price interruptible contracts in gas markets and the thesis by Kluge [84], who also discusses the problem of pricing multiple options and solves this issue by employing least-squares Monte-Carlo techniques.

While the studies mentioned above focus on the correct description of wholesale price dynamics and the pricing within wholesale markets, to our knowledge, there is no published work on the pricing and comparison of retail electricity contracts. It is the novelty of this work to show how well-known concepts of mathematical finance (i.e. risk-neutral
6.1. INTRODUCTION

pricing of complex payoff-structures) can be adapted to this area. More specifically, we develop a pricing framework that joins different fields of mathematical finance. First, we model the underlying market by an appropriate stochastic process and calibrate the model to traded electricity products. The extracted information (i.e. market parameters) are then used to feed a more complex framework that produces information about the structured contract. This complex framework is basically an early-exercise (or Bermudan style) option with special (in the sense, how the retail contracts are designed) payment flows.

We pursue two different objectives when it comes to the callable retail contracts. On the one hand, we are interested in the (monetary) analysis of existing contracts. In other words, these contracts are already priced by the utility companies. Therefore, we examine whether or not these contracts are priced in a fair fashion. Naturally, these contracts can be too cheap or too expensive from a customers standpoint. On the other hand, we show how to price such contracts, i.e. how should the contract price be chosen such that the contract has a fair value for both, the utility and the customer? We refer to the latter case as value-neutral specifications, see Section 6.3.5.

6.1.2 The Retail Electricity Price

Since 1998, the German electricity market is deregulated and by now customers are, for most contracts, free to change their supplier. Utilities on the other hand, try to attract customers by more or less structured contracts. The structuring usually includes fixed price warranties for a certain time horizon and rights for the customer to terminate the contract unilaterally. From the customers’ perspective, main drivers for entering into a particular contract are the (perceived) retail electricity price and customers’ willingness to change the supplier. From the utilities’ side, the pricing of such contracts depends on several risk sources such as the uncertain development of electricity prices, electricity consumption (i.e. the customers demand), and credit behavior, to name just a few. Though, all components may have a significant influence on the contract specification, it is the retail electricity price and the uncertain electricity demand, which we analyze and increase transparency in. Even though, credit-worthiness might be an important risk factor that has to be taken into consideration, we leave this issue to future research.

Retail customers cannot base their decisions about electricity contracts on any publicly quoted reference prices. They can only compare the different delivery contracts on offer, analyze the price, and additional terms and conditions in order to assign a value to an offer, which can be translated into a retail electricity price. The main ingredients for this price are:

- Wholesale price - quoted at the EEX\(^1\), includes profits connected with the power generation.

\(^1\)European Energy Exchange, www.eex.com
- Fees and taxes - includes usage of transmission lines, VAT (value-added tax), concession levy, etc.

According to the BDEW\(^2\), about 63\% of the electricity price effectively paid by the customer are fees and taxes. Supplier, resp. customers have only little influence on this component. Also, there is almost no uncertainty in the amount of the additional costs. Wholesale prices are the main drivers of uncertainty in retail prices. Figure 6.1 shows the decomposition of the electricity price in Germany.

![Pie chart showing the decomposition of the electricity price in Germany.]

Figure 6.1: Decomposition of the electricity price in Germany. Status: December, 2009. Source: www.bdew.de

For the remainder of the text, we define the *retail price* as the wholesale price divided by 37\%. This retail price is the price, which is perceived as a fair price for the customer as it does not include any additional sales margins (but only the fees and taxes). Thus, it should serve as a reference price for analyzing retail contracts. As a matter of fact, it is possible to model wholesale prices directly and add relative (e.g. value-added tax) and absolute add-on’s (e.g. grid charges) separately. The proposed approach is chosen to focus on the main idea of this chapter and to reduce complexity.

### 6.1.3 Retail Contract Structure

The contracts under consideration can be described by four main characteristics. First, a *kilowatt-hour rate* is specified in ct/kWh. This reimburses the supplier for providing the actual electricity and payments are usually done on a monthly basis. We denote

---

\(^2\)Bundesverband der Energie- und Wasserwirtschaft, www.bdew.de
6.1. INTRODUCTION

this quantity by \( k > 0 \). Second, a demand rate is given in EUR/year, denoted by \( K > 0 \). This is a consumption-independent quantity and is meant to cover fixed costs for, e.g. keeping track of the customers account during the course of a year. Third, the customer has the unilateral right to resign from the agreement on a set of pre-specified dates \( t_1, \ldots, t_{n-1} \). The interval length can be, e.g. monthly or yearly. Fourth, a time horizon \( t_n \) is agreed upon during which the utility company must not change any of the two prices \( k \) and \( K \), respectively (fixed price warranty). At the end of this period, the supplier is free to change terms and conditions and the customer may or may not resign from the contract in response to that change. Thus, we consider the time horizon \( t_n \) as the maturity date of the agreement.

In the course of the agreement the customer consumes electricity and pays according to the agreed price structure. If the consumed volume is known in advance and assumed deterministic, the agreement is a fixed-for-floating swap, see Figure 6.2. The fixed leg is the payment, the floating leg is the electricity delivery, which has value according to the corresponding wholesale prices. Equipped with cancel rights, the agreement resembles a multi-callable swap known from interest rates markets. The difficulty in valuing such contracts lies in the fact that exercising the option does not only lead to a stop of the exchange of money versus electricity, but one also gives away all future cancel rights, which have a value within themselves. We refer to this value as option premium. Ideally, the holder of the cancel rights calls the contract if and only if the remainder of the agreement including all options has negative value. We show in Section 6.3.4, how to determine a proper exercise strategy.

6.1.4 Customer Behavior

It is important to note that retail customers do not behave rationally. While it is almost impossible for an average customer to compare all different contracts on offer and derive an optimal decision rule, many customers simply do not care at all about electricity contract details. This leads to a suboptimal strategy when it comes to decision making and exercising cancel rights, in particular in connection with early-exercise features, where optimal exercise is a mathematically challenging issue. But as the customer is exercising suboptimally, an analysis based on rational investors leads to a conservative boundary in the valuation process (conservative for the utility). Consequently, a rational customer can be interpreted as a worst-case scenario for the power delivering company, which is an important information in the planning process. Moreover, even if an optimal exercise strategy is known, it is questionable whether the retail market is so liquid that an alternative, cheaper delivery contract is available at all. We can counter this thought in two ways: First, each customer has the right to be part of a base supply contract, which is monitored by the government for its fairness to some degree. This can be regarded as a kind of spot delivery in exchange for spot prices as the customer can quit this contract on short notice and the supplier can change the kWh-rate \( k \), as well. Thus, there is always an alternative, which is supposed to be fair. Second, one can take the view of such a contract as an investment, neglecting the practical need to have
6.2 Valuation Framework for Structured Contracts

A natural starting point is the analysis of the payment flows that are relevant for the valuation process. Endowing certain payment flows with options to interrupt these flows leads to the necessity to employ no-arbitrage theory to assign a value to these options. The underlying model is used to extract market parameters that are plugged into the more complex framework, which in most cases can be validated using primarily numerical methods.

6.2.1 Analysis of Payment Flows

Figure 6.2 depicts the general flow of electricity vs. cash. The utility company offers to deliver electricity to the customer. This electricity is purchased in the wholesale market or produced in a power plant and the utility has to pay the forward price

\[ F(t, t_i) := F(t, t_i, t_{i+1}), \text{ for } i = 1, ..., n - 1. \]

For notational convenience we omit almost always the indication of the delivery period \([t_i, t_{i+1}]\). On the other side, the customer receives the electricity according to his demand and pays usually a consumption-dependent fee to the utility (as a function of \(k\) and \(K\)).

![Figure 6.2: Flow of electricity versus cash for the retail contracts under consideration.](image_url)

At the EEX, the forward prices are noted in EUR/MWh. However, we consider all monetary quantities in either cents or ct/kWh. The reason for the conversion is the fact that retail contracts are commonly advertised in ct/kWh. One EUR/MWh corresponds to ten ct/kWh, since 1 MWh = 1000 kWh and 1 EUR = 100 ct. In this current chapter, we examine a simpler case where the electricity demand of the customer is known in advance, and hence we can calculate the costs that accrue throughout the delivery period. If the customer knew his future consumption, he would know his charges from utility company at the end of each period \([t_i, t_{i+1}], i = 1, ..., n - 1\). For simplicity, we
assume throughout this chapter that the payments are in arrears, e.g. at the end of each month. Recall that the contract specification consists of a kilowatt-hour rate \( k \) measured in ct/kWh, which is consumption-dependent and of a fixed rate \( K \), usually on a yearly basis, measured in EUR (one could also consider a general fixed rate that has to be paid). It is important to get the units straight, thus we define the accruing costs \( AC \) for a period \([t_i, t_{i+1}]\) (measured in hours) as

\[
AC = (K \text{ in ct}) \cdot (t_{i+1} - t_i) + (k \text{ in ct/kWh}) \cdot \text{(Demand in } [t_i, t_{i+1}] \text{ in kWh}). \tag{6.1}
\]

Hence, \( AC \) is measured in cents and corresponds to the quantity that the utility charges for the electricity consumption in \([t_i, t_{i+1}]\). For the current case, the costs are only deterministic quantities (no random variables). Consider a contract with several months of delivery, say \([t_1, t_2], [t_2, t_3], ..., [t_{n-1}, t_n]\), maturity at \( t_n \) and each interval corresponds to a month. The payment structure resembles a so-called fixed-for-floating swap known from interest rates markets. The fixed leg are the accruing costs \( AC \) and the floating leg is the electricity delivery, which has value according to the corresponding wholesale prices (and which are floating). Swaps in energy markets are similar to swaps in financial markets and are natural generalizations of forward contracts (which are basically a one-period swap). We consider a stochastic process \( \{\hat{D}_t\}_{t \geq 0} \) on a filtered probability space, which denotes the instantaneous demand. For now, the specific form of \( \hat{D}_t \) is irrelevant. Define

\[
D_{ti} = \int_{t_i}^{t_{i+1}} \hat{D}_\tau d\tau, \text{ for } i = 1, ..., n-1. \tag{6.2}
\]

This is the accumulated demand in \([t_i, t_{i+1}]\). The integral exists if

\[
\int_{t_i}^{t_{i+1}} |\hat{D}_\tau| d\tau < \infty, \text{ a.s.}
\]

Furthermore, we introduce the conditional expected accumulated demand for \( t_j \leq t_i \) as

\[
E\left[D_{ti} | \hat{D}_{t_j}\right], \text{ for } i = 1, ..., n-1. \tag{6.3}
\]

The expectation is well-defined if

\[
\int_{t_i}^{t_{i+1}} E\left[|\hat{D}_\tau| | \hat{D}_{t_j}\right] d\tau < \infty, \text{ a.s.}
\]

\( \hat{D}_t \) is measured in kW and \( D_{ti} \) is measured in kWh. If \([t_i, t_{i+1}]\) is a one-month period and the instantaneous demand \( \hat{D}_t \) is assumed to be deterministic and constant throughout this period, i.e. \( \hat{D}_t = \hat{D} \), then we obtain \( D_{ti} = D(t_{i+1} - t_i) \). A realistic quantity for a single household is \( D_{ti} = 125 \text{ kWh} \). Since one month has 730 hours we obtain \( \hat{D} = \frac{125 \text{ kWh}}{730 \text{ h}} \approx 0.17 \text{ kW} \) for the instantaneous demand. In this current chapter we assume that the cumulated demand is constant over the entire horizon under consideration \( t_1, ..., t_n \). I.e. in each period we consider
\[ D_{t_i} \equiv D, \text{ for } i = 1, \ldots, n - 1. \]

An extension to a non-constant, but deterministic demand (such as a seasonal component) is straightforward and complicates the calculations only slightly. Furthermore, in (6.3) we condition on the instantaneous demand \( D_t \), but we could as well condition on cumulated demand \( D_{t_j}, t_j < t_i \). The central quantity that we work with is the net present value (NPV) of the contract. At time \( t_0 \), we calculate the net present value of the electricity contract by

\[
P V_{t_0} (F(t_0,.), AC) := \sum_{i=1}^{n-1} e^{-r(t_i-t_0)} (D \cdot F(t_0, t_i) - AC),
\]

where \( F(t_0, t_i) \) denotes the \( t_0 \)-forward price with monthly delivery, starting at the beginning of month \( i \). \( AC \) is given in (6.1) and \( r > 0 \) denotes the riskless and flat interest rate. To ease notation we sometimes drop the arguments of the net present value:

\[
P V_{t_0} := PV_{t_0} (F(t_0, .), AC).
\]

Since the demand \( D \) is known, there are two degrees of freedom, namely \( K \) and \( k \). Furthermore, the fair price of a swap at time \( t_0 \) can be obtained by setting

\[
P V_{t_0} = 0 \iff AC = \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} D \cdot F(t_0, t_i)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)}}. \tag{6.5}
\]

In other words, the accruing costs (which are the fair rate to be charged) can be determined by calculating the fair swap rate. We show in Section 6.3.3, how to calculate the fair swap rate for some real-world retail contracts. A positive net present value means that the contract is advantageous for the customer. A negative net present value has accordingly a benefit for the utility.

**Remark 6.2.1** The quantity \( AC \) can also be measured in ct/kWh and not in cents as defined in (6.5). For instance, given a deterministic demand \( D \) for the period \([t_i, t_{i+1}]\) measured in kWh, the accruing cost (per kWh) are given by

\[
\hat{AC} := \frac{AC}{D} = \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} F(t_0, t_i)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)}}. \tag{6.6}
\]
The calculation and analysis of the net present value for different contracts is straightforward and does not require any stochastic model. However, in order to examine the impact of options to interrupt the flow of electricity vs. cash, we need to impose a model for the underlying market dynamics.

6.2.2 Market Dynamics

We introduce the model that describes the underlying forward market. In Börger [23], a two-factor forward model has been proven useful for electricity markets, capturing volatility term-structure effects and the influence of delivery periods on the forwards’ volatility. It has successfully been calibrated to European options on forwards and we employ the calibration procedure described in Börger [23] to infer market parameters that are used to price the more complex retail contracts. In this section, we briefly discuss the model and the calibration procedure. The results are presented in Section 6.3, along with the analysis of the retail contracts.

Consider a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, Q), T' \in \mathbb{R}\) and a tenor structure \(T = \{t_i \leq T', i \in \mathbb{N}\}\) with \(t_i\) denoting the beginning of each calendar month. Let \(F(t, t_i)\) denote the time-\(t\) forward price of 1 MWh electricity to be delivered constantly over a one-month period starting at \(t_i \in T\). The conditional expectation is denoted by \(E_{t}[\cdot] := E[\cdot | \mathcal{F}_t]\). Assuming that the market is free of arbitrage and the interest rate is given by a constant and deterministic continuously compounded \(r\) we can impose martingale dynamics for all forwards \(F(t, t_i), t_i \in T\) under the same risk-neutral measure via the stochastic differential equation

\[
\frac{dF(t, t_i)}{F(t, t_i)} = e^{-\eta(t_i-t)}\sigma_1 dW_1(t) + \sigma_2 dW_2(t), \tag{6.7}
\]

for a fixed \(t_i \in T\). The Brownian motions \(\{W_1(t)\}_{t \geq 0}\) and \(\{W_2(t)\}_{t \geq 0}\) are assumed to be uncorrelated. The initial value of this SDE is given by the condition to fit the initial forward curve observed at the market. This takes care of the seasonality in the maturity variable \(t_i\). The model is related to the well-known Schwartz-Smith-model [112] in that it leads to mean-reverting spot prices and features some nice interpretations: The first factor covers short-term information increasing the volatility of close-to-delivery forwards such as weather and unplanned outages. The second factor corresponds to long-term uncertainties such as macroeconomic developments.

Model Calibration

Equation (6.7) can be solved explicitly by employing Itô’s Lemma (see Øksendal [102]):

\[
F(t, t_i) = F(t_0, t_i) \exp \left\{ -\frac{1}{2}(\sigma_1^2(t, t_i) + \sigma_2^2 t) + \sigma_1 \int_0^t e^{-\eta(t_i-s)} dW_1(s) + \sigma_2 W_2(t) \right\},
\]

with \(\sigma_1^2(t, t_i) = \frac{\sigma_1^2}{2} \left( e^{-2\eta(t_i-t)} - e^{-2\eta t_i} \right)\). This explicit solution allows for straightforward European option pricing by Black’s formula [19]. A European call option on \(F(t, t_i)\) with
maturity \( T_0 < t_i \) and strike \( \bar{K} \) (note that this strike \( \bar{K} \) is different from the fixed cost \( K \) introduced above) can be evaluated by Black’s formula:

\[
V(t) = e^{-r(T_0-t)} \left( F(t, t_i)\Phi(d_1) - \bar{K}\Phi(d_2) \right),
\]

where \( \Phi \) denotes the standard Normal distribution function and

\[
d_1 = \log \frac{F(t, t_i)}{\bar{K}} + \frac{1}{2} \text{Var} \left[ \log F(T_0, t_i) \right]
\]
\[
d_2 = d_1 - \sqrt{\text{Var} \left[ \log F(T_0, t_i) \right]}.
\]

The option pricing mechanism can be used for calibrating the model implicitly by using exchange traded electricity options. In a first step we extract the implied volatility \( \text{Var}_{\text{market}}[.] \) by inverting (6.8). In a second step, the model parameters are chosen such that relevant option prices are matched well in a least-squares sense. More specifically, we seek to find a set of parameters \( \theta = \{\sigma_1, \sigma_2, \eta\} \) such that the squared difference of market- and model-implied quantities is minimal. We use all market-implied volatilities of options on forwards \( F(t, t_{\alpha}, t_{\beta}) \) with \( t_{\alpha}, t_{\beta} \in T \) and exercise date \( s_i \) and minimize the distance

\[
\sum_i \left( \text{Var}_{\text{market}} \left[ \log F(s_i, t_{\alpha}^{(i)}, t_{\beta}^{(i)}) \right] - \text{Var}_{\text{model}} \left[ \log F(s_i, t_{\alpha}^{(i)}, t_{\beta}^{(i)}) \right] \right)^2 \to \min_{\theta}.
\]

Since this model is not capable of capturing volatility smiles, which can be observed in option prices very often, we have to use at-the-money options only. The market option prices, implied volatilities and the calibration results are discussed in Section 6.3. \( \text{Var}_{\text{market}}[.] \) is the implied volatility from the model and is listed in Table 6.2. \( \text{Var}_{\text{model}}[.] \) is the model-specific variance and can be found in Börger [23].

Having introduced the model specifics, the next step is to endow the present value definition (6.4) with options to cancel the swap prior to maturity.

### 6.2.3 Endowing Payment Flows with Call Rights

When the cash flows described in the previous section are endowed with finitely many rights to cancel the agreement, we come to a similar structure known as Bermudan Swaptions in interest rates markets. However, the retail contract structure makes the valuation different in several ways:

1. The exercise of the option ceases cash exchange (usually it triggers the start of cash exchange).
2. The holder of the option is the customer (usually it is the bank).

As a first step, we consider only one right to cancel the contract at some pre-specified date \( t_m \in \{t_1, ..., t_{n-1}\} \).
6.2. VALUATION FRAMEWORK FOR STRUCTURED CONTRACTS

Swap Structure with European Style Option

Consider for the time being that the holder of the contract has only one right to cancel the contract, say on \( t_m < t_n \). At each \( t_i, i = 1, ..., m - 1 \) the customer pays the accruing costs and receives electricity worth the corresponding wholesale price \( F(t_i, t_i) \) times his realized demand. For simplicity, assume \( m = n - 1 \). Then \( D \cdot F(t_{n-1}, t_{n-1}) > AC \) means that the customer can realize a payoff of \( D \cdot F(t_{n-1}, t_{n-1}) - AC > 0 \). This argument becomes very clear if we assume that the customer can resell the electricity to another counterpart. Exercising this option does not mean to cancel the contract, but to keep it at least until the beginning of the next delivery period. In general, the payoff at time \( t_m \) is given by

\[
\left( \sum_{i=m}^{n-1} e^{-r(t_i-t_m)} (D \cdot F(t_i, t_i) - AC) \right)^+.
\]

(6.12)

\( AC \) is measured in cents, and hence \( F(t_m, t_i) \) has to be converted into ct/kWh, since it is usually noted in EUR/MWh. Also, \( K \) has to be converted, since it is a fixed amount on a yearly basis. With these considerations, (6.4) can be extended to

\[
P_{V_{t_0}} = \sum_{i=1}^{m-1} e^{-r(t_i-t_0)} (D \cdot F(t_0, t_i) - AC) + e^{-r(t_{m-1}-t_0)} \mathbb{E}_{t_0} \left[ \left( \sum_{i=m}^{n-1} e^{-r(t_i-t_m)} (D \cdot F(t_i, t_i) - AC) \right)^+ \right].
\]

(6.13)

The object of interest in (6.13) is the discounted expectation. For \( m = n - 1 \), this expectation simplifies considerably and can be evaluated using Black’s formula, as introduced above. The option is on \( D \cdot F(t_{n-1}, t_{n-1}) \) with strike \( AC = D \cdot k + K \) and maturity \( t_{n-1} \). Thus, as in (6.8), we obtain for the option price at time \( t_0 \)

\[
V_{t_0} = e^{-r(t_{n-1}-t_0)} \mathbb{E}_{t_0} \left[ (D \cdot F(t_{n-1}, t_{n-1}) - AC)^+ \right] = e^{-r(t_{n-1}-t_0)} (D \cdot F(t_0, t_{n-1}) \Phi(d_1) - AC \Phi(d_2)),
\]

and \( d_1, d_2 \) become

\[
d_1 = \frac{\log \frac{D \cdot F(t_0, t_{n-1})}{AC} + \frac{1}{2} \text{Var} [\log F(t_{n-1}, t_{n-1})]}{\sqrt{\text{Var} [\log F(t_{n-1}, t_{n-1})]}}
\]

\[
d_2 = d_1 - \sqrt{\text{Var} [\log F(t_{n-1}, t_{n-1})]},
\]

which are similar to (6.9)-(6.10). We employed the obvious fact that the variance term can be written as \( \text{Var} [\log D \cdot F(t_{n-1}, t_{n-1})] = \text{Var} [\log F(t_{n-1}, t_{n-1})] \). Furthermore, \( V_{t_0} \)
is given in cents, not in ct/kWh. The variance term takes several forms depending on the underlying model. In a simple Black model with dynamics
\[ dF(t, t_i) = \sigma F(t, t_i) dW(t), \]
the variance becomes,
\[ \text{Var} [\log F(t_{n-1}, t_{n-1})] = \sigma^2 (t_{n-1} - t_0). \]
If we use the two-factor dynamics introduced in Section 6.2.2, we obtain
\[ \text{Var} [\log F(t_{n-1}, t_{n-1})] = \frac{\sigma^2}{2\eta} \left( 1 - e^{-2\eta t_{n-1}} \right) + \sigma^2 t_{n-1}. \]
For general \( t_m < t_{n-1} \), however, the payoff in (6.12) contains the sum of log-normally distributed random variables and their distribution is unknown. One way to circumvent this problem is an approximation suggested by Lévy [94] and Brigo & Mercurio [25], where the random variable \( \hat{F}(t_j) := \sum_{i=j}^{n-1} e^{-r(t_i - t_j)} (D \cdot k + K) \) is approximated by another random variable \( X \), which is log-normally distributed and coincides in mean \( \hat{\mu} \) and variance \( \hat{\sigma}^2 \) with \( \hat{F}(t_j) \). With this approximation and \( \hat{K} := \sum_{i=j}^{n-1} e^{-r(t_i - t_j)} (D \cdot k + K) \) one can apply again Black’s formula to obtain the value of the option as
\[ V_{t_0} = e^{-r(t_j - t_0)} \mathbb{E}_0 \left[ \left( \hat{F}(t_j) - \hat{K} \right)^+ \right] \approx e^{-r(t_j - t_0)} \mathbb{E}_0 \left[ \left( X - \hat{K} \right)^+ \right] = e^{-r(t_j - t_0)} \left( \hat{F}(t_0) \Phi(d_1) - \hat{K} \Phi(d_2) \right), \]
with
\[ d_1 = \frac{\log \frac{\hat{F}(t_0)}{\hat{K}} + \frac{1}{2} \hat{\sigma}^2}{\hat{\sigma}}, \]
\[ d_2 = \frac{\log \frac{\hat{F}(t_0)}{\hat{K}} - \frac{1}{2} \hat{\sigma}^2}{\hat{\sigma}}. \]
The approximation has been proposed by Lévy [94] in the context of pricing options on arithmetic averages of currency rates. The advantage of the approximation over Monte-Carlo simulations is the difference in speed in which an option valuation can be carried out.

**Swap Structure with Bermudan Style Option**

We extend the previous consideration by endowing the retail contract with finitely many exercise dates. These can be minimum two exercise rights up to \( n - 1 \) exercise rights on
6.2. VALUATION FRAMEWORK FOR STRUCTURED CONTRACTS

Figure 6.3: Points in time under consideration. At $t_0$, we assume that the contracts are priced and issued immediately. In $(t_0, t_1)$, the customer can sign the agreement. $[t_1, t_2], ..., [t_{m-1}, t_m]$ are periods of certain delivery and $[t_m, ..., t_{n-1}]$ are the possible dates to cancel the agreement.

$t_1, ..., t_{n-1}$. Figure 6.3 depicts generally the life circle of a retail contract.

There are two ways to look at an early-exercise pricing problem; via dynamic programming or via optimal stopping. Formally, the latter leads to the former case. In the sequel, we let $t_1 \leq t_m \leq t_{n-1}$. We need the notion of a stopping time.

**Definition 6.2.2** Let $\mathcal{F}_{t_i}$ be the filtration generated by the underlying market dynamics (6.7). A random time $\tau : \Omega \to \{t_m, ..., t_{n-1}\}$ is an $\mathcal{F}_{t_i}$-stopping time if

$$\{\tau \leq t_i\} \in \mathcal{F}_{t_i},$$

for all $t_i \in \{t_m, ..., t_{n-1}\}$.

The state process is given by the forward dynamics (6.7) on a general state space $E = \mathbb{R}_+$, endowed with a $\sigma$-algebra. An element $x \in E$ is a realization of the forward price $F$. The state process is obviously Markovian. As long as the delivery contract is not being canceled, a measurable, discounted reward is obtained:

$$R_{t_i}(x) = e^{-r(t_i-t_m)} (D \cdot x - AC), \forall x \in E, \quad (6.14)$$

for $i = m, ..., n-1$. If the customer decides to cancel the contract, no reward is received. Once again, we would like to emphasize the difference with respect to more common options. Exercising means that we adhere to the exchange flow and we keep the option to interrupt the flow at all future dates. If we choose an arbitrary stopping time $\tau$ with $P[\tau \leq t_{n-1}] = 1$, then we obtain a discounted reward

$$\sum_{k=m}^{j(\tau)-1} e^{-r(t_k-t_m)} (D \cdot F(t_k, t_k) - AC) = \sum_{k=m}^{j(\tau)-1} R_{t_k}(F(t_k, t_k)), \quad (6.15)$$

where $j(\tau)$ is the index of $t \in \{t_m, ..., t_{n-1}\}$, where $\tau = t$. We want to find the value
CHAPTER 6. PRICING OF STRUCTURED RETAIL CONTRACTS

\[ V^*_m(x) := \sup_{\tau \in \{t_m, \ldots, t_{n-1}\}} E_{t_m x} \left[ \sum_{k=m}^{j(\tau)-1} R_{tk} \right], \text{ for } x \in E. \] (6.16)

\( E_{t_m x} [\cdot] \) denotes the conditional expectation given the information \( x \in E \) at time \( t_m \) (we sometimes omit the subscript \( x \)). Under mild regularity conditions, an arbitrage argument justifies calling (6.16) the option value, see Duffie [47], p.32 for further discussions.

A stopping time \( \tau^* \) is optimal if

\[ V^*_m(x) = E_{t_m x} \left[ \sum_{k=m}^{j(\tau^*)-1} R_{tk} \right], \text{ for } x \in E. \]

We intend to state the Bellman equation for this stopping problem via the Snell Envelope. To ensure the existence of this recursive formula we have to guarantee that

\[ \left( \hat{I} \right) = \sup_{\tau \in \{t_i, \ldots, t_{n-1}\}} E_{t_i} \left[ \sum_{k=i}^{j(\tau)} R^{+}_{tk}(F(t_k, t_k)) \right] < \infty, \text{ for } m \leq i \leq n - 1. \] (6.17)

Compare this condition with condition (I) in Chapter 5. Indeed, \( \left( \hat{I} \right) \) is fulfilled as the dynamics (6.7) produce finite first moments. Let us consider a process \((Z_t)\) with

\[ Z_{t_{n-1}} = 0 \]

\[ Z_{t_i} = \max \left\{ 0, D \cdot F(t_i, t_i) - AC + e^{-r(t_{i+1}-t_i)} E_{t_i} [Z_{t_{i+1}}] \right\}. \] (6.19)

This process is called Snell Envelope, see Snell [115] and Karatzas & Shreve [77]. In particular, we obtain \( Z_{t_m} = V^*_m(F(t_m, t_m)) \) as the value of the optimal stopping problem (6.16) and \( \tau^* := \inf \{t_i, i = m, \ldots, n - 1 : Z_{t_i} = 0\} \) is an optimal stopping time. Furthermore, \( Z_{t_i} \) is the smallest process, which satisfies

\[ Z_{t_i} \geq D \cdot F(t_i, t_i) - AC + e^{-r(t_{i+1}-t_i)} E_{t_i} [Z_{t_{i+1}}] \]

\[ Z_{t_i} \geq 0. \]

(6.18)-(6.19) are the Bellman equations for the stopping problem and well-posed as \( \left( \hat{I} \right) \) is fulfilled. In other words, we found a special stopping time that gives us the justification to write (6.16) inductively by means of dynamic programming:

\[ PV_{t_0} = \sum_{i=1}^{m-1} e^{-r(t_i-t_0)} \left( D \cdot F(t_0, t_i) - AC \right) + e^{-r(t_m-t_0)} E_{t_0} [V^*_m(F(t_m, t_m))], \] (6.20)

For \( m = n - 1 \) we obtain exactly (6.13), i.e. the European case. If we have two rights to cancel the contract, say for simplicity on \( t_{n-2} \) and \( t_{n-1} \), respectively, then (6.20) becomes
The net present value of the retail contract with cancel rights at time $t_i$ is
\[
CV_{t_i} = e^{-r(t_{i+1} - t_i)} \mathbb{E}_{t_i} \left[ (D \cdot F(t_{i+1}, t_{i+1}) - AC + CV_{t_{i+1}})^+ \right],
\]
and
\[
CV_{t_{n-1}} = 0
\]
\[
CV_{t_0} = e^{-r(t_{m} - t_0)} \mathbb{E}_{t_0} \left[ (D \cdot F(t_m, t_m) - AC + CV_{t_{m+1}})^+ \right].
\]

The net present value of the retail contract with cancel rights at $\{t_m, ..., t_{n-1}\}$ is concisely given by
\[
PV_{t_0} = \sum_{i=1}^{m-1} e^{-r(t_i - t_0)} (D \cdot F(t_0, t_i) - AC) + CV_{t_0}.
\]

The optimal exercise time becomes
\[
\tau^* = \inf \{t_i, i = m, ..., n - 1 : D \cdot F(t_i, t_i) - AC + CV_{t_i} \leq 0\}.
\]

These considerations show in a simple fashion the equivalence of an early exercise problem and dynamic programming (DP). Furthermore, the DP formulation (6.21)-(6.23) is better suited for numerical experiments. As encountered in Chapter 5, there are frameworks that cannot be solved explicitly. This is also true for the current setting and we suggest to apply a numerical approximation method known as least-squares Monte-Carlo. In the next section, we collect some properties of the net present value $PV_{t_0}$, which are helpful for our case study in Section 6.3.
6.2.4 Further Properties of $PV_{t_0}$

The retail contracts can in general consist of two parts, namely $m - 1$ periods of certain electricity delivery in $[t_1, t_2], ..., [t_{m-1}, t_m]$, and options to interrupt the delivery for the remaining $n - m$ periods $[t_m, t_{m+1}], ..., [t_{n-1}, t_n]$, see also Figure 6.3. Furthermore, pricing the retail contract does not mean assigning a value to the options in the contract. The price for the contract is reflected in the rates $K$ and $k$, which in the deterministic demand case can be summarized in $AC$. In other words, $AC$ has to be chosen such that the net present value $PV_{t_0}$ of the contract becomes zero.

In this section we turn to the question, how a utility should price a structured retail contract. We derive the necessary properties and methods to compute a fair contract value in form of a root search procedure and apply this methodology in the case study in Section 6.3. For this purpose we write the net present value dependent on the accruing costs:

$$PV_{t_0}(AC) := PV_{t_0},$$

where $PV_{t_0}$ has the form as in (6.24). Our main concern is to find an $AC^* \in [0, \infty)$ with $PV_{t_0}(AC^*) = 0$. Then $AC^*$ is the contract price as it is the fair price for both, the customer and the utility. First, we show that there exist a unique root of $AC \mapsto PV_{t_0}(AC)$ at all.

**Proposition 6.2.3** Assume $PV_{t_0}(AC)$ is defined as in (6.24). For arbitrary $n, m \in \mathbb{N}$ and $0 < m < n$, there exists a unique accruing cost $AC^* \in [0, \infty)$ such that

$$PV_{t_0}(AC^*) = 0.$$

**Proof:** We employ the intermediate value theorem. $AC \mapsto PV(t_0, AC)$ is continuous, which follows from (6.24) and a simple induction argument, see also the proof of Theorem 5.2.1. Furthermore,

$$PV_{t_0}(0) = \sum_{i=1}^{m-1} e^{-r(t_i - t_0)} D \cdot F(t_0, t_i) + e^{-r(t_m - t_0)} D \cdot E_{t_0} \left[ F(t_m, t_m) + \sum_{i=m+1}^{n-1} e^{-r(t_i - t_{i-1})} F(t_i, t_i) \right] > 0,$$

for any $m < n$, $r > 0$, and $D > 0$. This is intuitively clear, since a contract with no cost (electricity for free) must be very beneficial for the customer, i.e. $PV(t_0)$ should be positive. On the other hand,

$$\lim_{AC \to \infty} PV_{t_0}(AC) = -\infty < 0,$$
since the value of the option goes to zero and the first term in (6.24) is unbounded below. Hence, by the intermediate value theorem, at least one \( AC^* \) with \( PV_{t_0}(AC^*) = 0 \) exists. Lastly, we need to show that the net present value is monotonically decreasing in \( AC \). The easiest way to show this is to consider (6.16). \( R_{t_k} \) is strictly decreasing in \( AC \), the expectation operator is monotone, and the supremum does not affect the monotonicity. Thus, \( AC \mapsto PV_{t_0}(AC) \downarrow \). This makes \( AC^* \) unique and completes the proof. 

Figure 6.4 depicts for \( n = 4 \) and \( m = 2, 3, 4 \) the net present value (6.24) in dependence of \( AC \).

An elegant method to compute the root of \( PV_{t_0}(AC) \) is a fixed point iteration. A sufficient condition for the existence of a fixed point \( x^* \) of a real-valued, continuous, and differentiable function \( f(x) \) is \( |f'(x)| < 1, \ x \in \mathbb{R} \). Then \( \exists x^* \in \mathbb{R} \) with \( f(x^*) = x^* \). For our framework we define \( f(AC) := PV_{t_0}(AC) + AC \) and examine the cases where this function has a fixed point \( AC^* \). If we assume that we can find a unique fixed point, then it is the root of the net present value, and hence the fair price of the contract. As it turns out, only for the European case \( n = 3 \) and \( m = 2 \) we can employ a fixed point iteration. For arbitrary cases, especially for Bermudan cases, the net present value can probably not be transformed into a contracting function:

\[
f'(AC) = 1 - \left( e^{-r(t_1-t_0)} + e^{-r(t_2-t_0)}\Phi(d_2) \right) \in (-1, 1),
\]
and $d_2$ is given in (6.10). In other words, there is a certain delivery in $[t_1, t_2]$ with the option to interrupt the delivery for the last period $[t_2, t_3]$. We are looking for the accruing costs $AC$ that have to be charged in order to make $PV_{t_0}(AC) = 0$. The contraction $f$ has the form

$$f(AC) = e^{-r(t_1-t_0)} (D \cdot F(t_0, t_1) - AC) + e^{-r(t_2-t_0)} (D \cdot F(t_0, t_3) \Phi(d_1) - AC \Phi(d_2)) + AC,$$

by Equation (6.8). $d_1$ and $d_2$ are given in (6.9)-(6.10) and depend on $AC$. Consequently, the fixed point algorithm

$$f(AC_0) = AC_1$$
$$f(AC_1) = AC_2$$
$$\vdots$$
$$f(AC^*) = AC^*$$

provides the fixed point $AC^* = 7.58$ EUR/MWh, where $AC_0 \in (0, \infty)$ is chosen arbitrarily. Figure 6.5 depicts the convergence of $AC$, which is readily computed since $f$ is given in closed form.

![Figure 6.5: Convergence of AC in a simple Black model with $\sigma = 0.3$, $F(t_0, t_1) = 57.25$ EUR/MWh, $F(t_0, t_2) = 61.38$ EUR/MWh, $r = 0.03$ and $D = 125$ kWh. Initial value is $AC_0 = 7$ EUR.](image-url)
Note that the contract value without one right to cancel, i.e. certain delivery in \([t_1, t_2]\) and \([t_2, t_3]\) is with \(AC = 7.41\) EUR/MWh slightly lower. This is due to the fact that the contract is endowed with more flexibility for the customer, and thus this should be reflected in a higher price. More specifically,

\[
PV_{t_0}(AC) = e^{-r(t_1-t_0)} (D \cdot F(t_0, t_1) - AC) + e^{-r(t_2-t_0)} (D \cdot F(t_0, t_2) - AC) \\
\leq e^{-r(t_1-t_0)} (D \cdot F(t_0, t_1) - AC) + e^{-r(t_2-t_0)} \mathbb{E}_{t_0} \left[ (D \cdot F(t_2, t_2) - AC)^+ \right],
\]

which follows simply by the value bounds for European call options on forwards.

In order to shed light on the entire valuation procedure as applied in Section 6.3, we give a step-by-step example in the next section.

### 6.2.5 Illustrative Example

Consider a one-year contract with monthly rights to cancel. Assume we want to analyze and price the contract on December 1st (\(= t_0\)), which starts delivering electricity on January 1st (\(= t_1\)) until December 31st (\(= t_{13}\)), the following year. This contract contains 11 cancel rights, each on \(t_2, ..., t_{12}\). We do not take \(t_{13}\) into consideration, because as soon as \(t_{12}\) passes by, the customer forfeits all rights to cancel. Thus, there is no option left in the contract. Consequently, we set \(n = 13\) and \(m = 2\). The main steps of the pricing procedure are comprised of:

1. Construction of sufficiently long forward curves,
2. Simulation of forward price paths according to (6.7),
3. Application of the least-squares Monte-Carlo (LSMC) algorithm,
4. Solving \(PV_{t_0}(AC) = 0\) for \(AC\).

#### 1. Construction of Long Forward Curves

A long forward curve on monthly granularity is required, but not observable in the market. However, several monthly, quarterly, and yearly forwards are regularly quoted. We artificially construct a curve along the following lines: We take the available data of a forward curve provided by the EEX on a specific day (e.g. December 1st, 2008). Say, the December 2009 is not yet traded. To obtain a price for this month we calculate the ratio between the December 2008 contract and the 2008 yearly contract and multiply the resulting factor with the 2009 yearly contract. The subsequent month can be calculated in the same manner. Though, more sophisticated methods are explained, e.g. in Benth et al. [14] or Fleten & Lemming [59], we work with this basic approach as it suffices to illustrate our ideas. Table A.7 and Figure 6.6 show the constructed forward curves that are used for the case study in Section 6.3.
CHAPTER 6. PRICING OF STRUCTURED RETAIL CONTRACTS

2. Simulation

In order to apply the LSMC, we have to simulate paths along the discrete time line \( t_0, \ldots, t_{12} \), where at time \( t_0 \) we observe the initial forward curve \( F(t_0, t_i) \). Recall our proposed two-factor model in (6.7). The discrete-time version reads as

\[
\log F(t, t_i) = \log F(t_0, t_i) - \frac{\sigma_1^2}{4\eta} \left( e^{-2\eta(t_i-t)} - e^{-2\eta t} \right) - \frac{t\sigma_2^2}{2} \varepsilon_j
\]

\[
= \mu(t, t_i) + \sqrt{\frac{\sigma_1^2}{2\eta}} \left( e^{-2\eta(t_i-t)} - e^{-2\eta t} \right) Z_1 + \sigma_2 \sqrt{T} Z_2,
\]

with \( Z_j \sim \mathcal{N}(0, 1) \), \( j = 1, 2 \) being independent random variables. The forward contracts, which are due at the beginning of each month are denoted by \( F(t, t_1), \ldots, F(t, t_{12}) \). The simulated paths can be seen as a lower triangle matrix, which we refer to as the path matrix. For a fixed path \( \omega \) it reads

\[
\begin{pmatrix}
F(t_0, t_0) & 0 & \cdots & 0 \\
F(t_0, t_1) & F(\omega, t_1, t_1) & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots \\
F(t_0, t_{12}) & \cdots & \cdots & F(\omega, t_{12}, t_{12})
\end{pmatrix}
\]

(6.25)

\( F(\omega, \ldots) \) denotes the realized forward price from the simulation. The first column \( F(t_0, \cdot) \) corresponds to the initial forward curve observable in the market and is used as an initial value. It is important to notice that on a specific grid point \( t_i \), the contracts (more specifically, the paths) share the same Brownian shock \( W^{(1)}(t_i), W^{(2)}(t_i) \).

3. Least-Squares Monte-Carlo

The expectations in (6.24) cannot be calculated explicitly. This is a well-known obstacle when it comes to pricing options with early-exercise features. The seminal paper on LSMC applied to Bermudan options has been published by Longstaff & Schwartz [92] and has been discussed in Egloff [50] and Protter et al. [38]. We fit this algorithm to our purposes, but present only the main idea, since technical details apply identically as in the mentioned publications. The starting point is equation (6.24). For simplicity, let \( m = 1 \). The general case \( m > 1 \) can be treated by discounting appropriately. \( \Omega \) is the set of all possible realizations of underlying stochastic system. A typical element \( \omega \in \Omega \) represents a sample path of the forward prices. The aim is to find an approximation for the ‘real’ option value: \( \hat{C}V_{t_0} \approx CV_{t_0} \). For a particular sample path \( \omega \), the continuation value can be written as

\[
CV_{t_i}(\omega) = \mathbb{E}_{t_i} \left[ \sum_{j=i+1}^{n-1} e^{-r(t_j-t_i)} (D \cdot F(\omega, t_j, t_j) - AC) \right].
\]
The cash flows are conditional upon the contract has not been canceled at or prior to time $t_i$. At any time $t_i$ the continuation value can be represented as a linear combination of a countable set of $\mathcal{F}_t_i$-measurable basis functions. This is justified by the fact that $CV_{t_i}$ is an element of the space of square-integrable functions, denoted by $L^2(\Omega, \mathcal{F}, Q)$. $L^2$ is a Hilbert space, and thus has a complete countable orthonormal basis. More specifically,

$$CV_{t_i}(\omega) = \sum_{j=0}^{\infty} a_j \phi_j(X),$$

where we choose $X$ to be the forward curve at some point in time. Other information are also conceivable, such as the swap rate, which can be obtained from the forward curve on a specific day. $a_j$ are constant coefficients and $\phi_j$ are basis functions that can be chosen from a large class of functions. Longstaff & Schwartz [92] report that simple polynomials perform exclusively well. We encountered a similar performance and choose to approximate the continuation value with a cubic polynomial:

$$\phi_0(X) = 1 \quad \phi_1(X) = X \quad \phi_2(X) = X^2 \quad \phi_3(X) = X^3.$$ 

The continuation value becomes approximately

$$CV_{t_i}(\omega) \approx CV_{t_i}^{\text{approx}}(\omega) = a_0 + a_1 X_{t_i}(\omega) + a_2 X_{t_i}^2(\omega) + a_3 X_{t_i}^3(\omega).$$

$CV_{t_i}^{\text{approx}}(\omega)$ is estimated by projecting the discounted cash flows onto the basis functions. We only use paths where the option is in the money as they are the relevant ones. This yields estimates $\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3$. We denote the estimated continuation value by $\hat{CV}_{t_i}(\omega)$. It can be shown that $\hat{CV}_{t_i}(\omega)$ converges to $CV_{t_i}^{\text{approx}}(\omega)$ in mean-square and probability as $N \to \infty$, see Longstaff & Schwartz [92]. The approximated continuation value is used to decide whether or not to stop at some point in time. Assume we have approximations of all continuation values available. Consider the last right to cancel $t_{12}$, which is on December 1st of the following year. A rational customer stays in the contract if

$$D \cdot F(\omega, t_{12}, t_{12}) > AC,$$

because the customer pays less compared to the price for electricity in the market. In $t_{11} =$ November 1st, a rational customer stays the contract if the current forward price $F(t_{11}, t_{11})$ plus the expected residual value $\hat{CV}_{t_{11}}(\omega)$ exceeds the accruing costs:

$$D \cdot F(\omega, t_{11}, t_{11}) + \hat{CV}_{t_{11}}(\omega) > AC.$$
CHAPTER 6. PRICING OF STRUCTURED RETAIL CONTRACTS

Going back in time until we reach \( t_1 \). The value of the option for a specific path \( \omega \) is denoted by \( CV_{LSMC}(\omega) \). To obtain the overall value for the options within the retail contract, we have to average over all paths:

\[
CV_{t_0} = \frac{1}{N} \sum_{i=1}^{N} CV_{LSMC}(\omega_i).
\]

Now let \( m \in \mathbb{N} \) be arbitrary. The value of the delivery contract is approximately given by

\[
PV_{t_0}(AC) = m - 1 \sum_{i=1}^{m-1} e^{-r(t_i-t_0)} (D \cdot F(t_0,t_i) - AC) + CV_{t_0} \approx m - 1 \sum_{i=1}^{m-1} e^{-r(t_i-t_0)} (D \cdot F(t_0,t_i) - AC) + \tilde{CV}_{t_0},
\]

compare (6.24). This value determines the fairness of the contract. \( PV_{t_0}(AC) > 0 \) means that the contract is advantageous for the customer, i.e. too cheap from a utility’s point of view. A negative present value corresponds to an expensive contract. In Proposition 6.2.3 we proved that \( PV_{t_0}(AC) \) is a decreasing function in \( AC \). This implies that, e.g. a positive present value can be made zero by increasing the charges \( AC \) against the customer. For the analysis of the real-world contracts we take \( K \) and \( k \) as given and under an assumption on the demand rate \( D \) we can examine if the contracts are priced in an (un-)fair fashion. In a second step, we show how to price such a contract, i.e. finding \( K \) and \( k \) such that \( PV_{t_0}(AC) = 0 \).

Convergence Results

Clément et al. [38] prove that the LSMC algorithm converges weakly when the number of simulated paths goes to infinity, given that the number of basis functions \( \phi_j \) is sufficiently large. The limiting price distribution is even Gaussian. In particular,

\[
\lim_{N \to \infty} P \left[ \left| CV_{t_0} - \frac{1}{N} \sum_{i=1}^{N} CV_{LSMC}(\omega_i) \right| > \varepsilon \right] = 0.
\]

We employed several types of basis functions \( \phi_j \). Popular choices are Laguerre, Hermite, Legendre, and Jacobi polynomials. As already reported in Longstaff & Schwartz [92], simple polynomials perform surprisingly well. This coincides with our experience and we choose a cubic polynomial along with \( N = 10,000 \) simulated forward price curves to analyze the structured contracts.

The pricing algorithm is implemented in Matlab. Variance reduction techniques as well as vector-based coding would lead to either faster evaluations, given a required numerical accuracy, or higher accuracy in the simulation results within the same time frame. However, since we only analyze a small number of contracts using a few forward
6.3. CASE STUDY

curves we encountered acceptable effort. It takes about 2 minutes to run the algorithm on the contract described above. The algorithm includes the simulation of the paths, the regression of the continuation values, and the evaluation of the contract. For the analyzed contract we generated 10,000 paths and evaluated the continuation value at each node. We encountered stable results in the first decimal point. For this reason all numbers presented in the sequel are rounded to 1 digit after the decimal point. In Section 6.3 the numbers for the net present value (net present value) and the option premium have to be interpreted as ct/kWh ± 0.1 ct/kWh.

4. Root Search Procedure

We employ the Secant Method as the classical Newton Method is not applicable due to the lack of explicit derivatives of the net present value. Choosing $AC_0 = 0$ and $AC_1 > 0$ sufficiently large, we can calculate

$$AC_k = AC_{k-1} - PV_{t_0} (AC_{k-1}) \frac{AC_{k-1} - AC_{k-2}}{PV_{t_0} (AC_{k-1}) - PV_{t_0} (AC_{k-2})},$$

for $k = 2, \ldots$, until the difference $|AC_k - AC_{k-1}| < \varepsilon$ for some small $\varepsilon > 0$. The convergence is surprisingly fast because the slope of the net present value is steep and the convexity is moderate, compare Figure 6.4.

6.3 Case Study

We have to point out one important thing. The case study is conducted for a different unit on the value of the contracts. In the previous sections we derived the framework that provides quantities (such as net present values) the absolute monetary value. Since the advertised contracts and, in general, all market prices are always noted in cents per kWh (or EUR per MWh), we adapt this convention. This can simply be performed by dividing all previously introduced quantities by the deterministic demand $D$.

6.3.1 Example Contracts

Let us introduce three example contracts that were publicly available by the end of 2008. As described in the introduction, it was of utmost importance to gather as much relevant information about these contracts as possible. This was accomplished by using publicly available sources such as the online websites and telephone hotlines. In other words, all data that is used for this analysis is a result of our own research. For reasons of data protection we do not name the three different utility companies explicitly, but attach the letters $A$, $B$, and $C$ to those companies. We collected all contract information as follows: Concession levies (license fees) and fees for network access differ locally in Germany. Thus, we fix a region as a reference in southern Germany. Moreover, the utility companies offer contracts dependent on the yearly consumption of electricity. Therefore, we assume a typical single person household with a yearly consumption of...
1500 kWh \((D = 125 \text{ kWh})\). The yearly demand rate is for simplicity distributed evenly into the variable kilowatt-hour rate, i.e. we add

\[
\text{fixed yearly price } K \text{ in ct,}
\]
\[
\frac{1500 \text{ kWh},}{\text{1500 kWh}},
\]

(6.26)

to the variable price \(k\). We call this the \textit{modified kilowatt-hour rate} (modified kWh-rate). As a basis for our analysis we take the specification of three different real-world retail contracts into consideration.

\textbf{Contract A}

This contract was first offered by the end of 2008, we set \(t_0 = \text{December 1st, 2008}\). It contains a price guarantee until 12/31/2011 and yearly call rights, i.e. the first call right at the end of 2009 and the second call right at the end of 2010. Signing on 01/01/2009 implies a price warranty 3 years. The kWh-rate amounts to \(k = 21.9 \text{ ct/kWh}\) and the demand rate is \(K = 92.82 \text{ EUR/year}\) yielding a modified kWh-rate of \(AC = 28.1 \text{ ct/kWh}\) according to (6.26).

\textbf{Contract B}

In contrast to the A contract, the B contract offers a one year price guarantee and monthly call rights. In other words, the price guarantee is independent of the point in time when the customer signs the contract. This contract was issued in November 2008, i.e. we set \(t_0 = \text{November 1st, 2008}\). The kWh-rate amounts to \(k = 18.5 \text{ ct/kWh}\) and the demand rate is \(K = 90.00 \text{ EUR/year}\) yielding a modified kWh-rate of \(AC = 24.5 \text{ ct/kWh}\).

\textbf{Contract C}

In this case, the company offers again monthly call rights but the price guarantee is fixed by the end of June 2011. I.e., as for contract A we have to take the signing date into consideration. This offer was issued in October 2008, i.e. we set \(t_0 = \text{October 1st, 2008}\). The kWh-rate amounts to \(k = 19.9 \text{ ct/kWh}\) and the demand rate is \(K = 118.80 \text{ EUR/year}\) yielding a modified kWh-rate of \(AC = 27.8 \text{ ct/kWh}\). Table 6.1 summarizes the essential features of the contracts under consideration.

<table>
<thead>
<tr>
<th>Contract</th>
<th>Call Dates</th>
<th>Price Guarantee</th>
<th>Demand Rate (K) in EUR/Year</th>
<th>kWh-Rate (k) in ct/kWh</th>
<th>Mod. kWh-Rate in ct/kWh</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yearly</td>
<td>12/31/2011</td>
<td>92.82</td>
<td>21.9</td>
<td>28.1</td>
</tr>
<tr>
<td>B</td>
<td>Monthly</td>
<td>1 Year</td>
<td>90.00</td>
<td>18.5</td>
<td>24.5</td>
</tr>
<tr>
<td>C</td>
<td>Monthly</td>
<td>06/30/2011</td>
<td>118.80</td>
<td>19.9</td>
<td>27.8</td>
</tr>
</tbody>
</table>

Table 6.1: Summary of contracts offered by several utility companies in Germany assuming a constant consumption of 1500 kWh/year; the contracts were placed in the market by the end of 2008.
The contracts have quite different specifications and span a reasonable variety of possible contracts. In the sequel we show how the quantities influence the contracts’ values for the customer. Furthermore, we analyze each contract based on the curve where it was published. I.e. we assume that the contracts are signed within \([t_0, t_1]\). More specifically, we assume that contract A was signed in December 2009, contract B was signed in November 2009, and contract C was signed in October 2009.

### 6.3.2 Calibration Results

We described the underlying model for the forward market and the calibration procedure in Section 6.2.2. The options used for calibration are given in Table 6.2 and the calibrated parameters are listed in Table 6.3. These are reasonable in the sense that a long-term volatility of 18\% for forward contracts are implied, while fluctuations increase up to 63\% short before maturity of a forward contract. This is in line with the study in Börger [23].

<table>
<thead>
<tr>
<th>Product</th>
<th>Delivery Start</th>
<th>Strike in EUR</th>
<th>Forward Price in EUR</th>
<th>Option Price in EUR</th>
<th>Implied Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>November 08</td>
<td>93</td>
<td>93.15</td>
<td>5.506</td>
<td>50.36%</td>
</tr>
<tr>
<td>Month</td>
<td>December 08</td>
<td>87</td>
<td>87.00</td>
<td>6.472</td>
<td>45.91%</td>
</tr>
<tr>
<td>Month</td>
<td>January 09</td>
<td>94</td>
<td>94.05</td>
<td>8.084</td>
<td>43.20%</td>
</tr>
<tr>
<td>Month</td>
<td>February 09</td>
<td>94</td>
<td>93.99</td>
<td>8.969</td>
<td>41.75%</td>
</tr>
<tr>
<td>Month</td>
<td>March 09</td>
<td>84</td>
<td>83.50</td>
<td>8.046</td>
<td>39.16%</td>
</tr>
<tr>
<td>Quarter</td>
<td>July 09</td>
<td>67</td>
<td>70.32</td>
<td>8.633</td>
<td>29.82%</td>
</tr>
<tr>
<td>Quarter</td>
<td>October 09</td>
<td>76</td>
<td>81.87</td>
<td>12.049</td>
<td>29.28%</td>
</tr>
<tr>
<td>Quarter</td>
<td>January 10</td>
<td>85</td>
<td>84.50</td>
<td>9.778</td>
<td>27.60%</td>
</tr>
<tr>
<td>Quarter</td>
<td>April 10</td>
<td>65</td>
<td>64.52</td>
<td>7.625</td>
<td>26.08%</td>
</tr>
<tr>
<td>Year</td>
<td>January 09</td>
<td>78</td>
<td>77.29</td>
<td>3.565</td>
<td>25.33%</td>
</tr>
<tr>
<td>Year</td>
<td>January 10</td>
<td>74</td>
<td>74.25</td>
<td>7.397</td>
<td>22.89%</td>
</tr>
</tbody>
</table>

Table 6.2: Market observed at-the-money call options and implied volatilities. All prices are in EUR. Date of calibration is 09/30/2008.

There is a pronounced maturity effect observable for the forward contracts (implied volatility in Table 6.2). This is in contrast to the yearly delivery contracts scrutinized in a different context in Chapter 2. We want to point out that the market for options on electricity forwards is still fairly thin. It is difficult to find a reasonable number of at-the-money options on a specific day. In our situation we are fortunate, since calibrating the model based on market data on 09/30/2008 was necessary.
CHAPTER 6. PRICING OF STRUCTURED RETAIL CONTRACTS

\[
\begin{array}{c|c|c}
\sigma_1 & \sigma_2 & \eta \\
0.45 & 0.18 & 1.55 \\
\end{array}
\]

Table 6.3: Calibration results from the forward model introduced in Section 6.2.2.

6.3.3 Analysis of Contracts without Call Rights

As pointed out in the previous section, each contract can be viewed as a fixed-for-floating swap equipped with rights to interrupt the delivery where the fixed leg is the payment in EUR/MWh (assuming a deterministic energy demand \(D\)) and the floating leg is the delivery of electricity with varying value measured in EUR/MWh (= 0.1 ct/kWh). The fair swap rate (time to maturity \(n\) months, interest rate of \(r\), neglecting call rights) can be calculated by

\[
\hat{AC} = \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)}F(t_0, t_i)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)}},
\]

compare (6.6). Based on the prevailing power base forward curve as of December 1st, 2008 and with \(n = 36\), \(r = 3\%\), this yields 16.2 ct/kWh. As the typical household consumes with a particular load profile and not constantly at baseload hours, a mixture between baseload and peakload must be assumed. Performing the calculation with baseload forwards and peakload forwards yields the interval

\[
[\hat{AC}_b, \hat{AC}_p] := [16.2 \text{ ct/kWh}, 23.8 \text{ ct/kWh}].
\]

This interval should cover the fair retail price for a fixed price contract when no call rights are embedded (with a large probability). This can be interpreted as a confidence interval (though the probability level is, of course, not known). Compare also Table A.7 in the appendix to comprehend the numbers presented here. The difference between baseload and peakload is explained in Chapter 1.

To obtain more realistic intervals in the sequel, one could perform the analysis based on a standardized load profile, which reflects typical consumption profiles of household customers. Since we want to show the general behavior of such contracts, we omit a detailed analysis based on standardized load profiles in order to gain focus on the main ideas. As a general rule experienced in practice, the typical household consumption is slightly biased to peak hours.

As each of the contracts has been structured and put into the retail market at different times, it is somehow ‘unfair’ to compare the structures at only one historical date. For example, on December 1st, 2008, the retail contracts B and C had already been published for quite some time, while A had just been issued. Thus, we provide results
for the initial publication date of each retail contract, i.e. for October 1st, 2008 (C), November 1st, 2008 (B), and for December 1st, 2008 (A). Figure 6.6 depicts the life time of the contracts under consideration and other market information.

Corresponding to the maturity of each contract, we can calculate the fair price range for each contract omitting the call rights for the moment. The results are presented in Table 6.4, column 5. The numbers so far are independent of a particular model. As we have not included the call rights yet, which are of additional value to the customer, it seems reasonable that the modified kWh-rate offered by suppliers does not necessarily fall into the fair modified kWh-rate interval (cp. Table 6.4, column 5). However, incorporating the options yields a different interval for each contract and for each forward curve and we expect the (offered) modified kWh-rates to fall into those intervals.

### 6.3.4 Analysis of Contracts with Call Rights

In order to evaluate the multiple call rights held by the retail customer we need the model described in Section 6.2.2. The pricing algorithm returns a price (more specifically, a net present value) of the structured contract. In case of a fair, i.e. value-neutral setup, this net present value should be zero. If it is positive, it is advantageous for the customer. If it is negative, the supplier might have put on a sales margin. Using the forward model for wholesale prices and translating the results into retail prices, we obtain numbers summarized in Table 6.4, columns 6 and 7. Note that each contract is analyzed on a
basis of three different forward curves. Each curve corresponds to the initial offers of one of the contracts. Again, the intervals that we list correspond to base and peak curves as input to our model and should indicate a realistic confidence interval.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct</td>
<td>A</td>
<td>38</td>
<td>28.1</td>
<td>[20.8, 29.8]</td>
<td>[12.5, 96.2]</td>
<td>[19.8, 94.4]</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>12</td>
<td>24.5</td>
<td>[21.7, 30.6]</td>
<td>[15.2, 88.0]</td>
<td>[18.0, 81.9]</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>33</td>
<td>27.8</td>
<td>[21.0, 30.0]</td>
<td>[6.8, 147.9]</td>
<td>[13.6, 145.7]</td>
</tr>
<tr>
<td>Nov</td>
<td>A</td>
<td>37</td>
<td>28.1</td>
<td>[18.5, 26.9]</td>
<td>[3.5, 51.6]</td>
<td>[13.1, 52.8]</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>12</td>
<td>24.5</td>
<td>[18.3, 26.4]</td>
<td>[−2.5, 45.6]</td>
<td>[3.7, 43.7]</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>32</td>
<td>27.8</td>
<td>[18.6, 27.0]</td>
<td>[−6.5, 83.0]</td>
<td>[2.7, 83.8]</td>
</tr>
<tr>
<td>Dec</td>
<td>A</td>
<td>36</td>
<td>28.1</td>
<td>[16.2, 23.8]</td>
<td>[−4.8, 26.6]</td>
<td>[7.1, 31.2]</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>12</td>
<td>24.5</td>
<td>[15.5, 22.4]</td>
<td>[−8.0, 9.7]</td>
<td>[1.0, 11.8]</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>31</td>
<td>27.8</td>
<td>[16.0, 23.5]</td>
<td>[−11.8, 12.6]</td>
<td>[0.0, 17.3]</td>
</tr>
</tbody>
</table>

Table 6.4: This table contains the offered modified kWh-rates, the fair modified kWh-rates (excl. options), the net present values and the values of the options (option premia).

Basic Analysis Contract A

Lets focus on the results based on the December curve. The net present value ranges from -4.8 to 26.6 ct/kWh. A positive net present value means that the contract is beneficial for the customer. Assuming that the customer is consuming 1 kWh each month at peak prices, he pays 28.1 ct each month but obtains energy with monetary value of 22.8 ct in the first month (on average), 21.2 ct in the second and so on (cp. the forward curves in Table A.7). Averaged over the entire life time of the contract, the power delivery has a value of 23.8 ct/kWh. Arriving at the first call date he rationally cancels the contract if the expected future value of the remainder of the delivery contract is negative for him. Depending on the simulated path this might be sooner or later. After discounting and averaging over all paths, he would obtain 26.6 ct more value (in form of electricity) than he has paid for.

This 26.6 ct might realize over the course of all 36 months indicating that he is making a profit of 26.6 ct/36 months = 0.7 ct/month and kWh. Using the initially assumed consumption of 1500 kWh per year, he has an advantage of 125 kWh×0.7 ct/(kWh and month) = 92.5 ct/month and 34.23 EUR in total over the course of three years. If we assumed that base prices are more realistic, the utility would realize a margin of 5.98 EUR in total (16.6 ct/month). Neglecting the value of call rights, base and peak price assumptions leads to a non-profitable evaluation of the contract from a customer’s point of view (consuming energy on average for 23.8 ct/kWh and paying 28.1 ct/kWh). It is
then advisable to look for an alternative supplier.

Considering the exercise rights, the deal is much more attractive for the customer. In fact, based on peak prices the intrinsic loss is \(23.5 - 28.1 \text{ ct/kWh} = -4.6 \text{ ct/kWh}\). Compared to the profit of 26.6 ct/kWh, the additional value generated by the call rights amounts to 31.2 ct/kWh over the course of the 36 months (only 7.1 ct/kWh when considering base prices). We would like to emphasize that based on a rational exercise strategy the value of the options can be quite large. It is up to the customer if he wants to pay that much for the right to call the supply contract.

**Basic Analysis Contract B**

As B was issued in November, we base the discussion now on the November forward curve. The net present value ranges from -2.5 to 45.6 ct/kWh, an even wider range, though this profit/loss is distributed over 12 months only. This stems from the fact that even without call rights, the modified kWh-rate is already within the interval (at-the-money) and the customer has not only two dates at which he can call (as in the case of A), but 11. As soon as the contract turns against the customer, he can quit, generating no losses but only profits. Here, it is even more questionable whether the utility company receives a proper margin. However, one should not forget that the risk profile of a fixed price delivery contract over 12 months is quite different compared to a 36 months contract.

**Basic Analysis Contract C**

For the analysis of contract C we focus on the October forward curve. While the offered modified kWh-rate (27.8 ct/kWh) lies already within the interval of base/peak prices when no call option is considered (similar to B, in contrast to A), the inclusion of call rights shifts the profitability entirely in favor of the retail customer, yielding a profit of 6.8 ct/kWh (147.9 ct/kWh), spread over 33 months.

**Comparison among Contracts**

The offered modified kWh-rate is 28.1 ct/kWh (A), 27.8 ct/kWh (B), and only 24.5 ct/kWh for contract C. Typically, services offering comparison between different electricity supply contracts take this quantity as the main criterion for evaluation of a contract (more precisely, they calculate the actual cost according to a specified consumption) and mention other terms and conditions only qualitatively. Doing so, the ranking would be (1) B, (2) C, and (3) A. It is up to the customer to assign a value to the condition of being allowed to leave the contract at certain times and having a fixed price for some time period. Based on the net present value of our analysis, the ranking is by far not that obvious. Would all contracts have existed already in October and November, contracts A and B were quite similar and contract C not really arrangeable. In December, contract A would even take the lead as it is most profitable to the cus-
tomer in both cases, base and peak. B and C would share ranks two and three. This is a quite different picture.

Analysis of Calling Probabilities

![Graph showing cumulative cancel probabilities for actual offered contracts provided by the LSMC algorithm.](image)

Figure 6.7: Cumulated cancel probabilities for the actual offered contracts provided by the LSMC algorithm.

The pricing algorithm delivers the distribution of calling probabilities over the dates as a side. We take each path and determine the point in time $t_i$ where it was reasonable to cancel the contract. Doing this for all paths at time $t_i$ and dividing the resulting number of paths by the total number of simulated paths $N$ yields the calling probability at time $t_i$.

For brevity, we consider each contract with the data at its published date. Figure 6.7 shows the cumulated cancel probabilities. The result is twofold: If the analysis is based on base prices, the cumulative probabilities exhibit a concave shape. That means, the customer is likely to cancel the contract at the beginning of the contract life. This is very pronounced in the case of contracts A and B and can be explained as follows. The (intrinsic) net present value in Table 6.4 is negative for the base case, i.e. the contract is unfavorable for the customer and he probably cancels early. For example, the A contract reveals a cancel probability of about 88% after the first year. In the case of B, we obtain a cancel probability of about 50% after the first month. Contract A would be canceled in only 66% of the cases (compared to the roughly 88% from before).

If one knew that a particular percentage of customers does not behave rational and does
not quit optimally, it would be a starting point to simply multiply the probability curves by a fundamental factor, which represents the non-rationality. Such a modified curve could serve as a tool when analyzing scenarios during the structuring of such contracts. This would also be a relevant quantity for suppliers hedging their portfolio, taking into account that customers might switch and load has not to be bought in wholesale forward markets.

The probabilities are conservative estimates for the utility company as this requires optimal exercise by the customer, which is not the case. If the customer exercises suboptimally, he can

- exercise, even though the swap is profitable for him. If a customer is actively thinking about the delivery contract and willing to change, then he is price-sensitive and this event is very unlikely. If this still happens (e.g. due to competitors’ marketing), such canceling is favorable for the utility company from a purely financial point of view.

- not exercise, even though the swap is unfavorable for him. Due to laziness, this is the more likely scenario. Thus, the probability of 30% is rather an upper bound.

It is possible to analyze the customers’ laziness and find a, say 20%, which exercise suboptimally and multiply the 30% from above by 0.8. This yields prognosis of 24% of customers who cancel their contracts at the end of the first year, seen as of today. This can be accounted for when hedging such options. Such an analysis assumes that irrational customers act suboptimally whatever happens, even though the contract might be far out of the money. It seems more realistic to assume that a customer acts efficiently when the potential gain is large enough. Let’s take the standardized example of a one person household with a demand of 1500 kWh per year. He would change his supplier if he saved at least, say 50 EUR per year. Below that amount he is too lazy. This means, he would accept an additional

\[
\frac{50 \text{ EUR}}{1500 \text{ kWh}} = 0.3 \text{ ct.}
\]

What is the probability of such a customer exercising the contract? This can be incorporated easily into the framework as an additional strike. While for the rational customer the payoff reads

\[
\text{Exercise when the NPV of remaining contract is negative.}
\]

The suboptimal investor might act according to

\[
\text{Exercise when the NPV + 0.3 ct is negative.}
\]
6.3.5 Value-Neutral Specifications

In this section, we show how to price retail contracts such that they are fair for both, the customer and the utility. In other words, we seek an \( \hat{AC}^* \) which makes \( PV_{t_0}(\hat{AC}^*) = 0 \). From a utility company’s perspective one could use our proposed model to find a fair price as a first step and add some margin for profit in a second step. Since we do not want to compare the contracts here, we use the specific curve for each contract where it was published initially. While the fair modified kWh-rate \( \hat{AC} \) produces a zero net present value for the non-callable swap (base/peak), the augmented rate includes the value of options. Table 6.5 shows the fair price intervals \( [\hat{AC}_b, \hat{AC}_p] \) and compares it with the actual offered contracts.

<table>
<thead>
<tr>
<th>Contract</th>
<th>Offered Modified kWh-Rate [ct/kWh]</th>
<th>Fair Modified kWh-Rate [ct/kWh]</th>
<th>Augmented Modified kWh-Rate [ct/kWh]</th>
<th>Value of Options [ct/kWh]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>28.1</td>
<td>[16.2, 23.8]</td>
<td>[25.3, 36.8]</td>
<td>[9.1, 13.0]</td>
</tr>
<tr>
<td>B</td>
<td>24.5</td>
<td>[18.3, 26.4]</td>
<td>[23.5, 33.5]</td>
<td>[5.2, 7.2]</td>
</tr>
<tr>
<td>C</td>
<td>27.8</td>
<td>[21.0, 30.0]</td>
<td>[29.4, 41.4]</td>
<td>[8.4, 11.3]</td>
</tr>
</tbody>
</table>

Table 6.5: The fair modified kilowatt-hour rate incl./excl. options would make the contract with/without call rights have zero net present value.

Recall that the fair modified rate is the sum of the augmented modified rate and the value of the options to interrupt the delivery. Contract A offered a delivery price of 28.1 ct/kWh. This value lies within the augmented modified price range of [25.3, 36.8], compare Table 6.5. In other words, dependent on the base-peak mixture, this contract could be slightly advantageous for the customer or for company A. In any case, the price appears reasonable from our model point of view.

Company B, on the other hand, offers 24.5 ct/kWh. This value is also within the corresponding interval of [23.5, 33.5], however, fairly close to the lower bound. That means the contract is quite cheap and therefore very profitable from a customers point of view. This could be due to the reason that the customer does not exercise optimally, but suboptimally and B might already take this into consideration.

The C contract offers 27.8 ct/kWh. This value is outside and below the interval [29.4, 41.4]. This is in line with the results in Table 6.4. Contract C is very cheap. This observation is reasonable because the contract guarantees monthly call rights for the next 33 months.
6.4 Conclusions & Discussion

In this chapter we showed how to price structured retail contracts. To our knowledge, this is the first contribution of this kind to the academic literature. We develop a framework to price Bermudan-style options on swaps. This includes in a first step the analysis of the cash flows and in a second step the augmentation of these flows with options. The contract structure differs in several ways from classical Bermudan (or American) options. Thus, the framework is tailored to our needs and we derive necessary properties of the net present value $PV$, which justify the application of a root search procedure. We analyze real-world contracts and compare their value (as advertised) to the value produced by our model framework. Specifically, we are able to assign monetary values to these structured products and draw valuable conclusions about the pricing mechanism.

In Section 6.3.5 we show how to price the delivery contracts initially. This includes the monetary valuation of the options embedded in the contracts via the least-squares Monte-Carlo algorithm and the determination of the charge $AC = D \cdot k + K$ against the customer via the root search procedure.

The most pronounced shortcoming of the model approach in this chapter is the assumption of a deterministic customer demand. Utility companies face several types of risk and volume risk is doubtlessly one of the most severe risk sources (default risk might be the other big source of risk). Consequently, this uncertainty in volume has to be regarded when it comes to pricing structured products. Apart from the pricing, hedging the contracts is as important. However, here we face an additional difficulty as forward markets do not suffice to hedge volume risk exposure. This is still future research. In the next chapter we discuss an appropriate model approach that incorporates a stochastic demand into the pricing framework.

6.5 Extensions: Incorporating Volume Risk

A major criticism of the pricing framework developed in this current chapter is the strong simplification of a deterministic and constant demand rate. We now elaborate on this issue by imposing a stochastic process that describes the random fluctuations of the electricity consumption of the customer. Utility companies label the uncertain demand as volume risk. Volume risk in electricity markets has barely received any attention in the academic literature. Utility companies are naturally aware of the potential risk incurred by volume uncertainty and most publications focus on the qualitative description of this circumstance, see e.g. Lemming [90]. Volume risk is especially central in the electricity sector. We contribute to the academic literature by discussing potential modeling approaches. While risk management application and hedging are as important as pricing under volume risk, we focus here merely on the latter issue. We conduct a simulation study that assigns a monetary value to the options that are included in a typical retail contract. We compare the results to the deterministic demand case and
argue that even in pricing such contracts, stochastic demand is a central quantity that should be taken into consideration as it has an influence of the contracts value.

**Model Approaches for Stochastic Demand**

Assume that the demand is modeled by some dynamics generating a Gaussian distribution as it simplifies the calculations considerably. We can exploit convenient properties of the Gaussian distribution in continuative calculations. As in the previous chapter, we consider a customer who observes the forward price \( F(t, t_i) := F(t, t_i, t_{i+1}) \) at time \( t \) with delivery in \([t_i, t_{i+1}]\). We introduced the expected accumulated demand in (6.3). For notational convenience we introduced

\[
D_{t_i} := \int_{t_i}^{t_{i+1}} \dot{D}_\tau d\tau.
\]

And therefore

\[
E \left[ \int_{t_i}^{t_{i+1}} \dot{D}_\tau d\tau \left| \dot{D}_{t_i} \right. \right] = E \left[ D_{t_i} \left| \dot{D}_{t_i} \right. \right]. \tag{6.27}
\]

The existence of the integral is guaranteed if

\[
\int_{t_i}^{t_{i+1}} |\dot{D}_\tau| d\tau < \infty, \quad a.s.
\]

and the expectation is well-defined if

\[
\int_{t_i}^{t_{i+1}} E \left[ \left| \dot{D}_\tau \right| \right] d\tau < \infty, \quad a.s.
\]

Modeling the instantaneous rate \( \dot{D}_t \) directly by a log-normal process (e.g. geometric Brownian motion) causes problems with the expected accumulated demand (6.27), since the resulting distribution of the inner integral is unknown. If this integral can properly be approximated by another log-normally distributed random variable, the entire expectation can be calculated explicitly. This is not desirable. An alternative approach is to model \( \dot{D}_t \) by some dynamics that generate a Gaussian distribution. Then, the integral in (6.27) is normally distributed, and hence the expectation can be calculated explicitly. Even though this latter model approach generates negative demand rates with positive probability, we employ this idea as it is e.g. common practice in interest rate markets. Moreover, under reasonable parameter settings this probability is very small. Consequently, we impose the following dynamics for the instantaneous demand:

\[
d\dot{D}_t = \gamma \left( \bar{D} - \dot{D}_t \right) dt + \nu dW_t, \tag{6.28}
\]

where \( \nu, \gamma > 0, \bar{D} \in \mathbb{R} \) and \( \{W_t\}_{t \geq 0} \) denotes a standard Brownian motion, which is independent of the Brownian motions in the forward dynamics (6.7). The following lemma summarizes relevant properties of this model approach.
Lemma 6.5.1

1. For $0 \leq t_j < t_i$, the solution of (6.28) is given by

$$
\hat{D}_{t_i} = \bar{D} + (\hat{D}_{t_j} - \bar{D})e^{-\gamma(t_i-t_j)} + \nu \int_{t_j}^{t_i} e^{-\gamma(t_i-u)} dW_u. \tag{6.29}
$$

2. Given the knowledge of $\hat{D}_{t_j}$ for $t_j \leq t_i$, we obtain

$$
D_{t_i} = \bar{D}(t_i+1 - t_i) + (\hat{D}_{t_j} - \bar{D}) \frac{e^{-\gamma(t_i-t_j)} - e^{-\gamma(t_{i+1}-t_j)}}{\gamma} + \frac{\nu}{\gamma} \left( \int_{t_i}^{t_{i+1}} 1 - e^{-\gamma(t_{i+1}-u)} dW_u + \int_{t_j}^{t_i} e^{-\gamma(t_i-u)} - e^{-\gamma(t_{i+1}-u)} dW_u \right). \tag{6.30}
$$

$D_{t_i}$ is normally distributed with mean

$$
E[D_{t_i}|\hat{D}_{t_j}] = \bar{D}(t_i+1 - t_i) + (\hat{D}_{t_j} - \bar{D}) \frac{e^{-\gamma(t_i-t_j)} - e^{-\gamma(t_{i+1}-t_i)}}{\gamma}.
$$

and variance

$$
Var[D_{t_i}|\hat{D}_{t_j}] = \left( \frac{\nu}{\gamma} \right)^2 \left[ (t_{i+1} - t_i) - \gamma + \frac{1}{2\gamma} \left( 2e^{-\gamma(t_{i+1}-t_i)} - e^{-2\gamma(t_{i+1}-t_j)} + 2e^{-\gamma(t_{i+1}+t_i-2t_j)} - e^{-2\gamma(t_i-t_j)} \right) \right].
$$

3. Define $\bar{\gamma} := \frac{e^{-\gamma(t_i-t_j)} - e^{-\gamma(t_{i+1}-t_j)}}{\gamma}$. Then $\bar{\gamma} \in (0, t_{i+1} - t_i)$.

Proof:

1. The solution can be obtained as in the proof of Lemma A.1.1.

2. Using the solution of (6.28) yields
\[ D_{t_i} = \int_{t_i}^{t_{i+1}} \left( \dot{D} + (\dot{D}_{t_i} - \bar{D}) e^{-\gamma (\tau - t_i)} + \nu \int_{t_i}^{\tau} e^{-\gamma (\tau - u)} dW_u \right) d\tau \]
\[ = \bar{D}(t_{i+1} - t_i) + (\dot{D}_{t_i} - \bar{D}) e^{-\gamma (t_{i+1} - t_i)} - e^{-\gamma (t_{i+1} - t_i)} \int_{t_i}^{t_{i+1}} e^{-\gamma (\tau - u)} dW_u d\tau \]
Fubini
\[ = \bar{D}(t_{i+1} - t_i) + (\dot{D}_{t_i} - \bar{D}) e^{-\gamma (t_{i+1} - t_i)} - e^{-\gamma (t_{i+1} - t_i)} \int_{t_i}^{t_{i+1}} e^{-\gamma (\tau - u)} dW_u d\tau \]
\[ + \nu \int_{t_i}^{t_{i+1}} \int_{t_j}^{\tau} e^{-\gamma (\tau - u)} d\tau dW_u \]
\[ = \bar{D}(t_{i+1} - t_i) + (\dot{D}_{t_i} - \bar{D}) e^{-\gamma (t_{i+1} - t_i)} - e^{-\gamma (t_{i+1} - t_i)} \int_{t_i}^{t_{i+1}} e^{-\gamma (\tau - u)} dW_u d\tau \]
\[ + \nu \int_{t_i}^{t_{i+1}} \int_{t_j}^{\tau} e^{-\gamma (\tau - u)} d\tau dW_u + \nu \int_{t_j}^{t_{i+1}} e^{-\gamma (\tau - u)} dW_u \]

The stochastic Fubini theorem (see Björk [18] and Protter [107]) can be applied, since we integrate on compact intervals over the nonnegative function \( e^{-\gamma (\tau - u)} \). The expectation is straightforward and the variance follows by Itô-Isometry.

3. Follows directly with l'Hospital.

**Remark 6.5.2** For \( t_j = t_i \) the formulae in Lemma 6.5.1 simplify considerably.

The explicit form of the first and second moment of the accumulated demand implies that (6.27) is well-defined. Let us discuss the choice of reasonable model parameters.

**Choice of Model Parameters**

It is difficult to infer parameters for the model of the customer demand (6.28), since there is no market for that. What are reasonable model parameters? The conditional expectation of \( D_{t_i} \) is measured in kWh (e.g. on a monthly basis), \( F(t_j, t_i) - k \) is measured in cents per kWh and \( K \) is given in cents. For the remainder of this chapter we consider \( t_{i+1} - t_i \equiv \text{const} \) to be a one-month period. For a reasonable monthly consumption of 125 kWh, we obtain an instantaneous mean-reversion level \( \bar{D} = \frac{125 \text{ kWh}}{730 \text{ h}} \approx 0.17 \text{ kW} \). Consider for \( t_j < t_i \)

\[ \text{Var} \left[ \hat{D}_{t_i} | \hat{D}_{t_j} \right] = \frac{\nu^2}{2\gamma} \left( 1 - e^{-2\gamma (t_i - t_j)} \right) \]
\[ \to \nu^2 (t_i - t_j), \text{ for } \gamma \to 0. \]
Choosing a variance of $\nu^2 = 0.05$ yields a fluctuation of $\pm \sqrt{0.05 \cdot 730} \approx \pm 6.04$ kWh. We assumed a mean-reversion level of $\bar{D} \cdot 730 = 125$ kWh per month. Thus, a variance of $\nu^2 = 0.05$ (in case of $\gamma = 0$) seems adequate. Note that $\nu$ is a volatility on an hourly granularity. We suggest to choose the parameters $\nu$ and $\gamma$ such that the stationary variance $\nu^2 / 2\gamma$ of the demand process is 25 kWh, which yields the relation $50\gamma = \nu^2$. The next step is to examine the influence of the demand model on the swap structure without options to cancel the contract.

If we do not know the future demand of the customer, the fixed leg of the swap structure becomes floating in some sense. There are two floating legs (i.e. it resembles a floating-for-floating swap), the future accruing costs that are uncertain and the forward price for electricity delivery in future months. However, the cost structure of the retail contracts provides a loophole such that we can still infer a fair swap rate as follows. The expected accruing costs are calculated based on the information available at time $t_0$. When we set $K > 0$ as a fixed quantity (which is reasonable from a utility’s standpoint, since it should cover fixed costs) and write for the present value

$$PV_{t_0} = \sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \left( (F(t_0, t_i) - k) \cdot \mathbb{E}\left[D_{t_i} | \hat{D}_{t_0}\right] - K \right),$$

we can recover a fair swap rate that corresponds to the consumption-dependent rate $k$ by setting $PV_{t_0} = 0$ and solving for $k$:

$$k = \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \left( \mathbb{E}\left[D_{t_i} | \hat{D}_{t_0}\right] \cdot F(t_0, t_i) - K \right)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \mathbb{E}\left[D_{t_i} | \hat{D}_{t_0}\right]},$$

given that $\mathbb{E}\left[D_{t_i} | \hat{D}_{t_0}\right] \neq 0$ for at least one $t_i$. Looking at the payment structure in this way gives us again a fixed-for-floating swap. Let us study the influence of the mean-reversion parameter $\gamma$. For $\gamma \to \infty$ we obtain

$$\hat{D}_t = \bar{D},$$

by (6.29) and

$$D_{t_i} = \int_{t_i}^{t_{i+1}} \hat{D}_t d\tau = \bar{D} \cdot (t_{i+1} - t_i),$$

by (6.30). A large $\gamma$ corresponds to the deterministic demand case, because it compensates the fluctuation generated by the Brownian shock. Furthermore,
\[ k \rightarrow \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \left( \tilde{D} \cdot (t_{i+1} - t_i) F(t_0, t_i) - K \right)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \tilde{D}(t_{i+1} - t_i)} , \quad \text{as } \gamma \rightarrow \infty \]

\[ = \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} F(t_0, t_i)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)}} \cdot \frac{K}{\tilde{D} \cdot (t_2 - t_1)} . \]

In other words, \( \tilde{D}_{t_0} \) becomes negligible. The first term is the swap rate at time \( t_0 \) reduced by the fixed cost per month. This is exactly the formula that we employed in the case study in Section 6.3 and equation (6.26) to make the example contracts comparable. On the other hand, l'Hospital yields

\[ k \rightarrow \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \left( \hat{D}_{t_0} \cdot (t_{i+1} - t_i) F(t_0, t_i) - K \right)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} \hat{D}_{t_0}(t_{i+1} - t_i)} , \quad \text{as } \gamma \rightarrow 0 \]

\[ = \frac{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)} F(t_0, t_i)}{\sum_{i=1}^{n-1} e^{-r(t_i-t_0)}} \cdot \frac{K}{\hat{D}_{t_0} \cdot (t_2 - t_1)} . \]

Thus, \( \gamma \) close to zero outweighs the mean-reverting property of the demand process. In this extreme case, the swap rate is simply a monotonically increasing function in the current instantaneous demand rate \( \hat{D}_{t_0} \) converging to \(-\infty\) for \( \hat{D}_{t_0} \rightarrow 0 \) and converging to \( \frac{\sum_{i=1}^{12} e^{-r(t_i-t_0)} F(t_0, t_i)}{\sum_{i=1}^{12} e^{-r(t_i-t_0)}} \) for \( \hat{D}_{t_0} \rightarrow \infty \). This indicates that a high current demand rate \( \hat{D}_{t_0} \) results in a higher swap rate since the customer consumes more than expected and the lack of a strong mean-reversion rate reflects this fact. We obtain a significant term structure w.r.t. \( \hat{D}_{t_0} \). More specifically, for \( \hat{D}_{t_0} < \tilde{D} \) the swap rate can become negative. This depends naturally on \( K \) but still shows that parsimonious customers are rewarded with a lower contract price.

Let us conduct further numerical experiments. Assume we have a one-year contract with monthly payment and the following parameter set:

\[ \bar{D} = 0.17 \text{ kW}, \ r = 3\%, \ K = 5 \text{ EUR (for the year)}. \]
The initial forward curve is taken from Table A.7, 4th column. This curve is observed on December 1st and we consider the base prices converted to ct/kWh. We divide these prices furthermore by 0.37 to obtain the retail prices, see Section 6.1. The first delivery is in January, i.e. in \([t_1, t_2]\), and the last delivery is the following year in December, i.e. in \([t_{12}, t_{13}]\). Consequently, \(n = 13\) in (6.32). We observe an initial instantaneous demand \(\hat{D}_{t_0} \in (0 \text{ kW}, 0.5 \text{ kW}]\) on December 1st.

Figure 6.8: Price \(k\) of the delivery contract in dependence of the initial demand \(\hat{D}_{t_0}\). Varying \(\gamma\) yields quite different values for \(k\). The mean-reversion level \(\bar{D} = 0.17 \text{ kW}\) is indicated as the black stem.

Figure 6.8 plots the function \(\hat{D}_{t_0} \mapsto k_\gamma (\hat{D}_{t_0})\) for different values of \(\gamma\), which is given in (6.32). The reasonable range of values for \(\gamma\) indicate that a stochastic demand has a significant impact in the value of the delivery contract. This value is merely calculated based upon the information available at time \(t_0\). A \(\gamma\) of 0.0005 induces a half-life of approximately 2 months and \(\gamma = 0.0015\) induces a half-life of a little less than a day. The half-life of an Ornstein-Uhlenbeck process is derived in the Appendix. As \(\gamma\) becomes large, the swap rate \(k\) converges to 11.49 ct/kWh regardless of the initial demand \(\hat{D}_{t_0}\). This number can also be recovered from the deterministic case and formula (6.5).

**Subsequent Challenges**

The current model setup exhibits two sources of randomness: The forward curve dynamics introduced in Section 6.2.2 and the demand dynamics (6.28). We assume both dynamics to be stochastically independent. To price the structured retail product we have to follow the same lines as in Chapter 6. However, there are some central questions:
1. What is an equivalent martingale measure?
2. Is condition $(\hat{\ell})$ fulfilled?
3. Is the LSMC algorithm extendable?

Question 1 is crucial. We need to know the equivalent martingale measure as it guarantees the option prices to be arbitrage-free. This requires a proper introduction of the market dynamics for the forward curves as well as for the demand rate. The above discussion is a heuristic approach and it is not clear how to set up a continuous-time model. Question 2 is the common requirement to guarantee the existence of the Bellman recursion (or Snell Envelope in this context), making the setup amendable for dynamic programming. Assuming the conditional expectation to be an element of the $L^2$ space is not sufficient for the employment of the LSMC algorithm. To our knowledge, there is no publication on the performance of this algorithm under our setting. Thus, it is important to justify the approximation scheme properly and to examine the errors incurred by the simple basis functions. We leave these issues for future research.
Appendix A

Technical Details & Supplementary Material

A.1 The Arithmetic Ornstein-Uhlenbeck Process

In this section we introduce the arithmetic Ornstein-Uhlenbeck process and derive several properties that are used in Chapter 4, 5, and 6. References are Ornstein & Uhlenbeck [103], Karatzas & Shreve [76], and Dixit & Pindyck [45]. A stochastic process \( \{S_t\}_{t \geq 0} \) is an Ornstein-Uhlenbeck process if it is stationary, Gaussian, Markovian, and continuous in probability. Sometimes \( S_t \) is also referred to as a Gauss-Markov process.

A fundamental theorem, due to Doob [46], ensures that \( S_t \) has to satisfy the equation

\[
s_t = \eta(\bar{S} - S_t)dt + \sigma dW_t, \tag{A.1}
\]

where \( \eta, \sigma > 0, \bar{S} \in \mathbb{R}, \) and \( \{W_t\}_{t \geq 0} \) denotes a standard Brownian motion. \( \bar{S} \) is the mean-reversion level and \( \eta \) is the mean-reversion rate.

Lemma A.1.1 (Basic Properties)

1. For \( 0 \leq s < t \), the solution of (A.1) is given by

\[
S_t = \bar{S} + (S_s - \bar{S})e^{-\eta(t-s)} + \sigma \int_s^t e^{-\eta(t-u)} \, dW_u.
\]

2. The conditional distribution of \( S_t \) is given by

\[
S_t | S_s \sim \mathcal{N} \left( \bar{S} + (S_s - \bar{S}) e^{-\eta(t-s)} + \frac{\sigma^2}{2\eta} \left( 1 - e^{-2\eta(t-s)} \right), \frac{\sigma^2}{2\eta} \right),
\]

with covariance

\[
\text{Cov}[S_s, S_t] = \frac{\sigma^2}{2\eta} \left( e^{-\eta(t-s)} - e^{-\eta(s+t)} \right).
\]
**Proof:** (A.1) is an Itô process, and therefore Itô’s Lemma can be applied to the function \( g(t, x) \in C^2([0, \infty) \times \mathbb{R}) \) with \( g(t, S_t) = S_t e^\eta t \) (see Øksendal [102]):

\[
d(S_t e^\eta t) = \eta S_t e^\eta t dt + e^\eta t dS_t
\]

Integrating both sides from \( s \) to \( t \) yields

\[
S_t e^\eta t - S_s e^\eta s = \bar{S} \left( e^{\eta t} - e^{\eta s} \right) + \sigma \int_s^t e^{\eta u} dW_u,
\]

which is equivalent to

\[
S_t = \bar{S} + \left( S_s - \bar{S} \right) e^{-\eta(t-s)} + \sigma \int_s^t e^{-\eta(t-u)} dW_u.
\]

The formula for the mean is straightforward and the variance, resp. covariance follows by Itô Isometry.

For Chapter 4 and 5 it is sufficient to consider the time-discrete version of (A.1). More specifically, (A.1) is the limiting case for \( \Delta t \to 0 \) of the following first-order autoregressive (AR(1)) process:

\[
S_t = (1 - e^{-\eta}) \bar{S} + e^{-\eta} S_{t-1} + \varepsilon_t
\]

\[
= \left( 1 - \kappa^t \right) \bar{S} + \kappa^t S_0 + \sum_{n=0}^{t-1} \kappa^n \varepsilon_{t-n}, \quad (A.2)
\]

where \( \varepsilon_t \) are independent Gaussian random variables with zero mean and variance \( \sigma^2_\varepsilon = \frac{\sigma^2}{2\eta} \left( 1 - e^{-2\eta} \right) \), \( S_0 \) is the initial price and for notational convenience, we set \( \kappa := e^{-\eta} \) and \( \sigma := \sigma_\varepsilon \). Consequently, we have to require that \( \kappa \in (0, 1) \) to ensure stationarity. A sound treatment of the continuous and discrete time relation can be found in Cumberland & Sykes [41].

**Parameter Inference**

In order to have realistic case studies throughout this thesis, we need to infer market parameters. The parameter estimation is naturally performed on discrete data, and henceforth (A.2) is being used as the initial equation. The Yule-Walker equations are a popular candidate to infer the unknown parameters, see Box & Jenkins [22]. Any standard statistic package (e.g. Matlab) provides convenient routines to estimate the unknown parameters. For a representative inference we used the data for the 2011 forward contracts on power base, ARA coal, and CO₂ emissions and calculated the spread for an efficient coal plant, compare Table 1.5. The results are given in Table 3.2. These quantities should merely give an indication for realistic market parameters. The focus
A.1. THE ARITHMETIC ORNSTEIN-UHLENBECK PROCESS

is more on the liquidation problem itself. Statistical studies in energy markets can be found in Kat & Oomen [78], [79] and Knittel & Roberts [85].

A subject that is essential in financial modeling is the interpretation of the parameters. \( S_t \) might denote the EUR-price of an asset and the granularity might be daily. To scale the volatility \( \sigma \) to a weekly, monthly or yearly basis, one has to multiply the quantity by \( \sqrt{5} \), \( \sqrt{20} \) or \( \sqrt{250} \), respectively. To understand the meaning of the mean-reversion parameter \( \kappa \) (or \( \eta \) in the continuous-time case), we determine the half-life of an Ornstein-Uhlenbeck process as follows. Consider (A.1) and for \( s < t \) we have

\[
E[dS_t] = \eta(\bar{S} - S_t)dt.
\]

Consequently, when no more shocks occur

\[
\frac{dS_t}{S - S_t} = \eta dt.
\]

Integrating both sides yields

\[
\int_{S_s}^{S_t} \frac{dS_u}{S - S_u} = \int_s^t \eta dt,
\]

which is equivalent to

\[
\log\left(\frac{S_t - \bar{S}}{S_s - \bar{S}}\right) = -\eta(t - s).
\]

Let \( t - s = H = \) half-life, then \( S_t - \bar{S} = 0.5(S_s - \bar{S}) \). Consequently,

\[
H = \frac{\log(2)}{\eta}.
\]

(A.3)

For instance, estimating an OU-process on a daily basis, which yields \( \eta = 0.14 \) has a half-life of \( H \approx 5 \) days. In other words, the process requires 5 days to revert half-way back from the current shock. To scale the mean-reversion parameter to a weekly, monthly, or yearly granularity one has to multiply this quantity by 5, 20 or 250. Furthermore, in the discrete time case in Chapter 5 we use \( \kappa = e^{-\eta} \), consequently \( H = \frac{\log 2}{\log \kappa} \). These considerations yield the values in Table 3.2.
A.2 Supplementary Material for Chapter 2

Figure A.1: Log-returns of forward contracts with delivery in 2009 and the resulting P&L’s $P_{\text{eff}}$ and $P_{\text{ineff}}$, 625 trading days: 04/02/2007 - 09/30/2009

Figure A.2: Log-returns of forward contracts with delivery in 2010 and the resulting P&L’s $P_{\text{eff}}$ and $P_{\text{ineff}}$, 625 trading days: 04/15/2008 - 09/30/2010
### A.2. SUPPLEMENTARY MATERIAL FOR CHAPTER 2

<table>
<thead>
<tr>
<th>Volatility Model</th>
<th>Parameters</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Volatility (p.a.)</td>
<td>( \hat{\sigma} )</td>
<td>15.65%</td>
<td>22.36%</td>
<td>43.72%</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>( \lambda )</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>( \lambda^* )</td>
<td>0.85</td>
<td>0.91</td>
<td>0.84</td>
</tr>
<tr>
<td><strong>ARCH(1)</strong></td>
<td>( \hat{\alpha}_0 )</td>
<td>4.21e-005</td>
<td>1.25e-004</td>
<td>4.29e-004</td>
</tr>
<tr>
<td></td>
<td>(2.72e-006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_1 )</td>
<td>0.7730</td>
<td>0.4276</td>
<td>0.3829</td>
</tr>
<tr>
<td></td>
<td>(0.0819)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma} )</td>
<td>1.36%</td>
<td>1.48%</td>
<td>2.64%</td>
</tr>
<tr>
<td><strong>GARCH(1,1)</strong></td>
<td>( \hat{\alpha}_0 )</td>
<td>2.22e-006</td>
<td>1.25e-006</td>
<td>4.23e-005</td>
</tr>
<tr>
<td></td>
<td>(5.87e-007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_1 )</td>
<td>0.2067</td>
<td>0.1042</td>
<td>0.1890</td>
</tr>
<tr>
<td></td>
<td>(0.0340)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>0.7823</td>
<td>0.8958</td>
<td>0.7557</td>
</tr>
<tr>
<td></td>
<td>(0.0327)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_1 + \hat{\beta} )</td>
<td>0.9890</td>
<td>1.0000</td>
<td>0.9447</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma} )</td>
<td>1.42%</td>
<td>250.38%</td>
<td>2.76%</td>
</tr>
<tr>
<td><strong>GJR-GARCH(1,1)</strong></td>
<td>( \hat{\alpha}_0 )</td>
<td>2.29e-005</td>
<td>7.66e-007</td>
<td>5.75e-005</td>
</tr>
<tr>
<td></td>
<td>(6.57e-006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_1 )</td>
<td>0.1936</td>
<td>0.1051</td>
<td>0.0864</td>
</tr>
<tr>
<td></td>
<td>(0.0329)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>0.7796</td>
<td>0.9150</td>
<td>0.7146</td>
</tr>
<tr>
<td></td>
<td>(0.0439)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\kappa} )</td>
<td>0.0329</td>
<td>-0.0403</td>
<td>0.2360</td>
</tr>
<tr>
<td></td>
<td>(0.0439)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_1 + \hat{\beta} + \frac{1}{2} \hat{\kappa} )</td>
<td>0.9897</td>
<td>1.0000</td>
<td>0.9190</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma} )</td>
<td>1.49%</td>
<td>195.74%</td>
<td>2.67%</td>
</tr>
</tbody>
</table>

Table A.1: Annualized volatility estimates for model 1 and values for \( \lambda \) for models 2-3. Furthermore, parameter estimates are provided for the volatility models 4-6. applied to the 2009 contracts. Standard deviations are given in parentheses.
Table A.2: Annualized volatility estimates for model 1 and values for λ for models 2–3. Furthermore, parameter estimates are provided for the volatility models 4–6. applied to the 2010 contracts. Standard deviations are given in parentheses.

<table>
<thead>
<tr>
<th>Volatility Model</th>
<th>Parameters</th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO₂ Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Volatility (p.a.)</td>
<td>̂σ</td>
<td>20.31%</td>
<td>31.59%</td>
<td>39.59%</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>λ</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>̂λ*</td>
<td>0.88</td>
<td>0.91</td>
<td>0.88</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>̂α₀</td>
<td>9.72e-005</td>
<td>2.87e-004</td>
<td>4.58e-004</td>
</tr>
<tr>
<td></td>
<td>(3.17e-006)</td>
<td></td>
<td>(1.49e-005)</td>
<td>(2.61e-005)</td>
</tr>
<tr>
<td></td>
<td>̂α₁</td>
<td>0.6628</td>
<td>0.3393</td>
<td>0.2781</td>
</tr>
<tr>
<td></td>
<td>(0.0750)</td>
<td>(0.0656)</td>
<td>(0.0584)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>̂σ</td>
<td>1.70%</td>
<td>2.08%</td>
<td>2.52%</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>̂α₀</td>
<td>5.49e-007</td>
<td>2.94e-006</td>
<td>1.30e-005</td>
</tr>
<tr>
<td></td>
<td>(1.72e-007)</td>
<td></td>
<td>(1.33e-006)</td>
<td>(4.82e-006)</td>
</tr>
<tr>
<td></td>
<td>̂α₁</td>
<td>0.1582</td>
<td>0.1246</td>
<td>0.0815</td>
</tr>
<tr>
<td></td>
<td>(0.0215)</td>
<td>(0.0241)</td>
<td>(0.0181)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>̂β</td>
<td>0.8418</td>
<td>0.8754</td>
<td>0.8969</td>
</tr>
<tr>
<td></td>
<td>(0.0179)</td>
<td>(0.0219)</td>
<td>(0.0212)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>̂α₁ + ̂β</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9784</td>
</tr>
<tr>
<td></td>
<td>̂σ</td>
<td>165.65%</td>
<td>383.14%</td>
<td>2.45%</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>̂α₀</td>
<td>5.27e-007</td>
<td>2.75e-006</td>
<td>1.53e-005</td>
</tr>
<tr>
<td></td>
<td>(1.95e-007)</td>
<td></td>
<td>(1.32e-006)</td>
<td>(5.21e-006)</td>
</tr>
<tr>
<td></td>
<td>̂α₁</td>
<td>0.1600</td>
<td>0.1300</td>
<td>0.0422</td>
</tr>
<tr>
<td></td>
<td>(0.0224)</td>
<td>(0.0259)</td>
<td>(0.0171)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>̂β</td>
<td>0.8439</td>
<td>0.8783</td>
<td>0.8947</td>
</tr>
<tr>
<td></td>
<td>(0.0189)</td>
<td>(0.0230)</td>
<td>(0.0220)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>̂κ</td>
<td>-0.0077</td>
<td>-0.0166</td>
<td>0.0693</td>
</tr>
<tr>
<td></td>
<td>(0.0227)</td>
<td>(0.0253)</td>
<td>(0.0264)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>̂α₁ + ̂β + 1/2 ̂κ</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9716</td>
</tr>
<tr>
<td></td>
<td>̂σ</td>
<td>162.38%</td>
<td>370.53%</td>
<td>2.32%</td>
</tr>
</tbody>
</table>
### Volatility Model

<table>
<thead>
<tr>
<th></th>
<th>Power Base</th>
<th>ARA Coal</th>
<th>CO\textsubscript{2} Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>— Contracts 2009 —</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(255.93)</td>
<td>(379.00)</td>
<td>(302.17)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>0.44%</td>
<td>28.38%</td>
<td>1.26%</td>
</tr>
<tr>
<td></td>
<td>(40.42)</td>
<td>(23.10)</td>
<td>(36.72)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>69.44%</td>
<td>33.65%</td>
<td>63.64%</td>
</tr>
<tr>
<td></td>
<td>(16.35)</td>
<td>(22.07)</td>
<td>(17.26)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(186.63)</td>
<td>(306.79)</td>
<td>(96.85)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>96.97%</td>
<td>46.16%</td>
<td>7.07%</td>
</tr>
<tr>
<td></td>
<td>(9.91)</td>
<td>(19.94)</td>
<td>(29.95)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>96.25%</td>
<td>38.40%</td>
<td>19.61%</td>
</tr>
<tr>
<td></td>
<td>(10.30)</td>
<td>(21.23)</td>
<td>(25.14)</td>
</tr>
<tr>
<td><strong>— Contracts 2010 —</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(333.28)</td>
<td>(369.63)</td>
<td>(423.05)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>1.38%</td>
<td>95.00%</td>
<td>5.75%</td>
</tr>
<tr>
<td></td>
<td>(36.39)</td>
<td>(10.85)</td>
<td>(30.83)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>11.50%</td>
<td>98.13%</td>
<td>23.70%</td>
</tr>
<tr>
<td></td>
<td>(27.77)</td>
<td>(9.13)</td>
<td>(24.12)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(150.72)</td>
<td>(183.15)</td>
<td>(135.58)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>10.62%</td>
<td>99.01%</td>
<td>3.53%</td>
</tr>
<tr>
<td></td>
<td>(28.14)</td>
<td>(8.25)</td>
<td>(32.82)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>11.69%</td>
<td>99.12%</td>
<td>9.25%</td>
</tr>
<tr>
<td></td>
<td>(27.69)</td>
<td>(8.09)</td>
<td>(28.76)</td>
</tr>
</tbody>
</table>

Table A.3: Ljung-Box test of the squared devolatilized log-returns $\hat{r}_1^2$, $\hat{r}_2^2$, and $\hat{r}_3^2$ for the delivery years 2009 and 2010, respectively.
### Volatility Model | Power Base | ARA Coal | CO₂ Emissions
---|---|---|---
#### Contracts 2009
Constant Volatility | 0.00% (11.81) | 0.00% (10.14) | 0.00% (5.56)
EWMA Risk Metrics | 0.00% (4.33) | 2.41% (2.26) | 2.14% (2.30)
EWMA Optimized | 53.33% (0.62) | 12.29% (1.54) | 41.38% (-0.82)
ARCH(1) | 9.78% (1.66) | 0.92% (2.60) | 77.35% (0.29)
GARCH(1,1) | 70.13% (-0.38) | 58.89% (0.54) | 11.05% (-1.60)
GJR-GARCH(1,1) | 71.75% (-0.36) | 29.24% (1.05) | 9.89% (-1.65)
#### Contracts 2010
Constant Volatility | 0.00% (10.64) | 0.00% (9.76) | 0.00% (7.95)
EWMA Risk Metrics | 51.26% (0.65) | 77.91% (0.28) | 24.99% (1.15)
EWMA Optimized | 23.71% (-1.18) | 58.42% (-0.55) | 89.65% (-0.13)
ARCH(1) | 0.13% (3.22) | 0.00% (4.98) | 3.27% (2.14)
GARCH(1,1) | 10.26% (-1.63) | 36.78% (-0.90) | 76.36% (0.30)
GJR-GARCH(1,1) | 12.23% (-1.55) | 41.95% (-0.81) | 87.52% (0.16)

Table A.4: BDS test of the devolatilized log-returns \( \hat{r}_1 \), \( \hat{r}_2 \), and \( \hat{r}_3 \) for the delivery years 2009 and 2010, respectively.
### Table A.5:  Backtesting results and p-values of several tests for VaR adequacy. The test statistics are given in parentheses.

<table>
<thead>
<tr>
<th>Model</th>
<th># Violations</th>
<th>Kupiec LR</th>
<th>Markov</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contracts 2009</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>15</td>
<td>35.80%</td>
<td>12.98%</td>
<td>20.80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.66)</td>
<td>(0.32)</td>
<td>(0.97)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>12</td>
<td>8.77%</td>
<td>5.01%</td>
<td>3.41%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.23)</td>
<td>(0.15)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>17</td>
<td>67.37%</td>
<td>21.40%</td>
<td>42.28%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.92)</td>
<td>(0.46)</td>
<td>(1.38)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>17</td>
<td>67.37%</td>
<td>20.31%</td>
<td>40.72%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.92)</td>
<td>(0.45)</td>
<td>(1.36)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>16</td>
<td>50.43%</td>
<td>70.78%</td>
<td>74.59%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.80)</td>
<td>(0.93)</td>
<td>(1.73)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>15</td>
<td>35.80%</td>
<td>62.26%</td>
<td>58.06%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.66)</td>
<td>(0.89)</td>
<td>(1.54)</td>
</tr>
<tr>
<td><strong>Contracts 2009</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>23</td>
<td>33.00%</td>
<td>61.87%</td>
<td>54.98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.62)</td>
<td>(0.88)</td>
<td>(1.51)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>14</td>
<td>23.96%</td>
<td>1.08%</td>
<td>1.94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.50)</td>
<td>(0.04)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>19</td>
<td>95.29%</td>
<td>32.50%</td>
<td>61.50%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.00)</td>
<td>(0.62)</td>
<td>(1.61)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>15</td>
<td>35.80%</td>
<td>26.28%</td>
<td>35.01%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.66)</td>
<td>(0.53)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>16</td>
<td>50.43%</td>
<td>23.17%</td>
<td>39.13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.80)</td>
<td>(0.49)</td>
<td>(1.29)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>14</td>
<td>23.96%</td>
<td>29.67%</td>
<td>29.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.50)</td>
<td>(0.58)</td>
<td>(1.08)</td>
</tr>
</tbody>
</table>
## APPENDIX A. TECHNICAL DETAILS & SUPPLEMENTARY MATERIAL

<table>
<thead>
<tr>
<th>Model</th>
<th># Violations</th>
<th>Kupiec LR</th>
<th>Markov</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>— Contracts 2010 P&amp;L_{eff} —</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>27</td>
<td>6.59% (0.18)</td>
<td>96.88% (1.00)</td>
<td>18.41% (1.18)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>4</td>
<td>0.00% (0.00)</td>
<td>76.87% (0.96)</td>
<td>0.00% (0.96)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>13</td>
<td>15.02% (0.36)</td>
<td>33.32% (0.63)</td>
<td>22.24% (0.98)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>22</td>
<td>45.30% (0.76)</td>
<td>77.57% (0.96)</td>
<td>72.46% (1.72)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>17</td>
<td>67.37% (0.92)</td>
<td>79.48% (0.97)</td>
<td>88.47% (1.88)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>14</td>
<td>23.96% (0.50)</td>
<td>29.67% (0.58)</td>
<td>29.05% (1.08)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th># Violations</th>
<th>Kupiec LR</th>
<th>Markov</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>— Contracts 2010 P&amp;L_{ineff} —</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant Volatility</td>
<td>27</td>
<td>6.59% (0.18)</td>
<td>15.65% (0.37)</td>
<td>6.75% (0.55)</td>
</tr>
<tr>
<td>EWMA Risk Metrics</td>
<td>9</td>
<td>1.05% (0.04)</td>
<td>50.53% (0.80)</td>
<td>3.02% (0.84)</td>
</tr>
<tr>
<td>EWMA Optimized</td>
<td>14</td>
<td>23.96% (0.50)</td>
<td>31.49% (0.60)</td>
<td>30.22% (1.10)</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>19</td>
<td>95.29% (1.00)</td>
<td>15.37% (0.36)</td>
<td>36.09% (1.36)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>16</td>
<td>50.43% (0.80)</td>
<td>23.17% (0.49)</td>
<td>39.13% (1.29)</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>13</td>
<td>15.02% (0.36)</td>
<td>33.32% (0.63)</td>
<td>22.24% (0.98)</td>
</tr>
</tbody>
</table>

Table A.6: Backtesting results and p-values of several tests for VaR adequacy. The test statistics are given in parentheses.
Figure A.3: Real losses (indicated as bars) for the 2009 contracts compared to the $\text{VaR}_{0.05}(t)$, $t = 251, ..., 625$, efficient power plant.
Figure A.4: Real losses (indicated as bars) for the 2009 contracts compared to the $\text{VaR}_{0.05}(t)$, $t = 251, \ldots, 625$, inefficient power plant.
Figure A.5: Real losses (indicated as bars) for the 2010 contracts compared to the VaR_{0.05}(t), \( t = 251, \ldots, 625 \), efficient power plant.
Figure A.6: Real losses (indicated as bars) for the 2010 contracts compared to the VaR_{0.05}(t), \( t = 251, \ldots, 625 \), inefficient power plant.
A.3 Supplementary Material for Chapter 5

A.3.1 Distribution of $W_{T+1}$ under $\pi^*$

$W_{T+1}$ is the sum of products of Gaussian random variables with means not necessarily equal to zero. The resulting distribution is unknown. However, we can calculate the first moment using (A.2), $\tilde{S}_t = S_t - \theta \bar{\pi}_t^*$, and (5.20):

$$\mathbb{E}_0 [W_{T+1}] = \bar{S} \sum_{t=0}^T \mathbb{E}_0 [\pi_t^*] + \sum_{t=0}^T \mathbb{E}_0 [\pi_t^*] \cdot \kappa^t \cdot (s_0 - \bar{S}) + \sum_{t=0}^T \mathbb{E}_0 \left[ \pi_t^* \sum_{n=0}^{t-1} \kappa^n \varepsilon_{t-n} \right].$$

(i) is of most interest. Thus,

$$\begin{align*}
(i) &= \sum_{t=0}^T \sum_{n=0}^{t-1} \kappa^n \mathbb{E}_0 [\pi_t^* \varepsilon_{t-n}] \\
&= \sum_{t=0}^T \sum_{n=0}^{t-1} \kappa^n \left\{ d_i \sum_{i=0}^{t-1} \kappa^i \mathbb{E}_0 [\varepsilon_{t-i} \varepsilon_{t-n}] - \sum_{i=0}^{t-1} \sum_{j=0}^{i-1} \kappa^j \mathbb{E}_0 [\varepsilon_{i-j} \varepsilon_{t-n}] \right\} \\
&= \sum_{t=1}^T \sum_{n=0}^{t-1} \sum_{i=0}^{t-1} d_i \kappa^{n+i} \mathbb{E}_0 [\varepsilon_{t-i} \varepsilon_{t-n}] - \sum_{t=2}^T \sum_{n=0}^{t-1} \sum_{j=0}^{t-1} \sum_{i=0}^{j-1} \frac{d_i}{T-i} \kappa^{n+j} \mathbb{E}_0 [\varepsilon_{i-j} \varepsilon_{t-n}] \\
&= \frac{\sigma^2}{\kappa^2 - 1} \sum_{t=1}^T \left( d_t (\kappa^t - 1) - \sum_{n=1}^{t-1} \frac{d_n}{T-n} \kappa^{t-n} (\kappa^{2n} - 1) \right).
\end{align*}$$

The last equation can be obtained by ruling out the summands where $\mathbb{E}_0 [\varepsilon_i \varepsilon_j] = 0$, which is the case for all combinations $i \neq j$. We therefore have (5.26).

A.3.2 Distribution of $W_{T+1}$ under $\hat{\pi}$

Using (A.2), $\tilde{S}_t = S_t - \theta \bar{\pi}_t$, and $\hat{\pi}_t = \frac{Q_{tot}}{T+1}$ yields

$$\sum_{t=0}^T \tilde{S}_t \hat{\pi}_t = \frac{Q_{tot}}{T+1} \sum_{t=0}^T \left( S_t - \theta \frac{Q_{tot}}{T+1} \right).$$

$$= \frac{Q_{tot}}{T+1} \left( \tilde{S} \sum_{t=0}^T (1 - \kappa^t) + s_0 \sum_{t=0}^T \kappa^t + \sum_{t=1}^T \sum_{n=0}^{t-1} \kappa^n \varepsilon_{t-n} - \theta Q_{tot} \right).$$

Consequently, $\sum_{t=0}^T \tilde{S}_t \hat{\pi}_t$ is normally distributed, since the innovations $\varepsilon_t$ are i.i.d. Gaussian random variables with mean zero and variance $\sigma^2$. The expectation of $\sum_{t=0}^T \tilde{S}_t \bar{\pi}_t$ is given by
\[
\mathbb{E}_0 \left[ \sum_{t=0}^{T} \tilde{S}_t \hat{\pi}_t \right] = \frac{Q_{\text{tot}}}{T+1} \left( \bar{S}(T+1) + \frac{\kappa^{T+1} - 1}{\kappa - 1} \left( s_0 - \bar{S} \right) - \theta Q_{\text{tot}} \right),
\]
and the variance is
\[
\text{Var} \left[ \sum_{t=0}^{T} \tilde{S}_t \hat{\pi}_t \right] = \left( \frac{Q_{\text{tot}}}{T+1} \right)^2 \mathbb{E}_0 \left[ \sum_{t=1}^{T} \sum_{n=0}^{t-1} \kappa^n \varepsilon_{t-n} \right] = \left( \frac{Q_{\text{tot}}}{T+1} \right)^2 \sum_{t=1}^{T} \left( \frac{\kappa^{T-t+1} - 1}{\kappa - 1} \right)^2 \mathbb{E}_0 \left[ \varepsilon_t^2 \right] = \left( \frac{Q_{\text{tot}}}{T+1} \right)^2 \left( T \left( \kappa^2 - 1 \right) + \kappa \left( \kappa^T - 1 \right) \left( \kappa^{T+1} - \kappa - 2 \right) \right) \left( \kappa - 1 \right)^2 \left( \kappa^2 - 1 \right).
\]
These considerations yield the expectation given in (5.27) and variance given in (5.28).

The variance is always positive:
\[
\text{Var} \left[ \sum_{t=0}^{T} \tilde{S}_t \hat{\pi}_t \right] \to \left( \frac{Q_{\text{tot}}}{T+1} \right)^2 T, \text{ as } \kappa \to 0,
\]
and
\[
\text{Var} \left[ \sum_{t=0}^{T} \tilde{S}_t \hat{\pi}_t \right] \to \infty, \text{ as } \kappa \to 1.
\]
A.4 Supplementary Material for Chapter 6

<table>
<thead>
<tr>
<th>Delivery</th>
<th>October (Base/Peak)</th>
<th>November (Base/Peak)</th>
<th>December (Base/Peak)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct 08</td>
<td>88.65 / 118.65</td>
<td>78.00 / 113.60</td>
<td>60.85 / 84.50</td>
</tr>
<tr>
<td>Nov 08</td>
<td>94.50 / 136.25</td>
<td>71.10 / 98.50</td>
<td>57.25 / 78.40</td>
</tr>
<tr>
<td>Dec 08</td>
<td>88.25 / 122.30</td>
<td>78.38 / 116.00</td>
<td>61.38 / 86.60</td>
</tr>
<tr>
<td>Jan 09</td>
<td>96.00 / 133.00</td>
<td>78.70 / 116.40</td>
<td>56.00 / 82.88</td>
</tr>
<tr>
<td>Feb 09</td>
<td>95.43 / 133.00</td>
<td>64.65 / 92.00</td>
<td>53.50 / 77.50</td>
</tr>
<tr>
<td>Mar 09</td>
<td>83.83 / 118.45</td>
<td>61.38 / 83.75</td>
<td>49.50 / 68.50</td>
</tr>
<tr>
<td>Apr 09</td>
<td>70.78 / 94.50</td>
<td>56.15 / 72.80</td>
<td>54.60 / 77.23</td>
</tr>
<tr>
<td>May 09</td>
<td>64.65 / 92.00</td>
<td>78.38 / 116.00</td>
<td>53.10 / 83.25</td>
</tr>
<tr>
<td>Jun 09</td>
<td>68.55 / 96.30</td>
<td>78.38 / 116.40</td>
<td>58.13 / 89.43</td>
</tr>
<tr>
<td>Jul 09</td>
<td>70.85 / 102.00</td>
<td>60.60 / 87.50</td>
<td>58.30 / 91.75</td>
</tr>
<tr>
<td>Aug 09</td>
<td>66.73 / 95.50</td>
<td>80.00 / 119.85</td>
<td>69.15 / 96.00</td>
</tr>
<tr>
<td>Sep 09</td>
<td>75.58 / 115.70</td>
<td>64.50 / 91.25</td>
<td>64.08 / 83.58</td>
</tr>
<tr>
<td>Oct 09</td>
<td>76.73 / 115.15</td>
<td>64.50 / 91.25</td>
<td>54.60 / 77.23</td>
</tr>
<tr>
<td>Nov 09</td>
<td>88.68 / 126.08</td>
<td>80.60 / 119.85</td>
<td>66.75 / 97.40</td>
</tr>
<tr>
<td>Dec 09</td>
<td>83.80 / 115.50</td>
<td>75.50 / 109.85</td>
<td>69.25 / 97.40</td>
</tr>
<tr>
<td>Jan 10</td>
<td>89.18 / 128.00</td>
<td>79.35 / 120.75</td>
<td>58.98 / 89.00</td>
</tr>
<tr>
<td>Feb 10</td>
<td>89.15 / 128.00</td>
<td>79.35 / 120.75</td>
<td>58.98 / 89.00</td>
</tr>
<tr>
<td>Mar 10</td>
<td>77.65 / 116.88</td>
<td>69.38 / 100.25</td>
<td>56.88 / 89.90</td>
</tr>
<tr>
<td>Apr 10</td>
<td>68.55 / 91.33</td>
<td>69.00 / 81.95</td>
<td>50.88 / 78.00</td>
</tr>
<tr>
<td>May 10</td>
<td>62.16 / 89.64</td>
<td>55.12 / 72.47</td>
<td>53.60 / 79.99</td>
</tr>
<tr>
<td>Jun 10</td>
<td>65.91 / 93.83</td>
<td>63.32 / 90.83</td>
<td>56.61 / 82.35</td>
</tr>
<tr>
<td>Jul 10</td>
<td>68.12 / 93.39</td>
<td>63.93 / 95.07</td>
<td>54.31 / 86.23</td>
</tr>
<tr>
<td>Aug 10</td>
<td>64.16 / 93.05</td>
<td>59.19 / 87.10</td>
<td>59.45 / 92.63</td>
</tr>
<tr>
<td>Sep 10</td>
<td>72.67 / 112.74</td>
<td>65.28 / 101.16</td>
<td>59.63 / 95.04</td>
</tr>
<tr>
<td>Oct 10</td>
<td>73.77 / 112.20</td>
<td>65.28 / 100.29</td>
<td>59.72 / 99.44</td>
</tr>
<tr>
<td>Nov 10</td>
<td>85.26 / 122.85</td>
<td>79.12 / 119.30</td>
<td>65.53 / 86.57</td>
</tr>
<tr>
<td>Dec 10</td>
<td>80.57 / 112.54</td>
<td>74.12 / 109.35</td>
<td>69.43 / 100.35</td>
</tr>
<tr>
<td>Jan 11</td>
<td>89.18 / 127.68</td>
<td>79.98 / 122.29</td>
<td>72.03 / 100.35</td>
</tr>
<tr>
<td>Feb 11</td>
<td>89.15 / 127.68</td>
<td>81.36 / 122.29</td>
<td>61.34 / 91.69</td>
</tr>
<tr>
<td>Mar 11</td>
<td>77.65 / 116.58</td>
<td>71.14 / 101.53</td>
<td>59.16 / 92.62</td>
</tr>
<tr>
<td>Apr 11</td>
<td>68.55 / 91.10</td>
<td>66.65 / 83.00</td>
<td>52.92 / 80.36</td>
</tr>
<tr>
<td>May 11</td>
<td>62.16 / 89.42</td>
<td>56.52 / 73.40</td>
<td>55.75 / 82.41</td>
</tr>
<tr>
<td>Jun 11</td>
<td>65.91 / 93.60</td>
<td>64.93 / 92.00</td>
<td>58.88 / 84.84</td>
</tr>
<tr>
<td>Jul 11</td>
<td>68.12 / 99.14</td>
<td>65.55 / 96.28</td>
<td>56.49 / 88.84</td>
</tr>
<tr>
<td>Aug 11</td>
<td>64.16 / 92.82</td>
<td>60.70 / 88.22</td>
<td>61.83 / 95.43</td>
</tr>
<tr>
<td>Sep 11</td>
<td>72.67 / 112.46</td>
<td>66.94 / 102.46</td>
<td>62.02 / 97.91</td>
</tr>
<tr>
<td>Oct 11</td>
<td>73.77 / 111.92</td>
<td>66.94 / 101.57</td>
<td>73.56 / 102.45</td>
</tr>
<tr>
<td>Nov 11</td>
<td>85.26 / 122.54</td>
<td>81.13 / 120.83</td>
<td>73.56 / 102.45</td>
</tr>
</tbody>
</table>

Table A.7: Three forward curves (base/peak) noted in EUR/MWh.
List of Tables

1.1 Typical multicommodity portfolio of a utility company, used for illustration. Baseload and peakload positions, coal, and CO\textsubscript{2} emissions are 2011 forward contracts. The prices are taken from 09/30/2010 and the volatility is computed historically from the corresponding log-returns data. ........................................... 5


1.3 ARA coal API #2 year futures 2009, 2010, and 2011 contract information. Source: EEX. The prices for the 2009 contract are taken from the trader Evolution, since the EEX traded futures price series are not sufficiently long. ........ 7


1.5 Comparison of an efficient and inefficient coal-fired power plant. Source: EnBW. The total capacity $Q_\text{tot}$ is not necessarily the total capacity of the power plant but merely the available capacity for the utility as plant sharing is common in Germany. ........................................... 9

2.1 Ljung-Box test of the log-returns to show that the ACF of the log-returns are not equal to zero. ................................................... 20

2.2 The volatility across the years and contracts shows very different patterns. This is mainly due to the speculation bubble in 2008. .................. 21

2.3 Ljung-Box test of the squared log-returns for all commodities and portfolios under consideration. The table lists the p-values with the corresponding test statistics in parentheses. ................................. 23

2.4 BDS test on log-returns $r_1(t), r_2(t), r_3(t)$, and the portfolios under consideration. The table lists the p-values with the corresponding test statistics in parentheses. 24

2.5 Annualized volatility estimates for model 1 and values for $\lambda$ for models 2.-3. Furthermore, parameter estimates are provided for the volatility models 4.-6. applied to the 2011 contracts. Standard deviations are given in parentheses. 30

2.6 Ljung-Box test p-values of the squared devolatilized log-returns $\hat{r}_1^2$, $\hat{r}_2^2$, and $\hat{r}_3^2$ for the delivery year 2011. The corresponding test statistics are given in parentheses. 33

2.7 BDS test p-values of the devolatilized log-returns $\hat{r}_1$, $\hat{r}_2$, and $\hat{r}_3$ for the delivery year 2011. The corresponding test statistics are given in parentheses. .................... 34

2.8 Backtesting results and p-values of several tests for VaR adequacy. The test statistics are given in parentheses. ......................... 41
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary of the scheme that we examine in the next two chapters.</td>
<td>50</td>
</tr>
<tr>
<td>3.2</td>
<td>Initial model parameters that are used for the empirical analysis in the sequel. The granularity is daily. The data set used for estimation is comprised of the price series of the 2011 contracts of power base, ARA coal, and CO₂ emissions.</td>
<td>54</td>
</tr>
<tr>
<td>5.1</td>
<td>Model and input parameters for the numerical example, granularity is weekly.</td>
<td>106</td>
</tr>
<tr>
<td>5.2</td>
<td>Summary of the results of the numerical example. ( C_{T+1} = \theta \sum_{t=0}^{T} (\pi_t)^2 ) denotes the trading costs incurred by the impact parameter ( \theta ). The volatility is calculated from (5.28). All numbers are in EUR, except ( \kappa ).</td>
<td>108</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary of contracts offered by several utility companies in Germany assuming a constant consumption of 1500 kWh/year; the contracts were placed in the market by the end of 2008.</td>
<td>152</td>
</tr>
<tr>
<td>6.2</td>
<td>Market observed at-the-money call options and implied volatilities. All prices are in EUR. Date of calibration is 09/30/2008.</td>
<td>153</td>
</tr>
<tr>
<td>6.3</td>
<td>Calibration results from the forward model introduced in Section 6.2.2.</td>
<td>154</td>
</tr>
<tr>
<td>6.4</td>
<td>This table contains the offered modified kWh-rates, the fair modified kWh-rates (excl. options), the net present values and the values of the options (option premia).</td>
<td>156</td>
</tr>
<tr>
<td>6.5</td>
<td>The fair modified kilowatt-hour rate incl./excl. options would make the contract with/without call rights have zero net present value.</td>
<td>160</td>
</tr>
<tr>
<td>A.1</td>
<td>Annualized volatility estimates for model 1 and values for ( \lambda ) for models 2.-3. Furthermore, parameter estimates are provided for the volatility models 4.-6. applied to the 2009 contracts. Standard deviations are given in parentheses.</td>
<td>173</td>
</tr>
<tr>
<td>A.2</td>
<td>Annualized volatility estimates for model 1 and values for ( \lambda ) for models 2.-3. Furthermore, parameter estimates are provided for the volatility models 4.-6. applied to the 2010 contracts. Standard deviations are given in parentheses.</td>
<td>174</td>
</tr>
<tr>
<td>A.3</td>
<td>Ljung-Box test of the squared devolatilized log-returns ( \hat{r}_1^2 ), ( \hat{r}_2^2 ), and ( \hat{r}_3^2 ) for the delivery years 2009 and 2010, respectively.</td>
<td>175</td>
</tr>
<tr>
<td>A.4</td>
<td>BDS test of the devolatilized log-returns ( \hat{r}_1 ), ( \hat{r}_2 ), and ( \hat{r}_3 ) for the delivery years 2009 and 2010, respectively.</td>
<td>176</td>
</tr>
<tr>
<td>A.5</td>
<td>Backtesting results and p-values of several tests for VaR adequacy. The test statistics are given in parentheses.</td>
<td>177</td>
</tr>
<tr>
<td>A.6</td>
<td>Backtesting results and p-values of several tests for VaR adequacy. The test statistics are given in parentheses.</td>
<td>178</td>
</tr>
<tr>
<td>A.7</td>
<td>Three forward curves (base/peak) noted in EUR/MWh.</td>
<td>185</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Sources of power generation in Europe (left panel) and in Germany (right panel). Source: BDEW, IEA. ................................................................. 4

2.1 Autocorrelation function (ACF) of price series of power base (left panel) compared to their log-returns (right panel), contract with delivery in 2011. .......... 20

2.2 Log-returns of forward contracts with delivery in 2011 and the resulting P&L’s $P_{\text{L eff}}$ and $P_{\text{L ineff}}$, 625 trading days: 04/15/2008 - 09/30/2010. .......... 21

2.3 Annualized, rolling volatility with a window of 100 days. All commodities exhibit volatility clustering. ................................................................. 22

2.4 Autocorrelation function of squared log-returns of power base, ARA coal, and CO$_2$ emissions 2011 futures contracts and both portfolios, $P_{\text{L eff}}$ and $P_{\text{L ineff}}$, respectively. ................................................................. 23

2.5 The plot shows the different p-values from the Ljung-Box test for varying values of $\lambda$. ................................................................. 31

2.6 Autocorrelation function of the squared returns of power base, ARA coal, and CO$_2$ emissions for the models (1) constant volatility and (2) EWMA Risk Metrics. 32

2.7 Autocorrelation function of the squared returns of power base, ARA coal, and CO$_2$ emissions for the models (3) EWMA Optimized and (4) ARCH. .......... 32

2.8 Autocorrelation function of the squared returns of power base, ARA coal, and CO$_2$ emissions for the models (5) GARCH and (6) GJR-GARCH. .......... 33

2.9 Histograms of the devolatilized residuals $\hat{r}_i(t)$ of each commodity and fitted NIG distribution (red line) and Gaussian distribution (blue line) for comparison, 2011 contracts. ................................................................. 34

2.10 Real losses (indicated as bars) for the 2011 contracts compared to the VaR$_{0.05}(t)$, $t = 251, ..., 625$, efficient power plant. ................................................................. 42

2.11 Real losses (indicated as bars) for the 2011 contracts compared to the VaR$_{0.05}(t)$, $t = 251, ..., 625$, inefficient power plant. ................................................................. 43

4.1 This figure shows the admissible set $\Pi \cap \Gamma$ for different values of $\text{PR} \in \{500, 1000\}$ (left to right). Parameters for illustration are $T = 2$, $Q_{\text{tot}} = 70$ MW, $Q_{\text{max}} = 45$ MW, $\alpha = 0.02$, $\sigma = 2.2$, $\mu = 0.5$, $S_0 = 20$ EUR. ................................................................. 63
LIST OF FIGURES

4.2 Left panel: Model 1 and the blue area depicts the admissible pairs \((PR, \alpha)\), which yield \(\Pi \cap \Gamma \neq 0\). Right panel: Model 2 and the blue area depicts the admissible pairs \((PR, \alpha)\) which yield \(\Pi \cap \Gamma \neq 0\). .............. 71

4.3 Liquidation strategies for both models, random walk (upper panel) and mean-reverting process (lower panel). .................. 72

4.4 Liquidation strategies for both models, random walk and mean-reverting process. ............ 74

4.5 Left panel: For model 1, the varying steepness of the volatility term structure leads to almost opposite strategies. Right panel: For model 2 the varying steepness of the volatility term structure leads to a flattening of the liquidation strategy. 75

5.1 \(\pi_{\text{min}}^t\) (blue line) and \(\pi_{\text{max}}^t\) (red line) as functions of \(a_t^{t-1}\). Parameters are \(Q_{\text{tot}} = 500\), \(Q_{\text{max}} = 80\), \(T = 51\), \(t = 50\), and \(q_{t-1} = 120\) is fixed. .......................... 86

5.2 Portfolio position \(q_t\) in dependence of the drift parameter \(\mu\) for the random walk model. The other parameters are \(Q_{\text{tot}} = 500\) MW and \(\theta = 0.05\). .............. 98

5.3 Distribution of terminal wealth \(W_{T+1}\) under the optimal strategy \(\pi^*\) (blue area) and naive strategy \(\hat{\pi}\) (green area). Model parameters are \(s_0 = 20\) EUR, \(\bar{S} = 20\) EUR, \(\sigma = 1\), \(\theta = 0.005\), \(Q_{\text{tot}} = 500\) MW, \(T = 51\), Monte Carlo iterations: 50,000. Left graph: \(\kappa = 0.25\), middle graph: \(\kappa = 0.5\), right graph: \(\kappa = 0.75\). 107

5.4 Optimal liquidation under three different price realizations \(s_t\). The parameters are given in Table 5.1 and \(\kappa = 0.5\). ........................................ 109

5.5 The monotonicity of the function \(a_{T-1} \mapsto s_{T-1}a_{T-1} + \lambda(1 - \Phi(h(a_{T-1})))\) can be categorized. In two out of four cases the maximum is uniquely determined (left panel). .............. 115

5.6 Three representative spread price paths with \(s_0 = 20\) EUR. .............................. 120

5.7 Sanity check for the random walk: \(\sigma = 1\), \(\mu = 0.5\), \(s_0 = 20\) EUR, \(PR = 20000\) EUR (indicated by the shorter bar in the lower, right figure), \(\lambda = 0\) EUR. The longer bar in the lower, right figure corresponds to the mean of the terminal wealth, which is equal to 22,118 EUR (this number can also be recovered by using Corollary 5.2.8: \(E_0[W_{T+1}] = 22,090\) EUR, i.e. a Monte-Carlo error of less than 0.2%). .......................... 121

5.8 Optimal liquidation under a target wealth of 12,000 EUR and a penalty of \(\lambda = 2500\) EUR. ........................................ 122

5.9 Optimal liquidation under a target wealth of 12,000 EUR and a penalty of \(\lambda = 7500\) EUR (upper panel) and \(\lambda = 100,000\) EUR (lower panel). .................. 123

5.10 Probability of reaching the planning result \(PR\) under a random walk specification. .................. 124


6.2 Flow of electricity versus cash for the retail contracts under consideration. .......................... 134

6.3 Points in time under consideration. At \(t_0\), we assume that the contracts are priced and issued immediately. In \((t_0, t_1)\), the customer can sign the agreement. \([t_1, t_2], ..., [t_{m-1}, t_m]\) are periods of certain delivery and \(\{t_m, ..., t_n\}\) are the possible dates to cancel the agreement. .................. 141
6.4 The function \( AC \mapsto PV_{\text{t}}(AC) \) is monotonically decreasing. This property is crucial to guarantee the uniqueness of \( AC^* \). The figure shows three different net present values: (1) \( m = 2, n = 4 \), (2) \( m = 3, n = 4 \), (3) \( m = 4, n = 4 \).

6.5 Convergence of AC in a simple Black model with \( \sigma = 0.3 \), \( F(t_0, t_1) = 57.25 \text{EUR/MWh} \), \( F(t_0, t_2) = 61.38 \text{EUR/MWh} \), \( r = 0.03 \) and \( D = 125 \text{kWh} \). Initial value is \( AC_0 = 7 \text{EUR} \).

6.6 Life time of the contracts under consideration. The figure contains the monthly forward curves for the October, November, and December 2008 base (left axis) and peak (right axis) contracts. Moreover, the calling dates are indicated as white bars in the lower, gray stripes.

6.7 Cumulated cancel probabilities for the actual offered contracts provided by the LSMC algorithm.

6.8 Price \( k \) of the delivery contract in dependence of the initial demand \( \hat{D}_t \). Varying \( \gamma \) yields quite different values for \( k \). The mean-reversion level \( \bar{D} = 0.17 \text{kW} \) is indicated as the black stem.

A.1 Log-returns of forward contracts with delivery in 2009 and the resulting P&L’s \( P\&L_{\text{eff}} \) and \( P\&L_{\text{ineff}} \). 625 trading days: 04/02/2007 - 09/30/2009.

A.2 Log-returns of forward contracts with delivery in 2010 and the resulting P&L’s \( P\&L_{\text{eff}} \) and \( P\&L_{\text{ineff}} \). 625 trading days: 04/15/2008 - 09/30/2010.

A.3 Real losses (indicated as bars) for the 2009 contracts compared to the VaR_{0.05}(t), \( t = 251, ..., 625 \), efficient power plant.

A.4 Real losses (indicated as bars) for the 2009 contracts compared to the VaR_{0.05}(t), \( t = 251, ..., 625 \), inefficient power plant.

A.5 Real losses (indicated as bars) for the 2010 contracts compared to the VaR_{0.05}(t), \( t = 251, ..., 625 \), efficient power plant.

A.6 Real losses (indicated as bars) for the 2010 contracts compared to the VaR_{0.05}(t), \( t = 251, ..., 625 \), inefficient power plant.
Bibliography


BIBLIOGRAPHY


Zusammenfassung


Beitrag von Teil 1/Kapitel 2 - Eine Multivariate Rohstoffmarkt-Analyse mit Anwendungen im Risikomanagement


Viele Modelle befassen sich mit geeigneten Verteilungsfunktionen in Kombination mit einem passenden Abhängigkeitskonzept, z.B. Copulas. Sie vernachlässigen jedoch die

Beitrag von Teil 2/Kapitel 4 - Optimale Statische Liquidierungsstrategien


Eine realistische Risikobeschränkung ist ein Zielumsatz, welcher mit einer gewissen Wahrscheinlichkeit bis zum Ende der Liquidierungsperiode erreicht werden soll. Wir bezeichnen diese Nebenbedingung als Managementregel (management rule), weil sie oft aus praktischen Anwendungen stammt und mehr oder minder heuristisch bestimmt wird. Wir entwickeln ein numerisches Tool, welches der Firma Kriterien liefert, nach denen es den Zielumsatz von vorneherein fixieren kann. Wir untersuchen zwei verschiedene Modelle und die sich daraus ergebenen Liquidierungsstrategien, und wir quantifizieren die Umsätze indem wir realistische Szenarios verwenden.

Das Hauptergebnis ist, dass der Stromversorger einen optimalen Kompromiss zwischen maximalem erwarteten Umsatz und dem Risiko, den Zielumsatz nicht zu erreichen, sucht. Die Driftkomponente, die durch das jeweilige Modelle induziert wird, hat einen signifikanten Einfluss auf die Strategien in einem risikoneutralen Setting. Jedoch sorgt der Kompromiss dafür, dass die Strategien so aussehen, dass der Händler schneller verkauft.

Beitrag von Teil 2/Kapitel 5 - Optimale Dynamische Liquidierungsstrategien


In einem abschließenden Teil geben wir einen Ausblick für zukünftige Arbeit. Wir sind überzeugt, dass das Konzept der dynamischen Programmierung geeignet ist, um optimales Liquidierungsverhalten zu untersuchen. Insbesondere ist die Risikobeschränkung ein interessanter Aspekt und in Kombination mit realistischeren Modellen ein vielversprechendes Untersuchungsfeld in der Zukunft.

Beitrag von Teil 3/Kapitel 6 - Bewertung von Strukturierten Stromlieferverträgen


Für die Analyse der Verträge entwickeln wir einen risikoneutralen Bewertungsrahmen, welcher die Annahme eines geeigneten Marktmodeells erfordert. Wir kalibrieren das Modell, um weitere Marktparameter zu erhalten. Die Vertragsstruktur erfordert die Formulierung eines Rahmens mittels des Snell Envelope, welcher als Bellman Gleichung
für dieses Problem angesehen werden kann. Wir leiten Struktureigenschaften her, wo sie benötigt werden. Die Formulierung als dynamisches Programm kann nicht explizit gelöst werden und wir schlagen ein numerisches Verfahren in Form des Least-Squares Monte-Carlo Algorithmus vor. Dieser erlaubt uns, die strukturierten Kontrakte monetär zu bewerten. Wir betrachten reale Verträge und untersuchen, ob diese innerhalb unseres Modells korrekt bewertet sind. In einem zweiten Schritt zeigen wir, wie man solche Verträge initial bewerten kann, was eine nützliche Hilfe für Stromlieferanten ist.

In einem abschließenden Teil schlagen wir einen erweiterten Rahmen vor. Ein bedeutender Nachteil ist die Annahme eines deterministischen Kundenstromverbrauchs. Man weiß, dass das Volumenrisiko eine Hauptrisikoposition im Portfolio eines Energieversorgers darstellt. Wir präsentieren eine Herangehensweise mit der man den Verbrauch sinnvoll modellieren kann, sodass benötigte Kenngrößen explizit berechnet werden können. Darüber hinaus ziehen wir Parallelen zum deterministischen Rahmenwerk, was unseren Vorschlag validiert. Wir benennen die wichtigsten nächsten Schritte, welche in zukünftiger Arbeit angesprochen werden können.
Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Ulm, den 25.01.2011 

(Kevin Metka)