UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

UNIVERSITY OF SOUTHAMPTON

FACULTY OF LAW, ARTS, AND SOCIAL SCIENCES
School of Social Sciences

# Evaluating Investment in Base Load Coal Fired Power Plant Using Real Options Approach

JURICA BRAJKOVIC

Thesis for the degree of Doctor of Philosophy

November 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF LAW, ARTS, AND SOCIAL SCIENCES
SCHOOL OF SOCIAL SCIENCES

Doctor of Philosophy

EVALUATING INVESTMENT IN BASE LOAD COAL FIRED POWER PLANT USING REAL OPTIONS APPROACH

by Jurica Brajkovic

This thesis investigates the impact of uncertainty on investment in a coal-fired power plant using a real options (RO) framework. It is organized in five chapters. In the first chapter I give an outline of the thesis.

In Chapter 2 I review the background material. I describe the electricity sector in the pre- and post-liberalization periods and discuss the implication of the transition on investment in new generation capacity. Further, I analyze the mainstream approach to investment analysis used by the majority of electricity companies, the discounted cash flow (DCF) approach. Next, I describe an alternative approach for evaluating investments, RO.

In Chapter 3 I perform an econometric analysis of dark spread prices. I select four different stochastic processes and fit them to the observed data. The goal is to find which of the four processes (arithmetic Brownian motion (ABM), Ornstein-Uhlenbeck (OU), Cox-Ingersoll-Ross (CIR) and the Schwartz one-factor process) can best describe the evolution of dark spread prices. The analysis shows that the CIR process is the most appropriate model to use to represent the evolution of dark spread prices.

In Chapter 4 I evaluate an investment in a coal-fired power plant assuming the dark spread is the only source of uncertainty and using the stochastic processes for which I estimated parameters in Chapter 3. First I calculate the optimal investment threshold using a traditional budgeting approach based on the DCF principle. Following this, using the RO framework, I calculate the optimal investment threshold for the four stochastic processes. I conclude that one should use mean reverting process to model the investment decision but the choice of mean reverting process does not significantly affect the investment threshold values.

In Chapter 5 I extend the analysis and model coal and electricity prices separately. Now the investment decision is affected by two factors: the price of electricity (output) and the price of coal (input). The goal of this chapter is to analyze whether this increase in complexity (going from a one-factor to a two-factor model) affects the result obtained in the previous chapter. Given the different dynamics of electricity and coal prices, I find that this approach enriches the investment analysis and gives additional insights. In particular, the higher the coal price, the greater the dark spread needs to be in order to undertake the investment. Finally, Chapter 6 concludes.

The thesis contributes to the existing knowledge in several ways. RO have been applied to the electricity sector before, but this is the first time they have been applied to the evaluation of investment in a coal-fired power plant. Secondly, this is the first time that dark spread, electricity and coal prices are modeled for use in a RO analysis. Finally, the thesis provides a comparison of investment analysis for a coal-fired power plant using RO based on single and two state variables, which has not been carried out so far.

# Contents

# List of Figures

# List of Tables

## Declaration of Authorship

I, Jurica Brajkovic, declare that the thesis entitled 'EVALUATING INVESTMENT IN BASE LOAD COAL FIRED POWER PLANT USING REAL OPTIONS APPROACH' and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based upon work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission.

Signed: .......................................................

Date: .......................................................

# Acknowledgments

It would be impossible for me to complete the PhD program without the support and encouragement of my family. I am grateful to all of them for being my inspiration and support throughout this journey.

My supervisor Robin Mason played a pivotal role in development of the thesis. Discussions with Robin helped me in structuring my ideas and understanding what it takes to implement them. I am also very indebted to Jean-Yves Pitarakis who offered guidance and advice in the area of time series econometrics. Both Robin and Jean-Yves had a significant impact on my development as an economist.

I wish to thank my fellow students at the Department with whom I shared this experience and who made it all more joyful, in particular: Reza Boostani, Juan Correa, Dimitrios Gkountanis, Corrado Giulietti, Alessandro Mennuni, Ronaldo Nazare, Derya Tas and Greg Taylor. Furthermore, I appreciate the support of the administrative staff at the Department for helping me solve various technical issues along the way.

I thank my colleagues at the *Energy Institute Hrvoje Pozar*, Zagreb, Croatia for useful discussions. I am very grateful to Institute's director, Goran Granic, for encouraging me to pursue the PhD degree.

The thesis was examined by Christian Schluter and Richard Green. I thank both of the examiners for stimulating discussion and useful comments. Moreover, during the examination they offered interesting ideas on potential avenues for future research that I hope to pursue.

Parts of the thesis were presented at three conferences where I received useful comments on my work: *14th Annual International Real Options conference*, Rome, Italy (Chapters 3 and 4); *Enerday: 5th Conference on Energy Economics and Technology*, Dresden, Germany (Chapters 3 and 4) and *2nd University of Bath Economic & Finance PhD Conference* (Chapter

5).

I gratefully acknowledge the financial support of the *Croatian Ministry of Education, Science and Sports* for academic years 2007/2008 and 2008/2009.

# Abbreviations

ABM Arithmetic Brownian motion process

CIR Cox-Ingersoll-Ross process

DCF Discounted cash flow

EEX Energy Exchange, Leipzig, Germany

GBM Geometric Brownian motion

GBM Geometric Brownian motion

IRR Internal rate of return

ln Natural logarithm

Max. Maximum value

MC Monte Carlo

MCLS Monte Carlo least squares

Min. Minimum value

ML Maximum likelihood

NPV Net present value

O&M Operation and maintenance costs

ODE Ordinary differential equation

OU Ornstein-Uhlenbeck proces

PB Payback period

RE Renewable energy

RMSE Root mean squared error

RO Real options

SDE Stochastic differential equations

Se. Standard error

Skew. Skewness

# Chapter 1

# Introduction

Today, in most of the European countries, the electricity sector is liberalized; competition has been introduced into the generation and distribution segment of the market. In evaluating investments in new generation capacity, practitioners mostly rely on techniques appropriate for monopolized markets. They use some kind of discounted cash flow approach (DCF) such as net present value (NPV) or internal rate of return (IRR). Such an approach was justified in monopolized markets where utilities faced almost no business risk. However, with the liberalization of the electricity sector, an approach based on DCF becomes inadequate (IEA (1999), IEA (2003)).

Investment in new capacity in the liberalized market is characterized by a large degree of uncertainty related to input and output prices, time to build, managerial flexibility and various operational options such as options to suspend, expand or shut down operations. Therefore, to value new investments properly in an uncertain environment one needs a tool capable of quantifying the effect of uncertainty on investment. To this end, in the thesis I employ real options (RO) to evaluate investment in a new coal-fired power plant.

In Chapter 2 I provide preliminary material that describes the overall context of the thesis. There I outline the characteristics of electricity markets as well as the principles behind traditional capital budgeting methods and the RO approach.

The research results are presented in Chapters 3, 4 and 5, where I evaluate investment in a new coal-fired power plant from two angles. First, I assume that the value of the project depends upon one state variable, dark spread[1].

[1]Dark spread is defined as the difference between the price of electricity and the price of coal required to generate one unit of electricity.

Following this, I analyze the investment decision assuming independent electricity and coal prices.

In Chapter 3 I estimate an appropriate stochastic process for dark spread prices. I choose among arithmetic Brownian motion (ABM), Ornstein-Uhlenbeck (OU), Cox-Ingersoll-Ross (CIR) and the Schwartz one-factor model. The purpose of this chapter is to determine which process is the most appropriate for modeling dark spread prices. The goodness of fit of the processes is assessed in two ways. Upon estimating the parameters of each stochastic process I perform Monte Carlo (MC) simulations to obtain the distributional properties of each stochastic process. Based on the MC simulations I rank the processes according to how close they come to the observed data in terms of minimum and maximum values, mean, median, standard error, skewness and kurtosis. Furthermore, as a more formal test, I compute the root mean squared error (RMSE) for each stochastic process. Using Diebold and Mariano (1995) statistic I test whether the difference in RMSE among the stochastic processes is statistically significant. As a result I find that the CIR model is the most appropriate one for modeling dark spread prices.

In Chapter 4 I determine the optimal investment policy assuming that the profitability of the power plant depends upon the dark spread. As a benchmark, I calculate the investment threshold price using NPV. Subsequently, I determine the investment threshold for each of the four stochastic processes. For the ABM and CIR processes it is possible to determine the threshold dark spread price in an analytical form while for OU and the Schwartz one-factor model I rely on numerical methods, namely on the implicit finite difference scheme. The focus of this chapter is to determine the difference in investment threshold when using DCF nad RO. Furthermore, I am interested to determine how does the choice of different stochastic processes affect the investment threshold.

In the final chapter I analyze whether a more complicated model consisting of separate electricity and coal prices yields a different result. Given that I have two state variables I use the MC least squares approach to determine the threshold electricity and coal prices. It turns out that using separate processes for coal and electricity prices enriches the investment decision process.

Chapter 2

# Preliminaries

## 2.1 Electricity markets

### 2.1.1 Liberalization process

Prior to the 1990s, most countries operated monopolized electricity markets. In such markets electricity companies were vertically integrated utilities in charge of generating electricity, passing it over transmission and distribution networks and delivering it to the final customers. Utilities were regulated in such a way that they could recoup all prudently incurred costs and earn a reasonable rate of return, through regulation schemes such as 'price cap', 'revenue cap' or 'rate of return'. In essence, they were able to pass all cost increases to the final consumer. As a consequence, utilities did not face any business risk: they were faced only with regulatory risk.

Prices in monopolized markets did not reflect actual costs of production, mainly due to utilities implementing various government policy objectives: some of the most common ones are outlined in Table 2.1. Implementing such policies implied utilities could not focus on their core business, production of electricity, but rather had to pursue sometimes conflicting goals. For instance, they could be asked to run their operations efficiently but on the other hand were also required to keep electricity prices low and to use coal from uncompetitive domestic coal mines. Even though some of these policies were justified, others were inappropriate and should have been pursued in other, more direct, ways.

| Policy objective | Implementation by utilities |
|---|---|
| **Social policy** | Keeping electricity prices below real costs |
| | Over use of labor |
| **Supporting domestic coal mining industry** | Purchase of domestic coal |
| | Utilities ownership of uneconomic coal mines |
| **Promoting energy security** | Selection of a particular fuel |
| **Promotion of renewable energy (RE)** | Utilities required to purchase electricity from RE producers |
| | Subsidized electricity prices for RE producers |
| **Pollution reduction** | Selection of particular generation technologies |
| | Selection of specific environmental control techniques, not necessarily required to meet environmental regulation |
| **Supporting domestic equipment suppliers** | Use of non-competitive procurement procedures |
| | Selection of non-commercial generation technologies |

TABLE 2.1: Policy objectives pursued by utilities in monopolized markets on behalf of governments (IEA (1999), page 21, Table 1)

It was in 1996 that the European Union introduced a directive on electricity market liberalization (*Directive 96/92/EC*). The driving idea behind market liberalization was to introduce market based pricing which should result in increased efficiency of utility companies. In the end this should translate in consumers paying prices that reflect the actual cost of production (IEA (1999)).

From the standpoint of this thesis, the major result of market liberalization is reallocation of risk. In monopolized markets it is mainly consumers who bear all the risk. In liberalized markets it is mainly investors who bear it. This in turn makes electricity generation no different than any other capital intensive market industry. Some of the most common risks in the liberalized electricity sector are given in Table 2.2.

| | Risk type | Outcome compared to expectation |
|---|---|---|
| CONSTRUCTION | Cost overruns | Construction costs more than planned |
| | Schedule delays | Long construction times: purchase of replacement power may be needed to meet the shortfall |
| | Technology | Poor plant performance, especially when using new technology |
| | | Low plant efficiency, low plant availability |
| | Financial | High plant financing costs |
| OPERATING | Market | Low electricity sales, leading to excess capacity |
| | | Major customers find alternative suppliers |
| | | Low electricity sales price |
| | O&M | High labor or material costs |
| | Fuel | High fuel costs |
| | | Inadequate fuel capacities |
| | Financial | Returns are lower than expected |
| | | Poor capital structure |
| POLICY | Market | Regulation results in higher costs |
| | | Administrative procedures cause delays |
| | | Tax burden increases |
| | Environmental | Environmental laws become stricter |
| | | Environment assessment criteria change |

TABLE 2.2: Risks in liberalized markets faced by utilities (IEA (1999), page 23, Table 2)

Because the utilities did not bear much risk in monopolized markets, they did not have an incentive to mitigate it. For instance, in large projects it was common to have cost overruns and schedule delays. But the utilities did not find it necessary to engage in aggressive negotiations with suppliers. Today for instance, investors require suppliers to guarantee timely completion of projects and also require performance guarantees which ensure the performance of equipment. These performance guarantees are backed by bank guarantees that suppliers can withdraw only when sufficient time has passed to ensure equipment works properly. The following table gives a sample of possible measures that utilities can use to minimize some of the most common risks in liberalized markets.

| Risk type | Problem | Solution |
|---|---|---|
| Costs | Construction costs more than planned | Fixed price construction contract |
| Schedule delays | Lost revenue, inability to meet contractual obligations | Penalty fees |
| Technology | Poor plant performance, low plant efficiency, low plant availability | Performance guarantees |
| Financial | High plant financing costs | More advanced capital structure |
| Market | Low electricity sales, leading to excess capacity | Hedging |
| | Major customers find alternative suppliers | |
| | Low electricity sales price | |
| O&M | High labor or material costs | Contracts with special purpose companies |

TABLE 2.3: Possible measures to manage risks

The fact that the electricity industry was monopolized had an impact on how investments were evaluated. The next section goes into detail showing how the investment environment has changed and what it means for project evaluation in electricity generation.

### 2.1.2 Investment appraisal

In a monopolized market investment is made by optimizing the whole system, i.e., by minimizing the system cost. This is done by utilities making long-term projections regarding demand, capital costs, operational performance, economic lifetime, electricity and fuel price. Based on these predictions utilities make investment decisions. In ideal circumstances this approach would mimic optimal market outcome. Nevertheless, frequently, the result of this approach was vastly different from expected (IEA (1999)).

For instance, utilities tended to predict too high a growth in electricity demand which resulted in substantial overcapacity, as shown in Table 2.4. There are several reasons for this upward bias. Firstly, it was acceptable to have very high level of reserve margin. There was no measure to incentivize utilities to accurately predict demand growth, i.e., to penalize them for upward bias in their predictions. Secondly, utilities were slow in recognizing decrease in electricity demand growth which kept slowing down in the post-war period. Table 2.5 illustrates demand growth rates for the US where we can see that after a period of strong demand growth in the post-war period, demand growth slowed down in the years after the 1970s.

| Country | 1985 | 1990 | 1995 | 1997 |
|---|---|---|---|---|
| Finland | 30 | 44 | 29 | 26 |
| France | 31 | 39 | 38 | 39 |
| Germany | 27 | 25 | 29 | 27 |
| Greece | 42 | 42 | 32 | 30 |
| Iceland | 42 | 29 | 32 | 22 |
| Ireland | 34 | 32 | 24 | 22 |
| Italy | n.a. | 36 | 40 | 41 |
| Japan | 35 | 27 | 26 | 31 |
| Netherlands | 43 | 39 | 41 | 31 |
| New Zealand | 37 | 29 | 32 | 28 |
| Norway | 27 | 37 | 27 | 31 |
| Spain | 46 | 42 | 44 | 44 |
| Sweden | 27 | 32 | 27 | 27 |
| Switzerland | 49 | 45 | 46 | 47 |
| Turkey | 37 | 44 | 32 | 23 |
| UK | 21 | 26 | 21 | 21 |
| US | 30 | 26 | 20 | 22 |

TABLE 2.4: Reserve margins in % (IEA (1999), page 50, Table 5)

| Period | Electricity growth rate (%) |
|---|---|
| 1952-1962 | 8.7 |
| 1962-1972 | 7.3 |
| 1972-1982 | 3.9 |
| 1982-1992 | 2.9 |
| 1992-2002 | 2.2 |

TABLE 2.5: Average annual electricity demand growth rates in the US (Administration (2010))

As another example, once the utilities decided on an expansion path, their decision was frequently modified by the government or regulator. Governments required utilities to pursue certain policy measures (such as those in Table 2.1) while regulators approved only those investments that were justified in their opinion. Consequently, the final outcome of centrally planned investment decision process turned out to be less than optimal: among other issues, companies operated more capacity than was needed and prices did not reflect actual costs.

Utilities have used a levelized cost approach for decades in evaluating investment decisions[1]. Implementation of the levelized cost approach involves projecting costs, energy production and revenues in the future and discounting it back to the present value. Utilities use some predetermined discount rates to discount the cash flows. Then they choose the technology having the highest present value. To deal with risk, utilities develop various scenarios.

Given that the business environment in a monopolized market is stable, this approach served its purpose. In liberalized markets, however, this approach is not appropriate to handle the complexity of the investment environment because it makes several implicit assumptions which do not hold in competitive markets:

- **All technologies are similar.** This used to be the case several decades ago when there were indeed only a few similar technologies to choose from. Nowadays, due to technological progress, there is plethora of technologies an investor can choose from and each of them has some unique characteristics (more on this in Section 2.1.3).
- **All costs are passed to final consumers.** This is true for monopolized markets, but in competitive markets less competitive producers cannot pass their costs on to consumers.
- **There is strong growth in demand.** As Table 2.5 illustrates, electricity demand was declining rapidly in the second part of the 20th century which led to overcapacity as shown in Table 2.4. The fact that demand growth is not persistent and predictable increases uncertainty surrounding new investments.

What matters to the investor in liberalized markets is risk related to the profitability of the project, as shown in Figure 2.1. The more risky the project is, the higher the rate of return the investor requires. Within the levelized cost approach this issue is dealt with using different hurdle rates and scenarios, which are insufficient in competitive markets. As for the scenario analysis, it does not resolve the issue of risk, as there is nothing which tells the investor which scenario to look at. The fact that power plant, for instance, can become very profitable, profitable or not profitable

[1] A good review of the method is given in IEA (2005). Many utilities used programs similar to the *Wien Automatic System Planning* (WASP) program developed by the IEA which is based on a levelized cost approach to determine investment schedules. A major drawback is that such programs cannot address the issue of risk and uncertainty in decision making nor do they take into account economic incentives (IEA (2007)). Such programs are appropriate only for centrally planned monopolized markets.

(representing high, medium and low scenario) does not help the investor make an investment decision. As far as hurdle rates are concerned, they represent improvement over the scenario analysis but still do not capture the whole issue of risk. Questions such as what is the appropriate hurdle rate, why is it constant and how is it determined, remain unanswered.

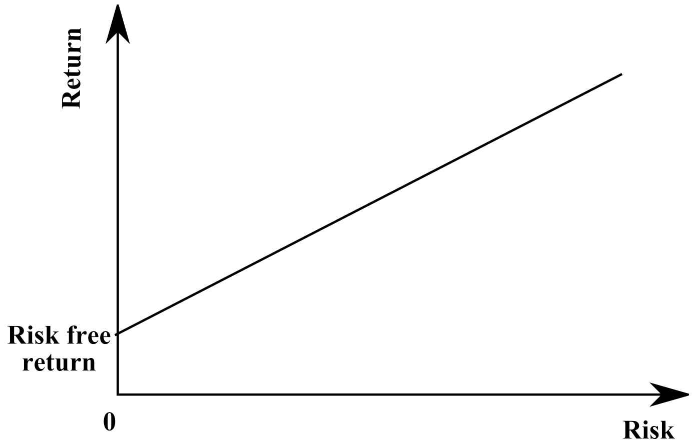



FIGURE 2.1: Required rate of return as a function of project risk

Practitioners have come up with improvements to solve the issue of risk, including probabilistic analysis. But the main problem still remains: the levelized cost method relies on a discounted cash flow (DCF) approach, and this approach by construction only looks at average values without taking any regard to riskiness of the project (e.g. extreme input and output prices) or managerial flexibility (ability of the managers to act and limit the downside and exploit the upside).

### 2.1.3 Impact of evaluation methodology on investment choices

Electricity generation is a complex business, and the risk associated with it depends upon the selected technology. The following table shows the risk profile for various technologies. As can be seen, a potential investor has a full range of technologies to choose from: the question is how to decide on the appropriate choice.

| Technology | Unit size | Lead time | Capital cost/kW | Operating cost | Fuel cost | C02 emis-sion | Regulatory risk |
|---|---|---|---|---|---|---|---|
| CCGT | Medium | Short | Low | Low | High | Medium | Low |
| Coal | Large | Long | High | Medium | Medium | High | High |
| Nuclear | Very large | Long | High | Medium | Low | None | High |
| Hydro | Very large | Long | Very high | Very low | None | None | High |
| Wind | Small | Short | High | Very low | None | None | Medium |
| Photovoltaics | Very small | Very short | Very high | Very low | None | None | Low |

TABLE 2.6: Risk profile of production technologies (IEA (2003), page 32, Table 2)

The fact that in a liberalized market electricity and fuel prices are non-constant affects profoundly the profitability of power plants. In turn, it influences the choice of investment technologies and results in a different outcome when compared to monopolized markets. Technologies with higher investment costs and long lead times but low variable costs such as nuclear are being shunned. Even though they have very low variable costs, the fact that they cannot be easily turned off when prices slump, makes them very unpopular in liberalized markets. On the other hand, technologies such as gas plants which have short lead times, relatively low investment costs and can be turned off quickly but have high fuel costs are being favored by electricity companies in liberalized markets. This tendency is illustrated in Figure 2.2 where it can be seen that gas turbines represent 61% of all capacity ordered between 1991 and 2002 (IEA (2003)).

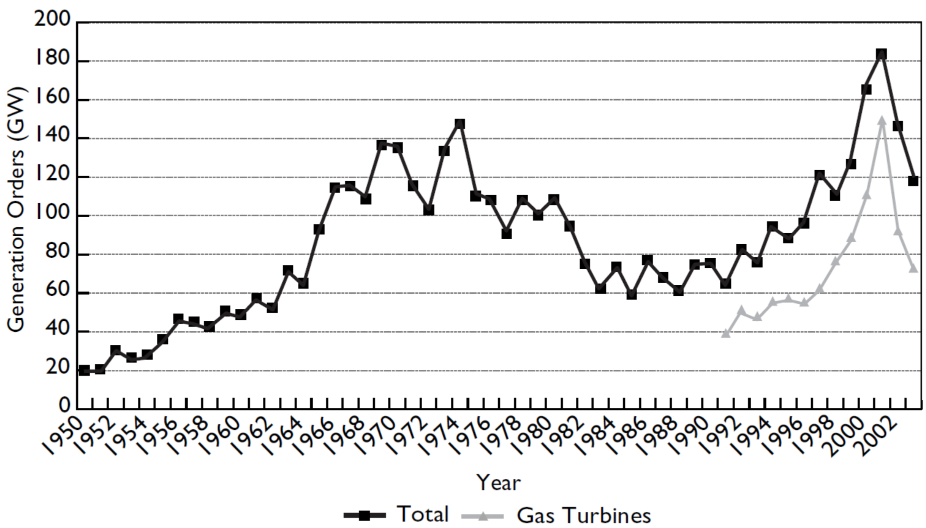



FIGURE 2.2: New orders of gas turbines (IEA (2003), page 23, Figure 1)

The reason why utilities are staying away from more capital intensive technologies, such as nuclear, and using more gas, lies in the way they evaluate investments. For instance, nuclear power plants that have very long lead times, high construction costs but low production costs cannot prove their profitability through a levelized cost approach due to the nature of that framework and the way investments are structured. As the study by the *Texas Institute for the Advancement of Chemical Technology* (2005) shows, nuclear power plants are not necessarily uncompetitive: they are only uncompetitive if one uses a DCF framework. Using an options approach where one resolves uncertainty by making incremental investments, i.e., by resolving different sources of uncertainty, one can structure investment in nuclear power in a way that minimizes the downside but maximizes the upside potential.

In the next section I will describe the principles of DCF approach and point to its weaknesses which make it an inappropriate tool for investment analysis in liberalized electricity markets.

## 2.2 Traditional capital budgeting

To make investment decisions companies predominantly use a DCF approach. A study by Graham and Harvey (2001) showed that the vast majority of US companies use net present value (NPV) , internal rate of return (IRR) or payback (PB) period in valuing new investments (Figure 2.3). The foundations of this approach have been laid down by Irving Fisher who formalized the notion of interest and time value of money in the first part of the 20th century (Brennan and Schwartz (1985), Moyen et al. (1996)). Since then, the approach has been refined but still serves as the main tool in valuing new investments.

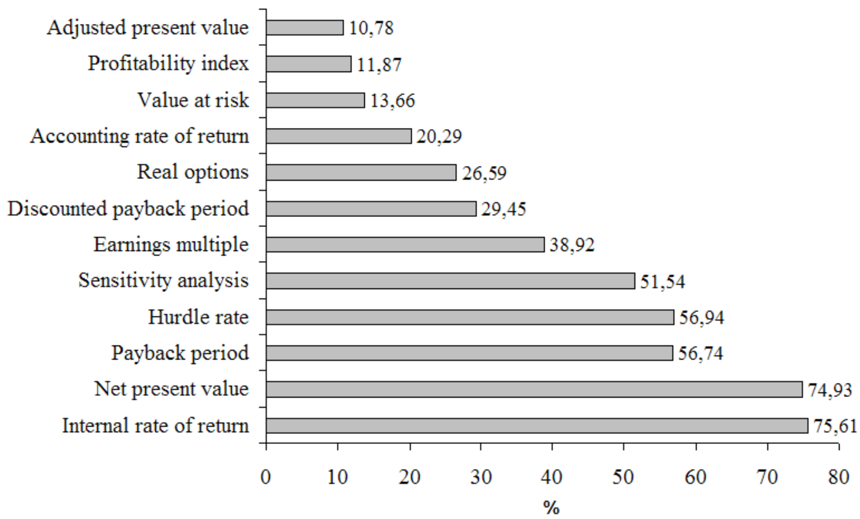



FIGURE 2.3: Percentage of CFOs who always or almost always use a given technique (Graham and Harvey (2001), page 197, Figure 2)

The reason for the popularity of DCF lies in its simple and intuitive nature. In order to determine the value of a project under the DCF framework one needs to forecast cash flows for revenues and costs, usually at an annual frequency. Following from this, one subtracts costs from revenues to arrive at the profit for each period. At this point there are several options available to determine the profitability of the project: most popular methods include NPV and IRR. As NPV and IRR use the same underlying logic, I will focus on NPV in the remainder of this section.

One discounts profits for each period, adds the profits of each period up and then subtracts from this sum the investment cost to finally arrive at the NPV of the project. If NPV is positive, the project should be undertaken because it is worth more to the firm than it costs. Otherwise it should be rejected. Equation 2.1 gives a formal expression to calculate the NPV of a project.

$$NPV = \sum_{t=i}^{T} \frac{E(\pi_t)}{(1+k)^t} - I \tag{2.1}$$

where definitions are as follows: $E(\pi_t)$ denotes expected profit at time $t$; $k$ denotes discount rate; $I$ denotes investment cost; $T$ denotes project life.

A crucial element in DCF methods is played by discount rate ($k$). Discount rate should adjust each period's cash flow to reflect two phenomena:

- **Time value of money**. This principle implies that cash flow today is worth more than cash flow tomorrow. To adjust cash flows for time value of money, one uses a risk-free interest rate.
- **Riskiness of a project**. Investors should be compensated for the risk they undertake when investing in a risky project. But it is not the overall project risk they should be compensated for, but only non-diversifiable or systematic (market) risk. In other words, investors should not be compensated for risk specific to the project (idiosyncratic risk) because they can eliminate it by investing in different projects. They should be compensated for the risk inherent in the market, risk that cannot be diversified away.

Therefore, to correctly discount the cash flows we should take into account time value of money and riskiness of the project. The sum of time value of money and riskiness of a project can be seen as an opportunity cost of capital. This opportunity cost represents a reward investors require when investing in a project of a given level of risk.

To calculate the opportunity cost of capital ($k$), also known as risk adjusted discount rate, one can use Capital Asset Pricing Model (CAPM) given by:

$$k = r_f + \beta(r_m - r_f) \tag{2.2}$$

where the variables have the following meaning: $r_f$ denotes risk-free rate (such as a return on short-term government bills); $r_m$ denotes return on a broad market portfolio (such as S&P 500). In effect $r_m - r_f$ can be interpreted as a market risk premium: how much should one require in terms of risk premium for holding a basket of securities which replicates the whole market. Furthermore, opportunity cost of capital ($k$) equals risk-free rate plus (or minus) market risk premium adjusted by proportionality factor beta.

While we can easily observe the value of risk-free rate and return on market portfolio, we need to calculate the value of beta. It measures the relation between the return on a security of the same level of risk as an analyzed project (assuming such security is traded) and the return on the market portfolio. Beta is calculated as follows:

$$\beta = \frac{Cov(r_i, r_m)}{Var(r_i)} \tag{2.3}$$

where I define as follows: $Cov(r_i, r_m)$ is the covariance between the return on a project $i$ of similar risk and market portfolio; $Var(r_i)$ is the variance of return on the project.

As this section shows, calculating project value using NPV is straight forward. The next issue I discuss is whether NPV is an appropriate tool to value investments in an uncertain environment.

### 2.2.1 Drawbacks of NPV analysis

NPV is a good method to evaluate investments in stable and static environments. It assumes prices and costs are predictable, and once a project starts, management does not interfere with it but passively sits on the side until the end of the project. Unfortunately, these are not realistic assumptions for most projects. In most projects, prices and costs are stochastic, investments are irreversible and management actively manages projects.

When evaluating a project under an NPV framework, the project should be undertaken as soon as NPV$\geq$0. This rule implies that an investment opportunity will vanish if not undertaken immediately. But this is only the case in limited circumstances. Actually, each project derives value from three different sources which have to be taken into account when making investment decisions (Ross (1995)):

- **Intrinsic value**. This is just the discounted value of the project, i.e., NPV of the project if the investment is made immediately.
- **Option to invest**. Opportunity to invest has value. Even though a project might not be economical at one point in time, it does not mean market conditions will not change in favor of the project. Therefore, to be able to invest has value.
- **Options embedded in the project**. This includes managerial and operational flexibility of the project.

The NPV approach only recognizes the first element of the investment decision, the intrinsic value of the project, while other elements are neglected. Focusing only on the intrinsic value of the project results in the NPV having some serious shortcomings, some of which I state now.

**SHORTCOMING I: Assumption that investments are made on now-or-never basis**

Most projects do not have to be undertaken immediately. Rather, the investor can choose the appropriate timing of the investment and hence acquire new information that can alter the investment decision. This aspect is not important if the investment environment is certain; for most projects, however, the investment environment is not certain.

To demonstrate the value of waiting (i.e., the ability to choose the timing of investment), I will use a simple example. Assume a plant costs 100, it can be built instantaneously and operates forever. Once built, the plant produces 1 unit of output forever at the cost of 40/unit. During the first year the revenue is 50/unit. From the second year on, *expected* revenue is 50/unit: there is a 50% chance that revenue will be 70/unit and a 50% chance that it will be 30/unit. Assume a discount rate of 10%. Value of the project is easily calculated using NPV and equals 10:

$$NPV = \sum_{t=0}^{\infty} \frac{E(\pi_t)}{(1+r)^t} - I = 10$$

Even though the NPV is positive, we know we should not invest immediately because of a likely price decrease. If the price in the second year goes up to

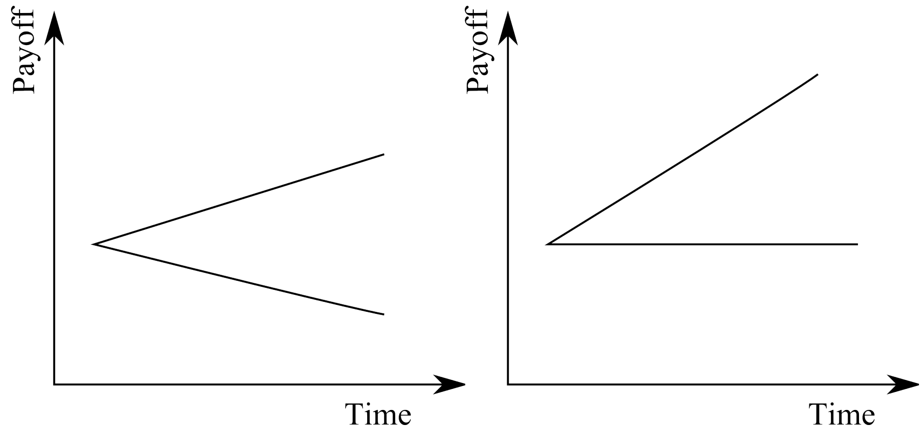
70/unit the investment will be more than profitable. But if the price drops to 30/unit the investment will be a bad idea. Therefore, even though NPV suggests the investment should be undertaken, it is optimal to wait a year for uncertainty to be resolved.

**SHORTCOMING II: In performing NPV analysis one has to forecast future cash flows and use point estimates**

In doing NPV calculations one uses point estimates for state variables which, in most cases, are arbitrary and subjective. This practice is extremely unrealistic as price and cost variables can take any value as they change over time. As an example, assume a project that once completed gives a certain one-time payoff of 110 million (Ross (1995)). Assume also, it takes a year to build the project and the investment cost is 100 million. Because the project payoff is certain, I use risk-free discount rate to calculate the project value. For simplicity, assume the risk-free rate is 10.5%. Under these assumptions the NPV of the project is 99.5 million, and the project should not be undertaken because it is worth less than it costs. But it is common knowledge that the risk-free rate changes over time. Therefore, even though the NPV says the project should be abandoned we know that if the risk-free rate goes below 10% the project will make economic sense. Thus, we can draw the following lesson from this example: Even though NPV suggests the project should be abandoned, we know the project is not worthless. It has value, and this value comes from the possibility that the risk-free rate will drop, making the project feasible. NPV gives a misleading result because it uses point estimate.

**SHORTCOMING III: NPV assumes projects will be managed passively**

NPV does not make any provision for active management. In real life management does not just sit aside and observe what happens with the project but rather it takes an active role. Thus, if a situation turns out better than expected, management takes advantage of the current situation and improves the outlook for the project, by expanding the upside potential. On the other hand, if market conditions turn out to be worse than expected, management acts to limit losses, eliminating the downside of the project. This is shown in Figure 2.4 where the left figure shows returns under NPV where the project is assumed to be managed passively, and where the right figure shows returns under RO where the project is assumed to be managed actively (Amram and Kulatilaka (1998)).




(A) DCF payoff (B) RO payoff

FIGURE 2.4: Payoff under NPV and RO valuation

**SHORTCOMING IV: NPV does not see projects as option**

NPV suggests projects should be undertaken as soon as NPV≥0 and not undertaken if NPV<0, as shown in the above example. On the other hand, sometimes even if the NPV is negative, it might be beneficial to go forward with the project. This is especially the case when a company wants to secure a strategic advantage by entering a market early or, for example, when it introduces a new product. By sustaining initial losses a company can learn a lot about the market. It gives the company an option to expand the production in later phases if the initial launch turns out to be successful. If the initial launch turns out to be unsuccessful, the company has limited its losses. In other words, initial investment gives the company an option for further actions.

**SHORTCOMING V: NPV assumes investments can be undone**

Given that NPV suggests investing as soon as NPV>0 even though the profits might drop and make the project uneconomical (such as in the example in Shortcoming I), one can claim that NPV implicitly assumes investments can be undone. While this might be true for some investments (e.g., the purchase of a house: you can easily sell it if you don't like it), most of the investments are irreversible: once the investment is made it is practically impossible to undo it. If an investor builds a new power plant and realizes it is not profitable, it cannot sell it on the secondary market: it remains stuck with it. Therefore, when evaluating investments one should take into consideration that the investment environment might change. Therefore, a simple rule that one should invest as soon as NPV>0, should be augmented by saying that one should invest when NPV is sufficiently large.

By investing when NPV is sufficiently large, the investor protects him/herself from future adverse market movements[2]. And this is very important as once the investment is made, it cannot be undone.

## 2.3 Real options

The above shortcomings of NPV framework have been known to practitioners and researchers for some time. The source of the problem with NPV is that the approach is not suited to valuing investments in an uncertain environment. It was envisaged to value annuities which by definition are payments of predetermined size, and as such, is inappropriate to value investments that are characterized by uncertainty, managerial and operational flexibility. An alternative to NPV emerged when Black and Scholes (1973) presented a revolutionary paper on option valuation. As I will show, unlike NPV, the options approach is perfectly suited to value investments in real projects.

In their paper, Black and Scholes derive a pricing formula for an option on a stock. A call (put) option gives a holder a right, but not an obligation to buy (sell) a security at a predetermined price, called a strike or exercise price, at a future date called maturity. If the option can be exercised only at maturity, it is called European option. If it can be exercised at any other time prior to expiration, it is called American option.

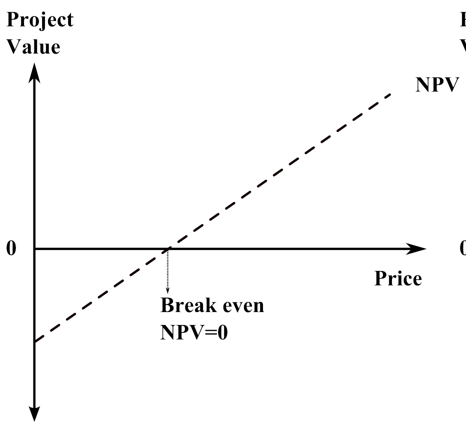
Investment opportunity can be seen as a call option on the project. If the investment does not need to be made immediately, then the investor has an option but not an obligation to invest in the project. Because for most projects the investment opportunity stays open for a certain period of time during which the investment can be undertaken, the investment option can be seen as a (perpetual) American option.

| Real option | Variable | Financial option |
|---|---|---|
| Investment cost | X | Exercise price |
| Value of the project | S | Stock price |
| Amount of time decision can be deferred | t | Time to maturity |
| Riskiness of underlying variable | $\sigma^2$ | Stock volatility |
| Discount rate | r | Risk-free rate of return |

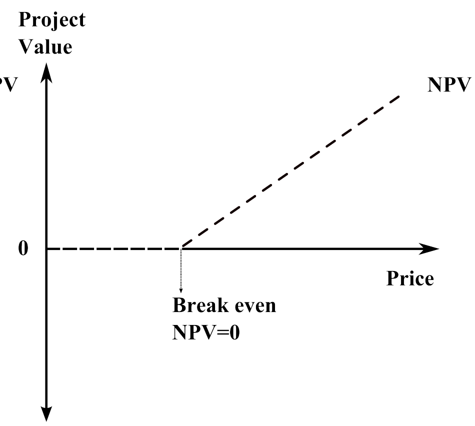
TABLE 2.7: Link between financial and real options

[2]Actually, the next section will show what is meant by 'NPV being sufficiently large'. In short, one should invest when NPV exceeds investment cost plus the value of option to invest.

Just as in the case of financial options, where value is strictly positive, value of the investment option is also strictly positive. This is in contrast to NPV where projects can even have a negative value. Figure 2.5a shows a typical NPV valuation where NPV is negative for prices below break-even price. This implies that if an investor were to sell this project before it was developed, he/she would have to pay someone to take it because NPV is negative: this is not very likely to occur in practice. A modified version of NPV shown in Figure 2.5b says that a project does not have any value if the price is below break-even price. This is an improvement over the most basic NPV analysis but still does not reflect the full reality of investment projects.




(A) Simple NPV




(B) Adjusted NPV

FIGURE 2.5: NPV valuation

Even though a project might not have a positive NPV, we still know it might be worth something. As was pointed out in Ross (1995), a part of a project's value is derived from the option to invest. For low prices, when NPV is negative, a project still has value: this is shown in Figure 2.6.

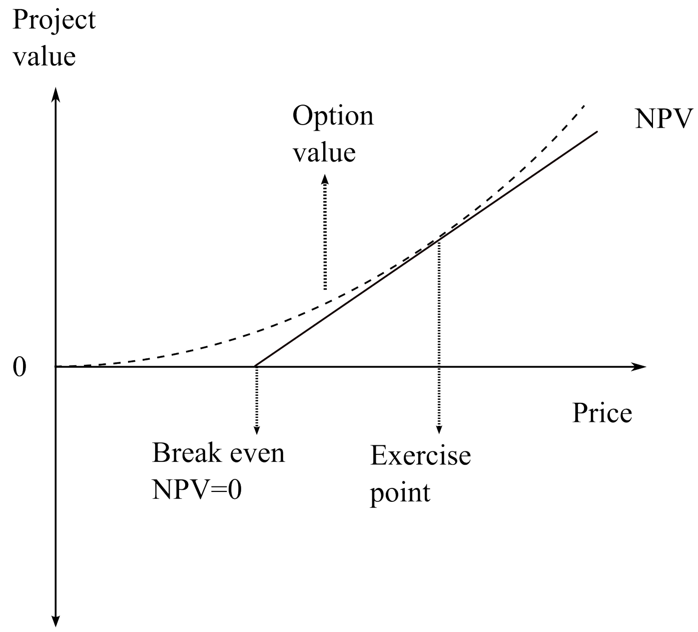



FIGURE 2.6: Option valuation

Figure 2.6 shows the logic of (real) option valuation. For below break-even prices, the option to invest (shown as a dashed line) in the project has a value, unlike the modified NPV framework where this value is zero. As the price rises, the value of the option increases and is actually greater than the intrinsic value of the project (NPV). Option theory gives us guidance on when to invest in the project. This should not happen as soon as NPV becomes positive, but rather when NPV is significantly positive. Because investment decision cannot be reversed, we have to take into account the opportunity cost of investing now rather than later. Therefore, as the option to invest has value, we should invest when the value of the project is greater than the investment cost plus option to invest. Assuming only one state variable $P$, I can state the following rule on investing[3]:

**Definition** Invest in a project when: $V(P) = F(P) + I$, where $V(P)$ denotes project value, $F(P)$ option value and $I$ investment cost. ■

Formally, optimal investment decision under option valuation framework is determined using two rules:

- **Value matching**. It implies investment should be done at a point where the following relation holds: $F(P) = V(P) - I$
- **Smooth pasting**. It implies that not only do the value of the option and NPV line ($V(P) - I$) have to equal, but they have to meet tangentially, i.e., $\partial F(P)/\partial P = \partial V(P)/\partial P$.

[3]One variable case can be easily extended to multi-variable case.

The RO solve shortcomings present in NPV analysis in the following way:

- Unlike NPV that assumes investments should be made on now or never basis, i.e. as soon as NPV>0, RO suggest investment should be made once a critical threshold is reached. For instance, invest once a critical price $P^*$ is achieved: in essence, RO assume investment opportunity will be around for some time, and the investor should find the optimal moment to invest. By investing when NPV is sufficiently large ($P^*$>(NPV>0)), RO acknowledge the irreversible nature of investments and uncertainty inherent in these investments.

- While NPV uses point estimates for state variables, in RO valuation one uses stochastic differential equations (SDE) to describe the evolution of state variables. The parameters used in SDE are usually estimated econometrically while they can also be calculated using an engineering approach. For instance, volatility of the prices could be estimated from historical data while long-run price level could be calculated using engineering data, or more preferably, inferred from long-term contracts.

- RO can explicitly incorporate the effect of managerial flexibility. For example, we can assign value to options that allow managers to expand production in good times and contract it in bad times.

- RO can value operational flexibility that comes with some technologies. For instance, we can easily value a peaking gas-fired power plant that operates only during periods of extremely high prices.

Finally, it should be said that RO does not represent a completely different approach to investment evaluation when compared to NPV analysis. Rather, NPV analysis is encompassed within the RO framework and inputs used in NPV analysis are also used in RO analysis. Actually results obtained under NPV and RO analysis will be the same if there is no uncertainty. Therefore, NPV represents a special case of RO analysis.

# Chapter 3

# Econometric modeling of dark spread prices

## 3.1 Introduction

When analyzing investment decisions in a coal-fired power plant, two factors play a crucial role: input prices (price of coal) and output prices (price of electricity). These two components represent a major source of uncertainty for potential investors[1]. Another significant component is investment cost, but risks related to this can be properly mitigated[2]. Furthermore, there are additional costs such as maintenance and labor costs but they are predictable and not significant, thus they do not affect the level of risk of the project.

In this chapter, a one-factor stochastic model is developed to model evolution of dark spread prices. Dark spread is defined as the difference between the price of electricity and price of coal required to generate one unit of electricity. One can think of dark spread as representing the profit flow for the power plant owner. Models estimated in this chapter are used in Chapter 4 as input to real options (RO) analysis. To model dark spread

[1] Another important source of uncertainty are emission constraints. The impact of emission constraints on future investments in coal-fired power plants is not clear as there exists no final global emission framework for the post-Kyoto period. Nevertheless, a merchant producer has several options at his disposal to mitigate this risk. For instance, it can build a power plant in a country that will have less stringent emission policy. Such is a project by the privately held EFT company to build the 'Stanari' coal power plant in Bosnia and Herzegovina. Another option is to build a coal power plant that is ready to be fitted with technology to reduce emissions.

[2] For instance, potential investor can obtain binding offers from power plant equipment producers regarding the price of equipment. The same could be done with other construction work. This in effect eliminates any construction risk such as escalation of costs or time to build.

prices, four different stochastic processes are chosen: arithmetic Brownian motion (ABM), Ornstein-Uhlenbeck (OU), Cox-Ingersoll-Ross (CIR) and Schwartz one-factor process.

Within the RO framework in which the optimal investment decision is calculated, only a few stochastic processes have analytical solutions: ABM and CIR are two of them. Other require use of numerical procedures to derive optimal investment policy. For investment problems consisting of one state variable, the most appropriate numerical method is finite differences (FD). Nevertheless, the FD method cannot deal with more complex stochastic processes, which limits the number of models that can be used. For example, it is extremely difficult to use models such as GARCH or regime switching models in valuing real options using FD. Most of these more sophisticated models are used to explain short-term variations in dark spread prices, such as temporary price spikes. I claim that the ability to explain short-term deviations is not crucial for long-term investment analysis. Therefore, the extra benefits these more complex processes have over the four processes that I selected to model the evolution of dark spread prices (ABM, OU, CIR and Schwartz) do not matter too much for long term investment analysis.

Once the parameters of the stochastic processes are estimated, a test is performed to see how well each fits the data. Testing is done in two ways. First, a less formal approach that is frequently used in the literature is adopted where 10,000 simulations are run for each process and their distributional properties are estimated (following Geman and Roncoroni (2006) and Seifert and Uhrig-Homburg (2007)): minimum and maximum value, mean, mode, standard error, skewness and kurtosis. Each of the processes is judged on the basis of the estimated parameters compared with observed data. Secondly, in a more formal approach, the root mean squared error (RMSE) is computed for in sample forecasts. Given this, different values of RMSE are obtained from different models and a test is performed to determine whether this difference is statistically significant: for this, the Diebold and Mariano (1995) test statistic is employed.

## 3.2 Literature review

It is now widely accepted that commodity prices follow some kind of mean reverting process (e.g. Pilipovic (1997), Pindyck (1999), Clewlow and Strickland (2000)). One can offer simple economic intuition as to why commodity prices should be mean reverting.

Equilibrium prices should reflect the marginal cost of production. When prices start to rise, new competitors with higher marginal costs will enter the market, increase the supply and thus exert downward pressure on prices. By the same token, prices cannot fall too low because a decrease in prices will drive less competitive companies out of business, reducing supply, which will in turn push prices up. Therefore, a model of dark spread prices should be mean reverting.

Additional properties of electricity prices include seasonality, non constant volatility and large price jumps ('spikes'). Researchers have proposed numerous models to capture such behavior including GARCH type, regime switching and multi-factor models (e.g. Schwartz (1997), Knittel and Roberts (2005), Geman and Roncoroni (2006), Seifert and Uhrig-Homburg (2007)). These models have been developed primarily for the purpose of derivative pricing or electricity price forecasting and not for RO valuation. Valuing RO adds an extra layer of complexity which prevents some of the proposed models being useful due to implementation difficulties.

To my knowledge, only a paper by Abadie and Chamorro (2009) deals with estimation of dark spread prices. The authors assume dark spread prices that follow the Ornstein-Uhlenbeck (OU) process. A similar paper is by Nasakkala and Fleten (2005) who model the development of spark spread prices[3]. The authors use a two factor model to describe the evolution of spark spread. I am not aware of any research paper investigating the appropriate stochastic model of dark spread prices for use in RO analysis.

## 3.3 Price characteristics

The coal prices used in the analysis are obtained from McCloskey Co., are published weekly and are given in €/GJ. The energy efficiency of a coal-fired power plant is assumed to be 36%, thus the total requirement for the production of 1 MWh of electricity is 10 GJ of coal.

Electricity prices are obtained from the European Energy Exchange (EEX) in Leipzig, Germany. Given that coal prices are observed on a weekly basis, electricity prices had to be converted to a weekly basis also. Conversion of electricity prices could have been done in two ways. The first approach would be to use volume-weighted weekly prices. The alternative approach,

[3]Spark spread is defined in the same way as dark spread but with gas instead of coal.

which is used in this chapter, is to first obtain arithmetic daily averages (this is actually provided by the EEX through its Phelix base index) and then to compute weekly arithmetic averages. The logic for using arithmetic averages is as follows. I analyze an investment in merchant base load coal-fired power plant. This plant operates 24 hours a day and sells all of its output on the exchange. Its output is very small compared to the overall market size and hence the plant is a price taker: it has to take whatever market price prevails at each hour. Given the assumption that the plant sells all of its output at each hour, the fact that market volume is time-varying does not affect the merchant's power plant operating decisions.

Both coal and electricity price series go from January 1, 2002 to November 13, 2009 and they consist of 410 weekly observations in total. Figure 3.1 shows graphs of electricity and coal prices for the observed period, while a plot of the dark spread prices for the same period is given in Figure 3.2. Summary statistics for electricity, coal and dark spread prices are reported in Table 3.1.

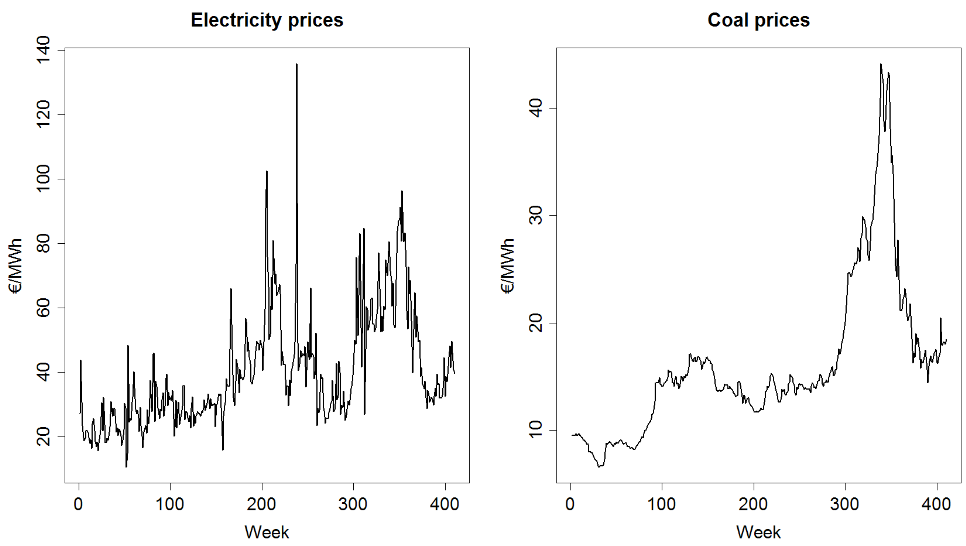



FIGURE 3.1: Weekly electricity and coal prices

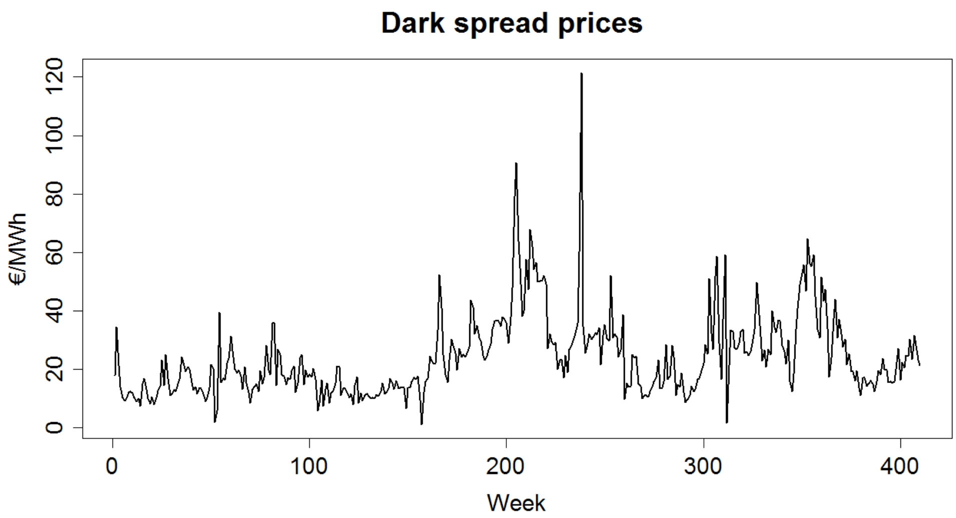



FIGURE 3.2: Weekly dark spread prices

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Electricity prices | 10.70 | 35.52 | 40.07 | 135.69 | 17.51 | 1.34 | 5.39 |
| Coal prices | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| Dark spread | 1.09 | 20.27 | 24.01 | 121.39 | 14.07 | 1.93 | 9.86 |

TABLE 3.1: Summary statistics for electricity, coal and dark spread prices

Table 3.1 shows the major characteristics of dark spread prices that I model. Namely, dark spread prices are characterized by positive skewness and large kurtosis. Positive skewness means that most of the observations are concentrated in the left tail of the distribution while large kurtosis implies the occurrence of large dark spread prices, i.e. 'spikes'.

## 3.4 Arithmetic Brownian motion

One of the most simple stochastic processes which is commonly used in finance is Brownian motion. Despite the fact that Brownian motion does not appear to be a good candidate for modeling dark spread prices, it is evaluated for the following reason[4]. It is an extremely simple process that has an analytical solution. If by any chance Brownian motion could give results comparable to ones obtained using more realistic mean reverting processes, then in some cases it would be reasonable to sacrifice accuracy for the sake of simplicity. Namely, Brownian motion could be used as a starting point in valuing more complex options.

[4]Brownian motion is an unbounded process meaning it can reach extremely high or low levels.

Given that, theoretically, dark spread prices can become negative, the arithmetic version of Brownian motion (ABM) is used to describe their evolution[5] The ABM process for dark spread prices ($\boldsymbol{p}$) is given by the following expression:

$$dp_t = adt + \sigma dz_t \tag{3.1}$$

Equation 3.1 states that the change in the value of dark spread ($dp_t$) consists of two parts: a deterministic and constant drift $a \in [-\infty, \infty]$ and a stochastic component driven by constant volatility $\sigma \in [0, \infty]$ and an increment of Wiener process $dz = \epsilon\sqrt{dt}$ where $\epsilon \sim N(0,1)$ and $Cov(\epsilon_t, \epsilon_s) = 0$ for $t \neq s$, and where $dt$ denotes the time increment. The expected value and variance of dark spread following ABM are given by the following two expressions:

$$E[p_t] = p_0 + at \tag{3.2a}$$

$$Var[p_t] = \sigma^2 t \tag{3.2b}$$

Equation 3.2a states that if the dark spread follows ABM, its expected value will increase linearly at a rate $a$ as time passes. Equation 3.2b states that the variance of the ABM process increases linearly with time as well.

### 3.4.1 Parameter estimation

Given a vector of observed dark spread prices $\boldsymbol{p}$, parameters $a$ and $\sigma$ are estimated and are considered to be elements of a parameter vector $\boldsymbol{\theta}$. For this, the maximum likelihood (ML) approach is adopted. Furthermore, a probability density function for dark spread ($\boldsymbol{p}$) conditioned on a set of parameters $\boldsymbol{\theta}$ is denoted as $\boldsymbol{f}(\boldsymbol{p}|\boldsymbol{\theta})$. Given that a variable that follows ABM is normally distributed, the likelihood ($L$) and log-likelihood (l$nL$) functions (where $ln$ denotes natural logarithm) are given as follows:

$$L(\boldsymbol{\theta}|\boldsymbol{p}) = \sum_{i=2}^{n} lnf(p_i|\boldsymbol{\theta}) \tag{3.3a}$$

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = -(n-1) \cdot ln\sqrt{(Var(p))} - \sum_{i=2}^{n} \left[ \frac{(p_i - E(p))^2}{2 \cdot Var(p)} \right] \tag{3.3b}$$

Using the expression for expected value of dark spread given in Equation

[5]One could think of using a geometric version of the same process (GBM). This option was actually tested, but the simulation results turned out to be extremely unsatisfactory.

3.2a, the log-likelihood function becomes:

$$lnL = (\boldsymbol{\theta}|\boldsymbol{p}) = -(n-1) \cdot ln\sigma - \sum_{i=2}^{n} \left[ \frac{(p_i - p_{i-1} - a)^2}{2\sigma^2} \right] \tag{3.4}$$

Values of estimated parameters are given in Table 3.2 where standard errors are given in brackets. The value of the drift parameter is positive and rather insignificant economically: when converted to annual values, the drift is equal to 0.42 € per annum. On the other hand, volatility is very high and on an annual basis it is equal to 71.4 €[6].

| | Weekly values |
|---|---|
| Drift ($a$) | 0.0081 (0.489) |
| Volatility ($\sigma$) | 9.9 (0.346) |

TABLE 3.2: Estimated weekly parameters for dark spread prices following ABM process

After estimating the parameters, I determine how well the process fits the real data. As a first measure of the goodness of fit, 10,000 simulations of ABM were run using the estimated parameters. To simulate ABM, the expression from Equation 3.5 is used where $p_t \sim N(p_0 + at, \sigma^2 t)$. Given that weekly data are used to estimate the process and the aim is to simulate weekly trajectories, a value of $1$ is put in $\Delta t$.

$$p_t = p_0 + a\Delta t + \sigma \cdot \sqrt{\Delta t} \cdot N(0,1) \tag{3.5}$$

In performing the simulations, it is necessary to select an initial dark spread price from which to start the simulations. As the initial point for the simulation, the value of the first observed dark spread price is taken, i.e. 17.83 €/MWh. Summary statistics for the simulated dark spread prices together with statistics for the observed dark spread prices are given in Table 3.3. Also, Figure 3.3 gives a histogram of simulated and observed prices.

Comparing the statistics of the simulated prices to those of the observed prices, it appears that ABM does a reasonably good job of capturing the mean and median of the observed data. With regard to all other statistics (minimum, maximum, standard error, skewness and kurtosis) the simula-

[6]To obtain annual values of the drift parameter, the weekly drift value has been multiplied by 52. To convert weekly volatility to annual volatility, it has been multiplied by $\sqrt{52}$ (Hull (2005)).

tion based on ABM yields values which are substantially different from the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Simulated | -819.66 | 18.64 | 19.48 | 830.51 | 142.37 | -0.01 | 4.03 |
| Observed | 1.09 | 20.27 | 24.01 | 121.39 | 14.07 | 1.93 | 9.86 |

TABLE 3.3: Summary statistics for observed and 10,000 simulated price trajectories for ABM process

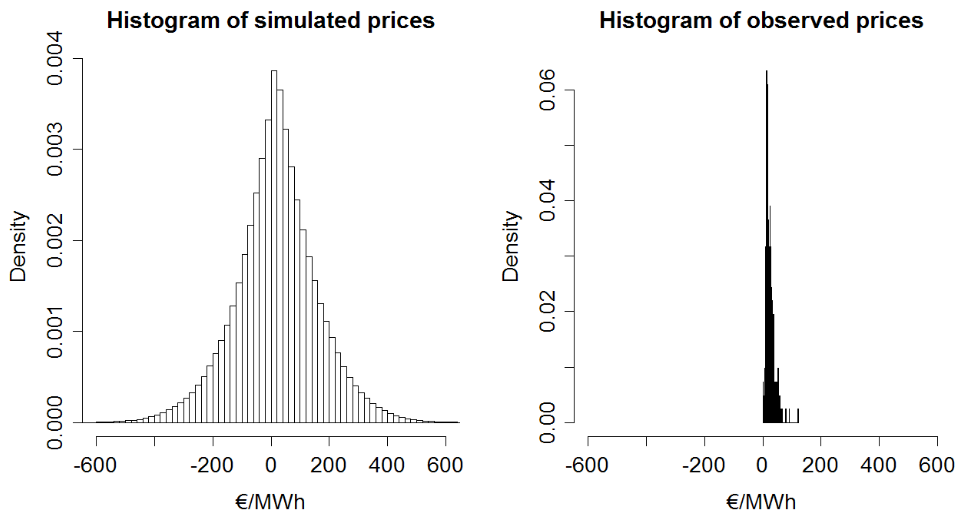



FIGURE 3.3: Histograms of observed and simulated prices for ABM

Nevertheless, the performance of ABM is even worse than it looks at first glance. The reason why ABM captures relatively well the mean and median of the observed prices is due to the selection of a starting value for the simulation. The starting value was 17.83 €/MWh which is relatively close to the mean of the observed data (24.01 €/MWh). Because the drift term in ABM is relatively small, one can expect that for half the time ABM will generate values greater than the starting value, and for the other half of the time, the generated values will be lower than the starting value. Thus, on average, the mean of the simulated prices should be around the value used to start the simulation (17.83 €/MWh), or a little above it due to the positive drift. Therefore, if the price at which the simulations are started is altered from 17.83 €/MWh to, for instance, 0 €/MWh then the values of the descriptive statistics change and are given by Table 3.4. Now, it appears that ABM does not capture any of the distributional properties of the observed prices.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Simulated | -837.50 | 0.81 | 1.64 | 812.67 | 142.37 | -0.01 | 4.03 |
| Observed | 1.09 | 20.36 | 24.03 | 121.39 | 14.08 | 1.93 | 9.83 |

TABLE 3.4: Summary statistics for observed and 10,000 simulated price trajectories for ABM

Performing Monte Carlo (MC) simulations shows that ABM is not capable of matching the distributional properties of dark spread prices. Furthermore, the performance of ABM is greatly influenced by the starting value used in the simulations. As an illustration, Figure 3.4 shows 4 random price realizations of ABM: none of them resemble the observed prices.

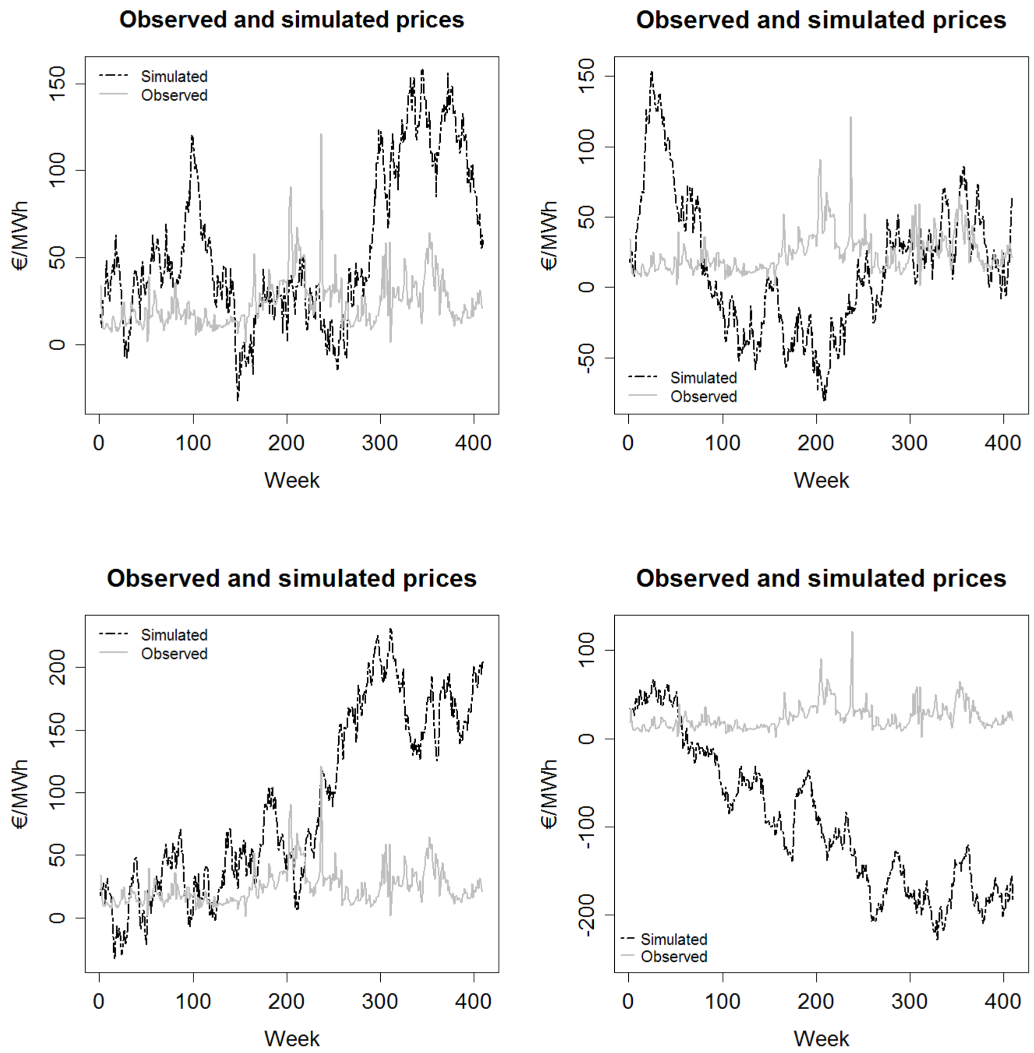



FIGURE 3.4: Simulated paths of dark spread prices following ABM

As a further measure of goodness of fit, the in sample RMSE is used: the RMSE is calculated for each of the 10,000 sample paths (see Appendix A.2).

Table 3.5 reports the mean value of the RMSE together with the standard error of the estimate.

| | mean | se. |
|---|---|---|
| RMSE | 126.07 | 67.68 |

TABLE 3.5: RMSE for in sample forecasts based on ABM process

## 3.5 Ornstein-Uhlenbeck model

In Section 3.4, it was shown that ABM cannot properly describe the behavior of dark spread prices. Economic logic suggests that mean reverting processes should be best suited to model the development of dark spread. The basic idea behind a mean reverting process is that prices cannot stay away from some long-run level for too long: soon after moving away, prices are pulled back to their long-run level.

The first mean reverting process that is analyzed is the Ornstein-Uhlenbeck (OU) process. The process is defined by the following stochastic differential equation (SDE):

$$dp_t = k(\mu - p_t)dt + \sigma dz_t \tag{3.6}$$

In Equation 3.6, $\mu$ represents the average, long-run price level: it can be thought of as a price level corresponding to the average cost of production. $k$ stands for the speed of reversion i.e. how quickly prices are pulled back when they move away from their long-run level. $\sigma$ is the volatility of price change, and $p_t$ and $dz_t$ represent the price level and the increment of a Wiener process.

### 3.5.1 Parameter estimation

To get the explicit solution to Equation 3.6, a new function $f(p,t) = pe^{kt}$ is defined (Iacus (2008)). Applying the Ito lemma to it, we get: $\frac{\partial f(p,t)}{\partial t} = pke^{kt}$, $\frac{\partial f(p,t)}{\partial p} = e^{kt}$ and $\frac{\partial^2 f(p,t)}{\partial p^2} = 0$. Furthermore:

$$df = f_t dt + f_p dp \tag{3.7a}$$

$$df = pke^{kt}dt + e^{kt}dp \tag{3.7b}$$

Inserting the value for $dp$ from equation 3.6 into equation 3.7b I obtain the following expression:

$$df = e^{kt} k\mu dt + e^{kt}\sigma dz \tag{3.8}$$

Integrating the equation 3.8 from $0$ to $t$ I get the expression for dark spread following the OU process:

$$p_t = p_0 e^{-kt} + \mu(1 - e^{-kt}) + \sigma \int_0^t e^{k(s-t)} dz_s \tag{3.9}$$

Thus, the mean and variance of the OU process are given by Equations 3.10a and 3.10b, respectively.

$$E[p_t] = p_0 e^{-kt} + \mu(1 - e^{-kt}) \tag{3.10a}$$

$$V[p_t] = \frac{\sigma^2}{2k}\left(1 - e^{-2kt}\right) \tag{3.10b}$$

To estimate the values of the parameters, a ML approach is adopted where the log-likelihood function is given by:

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = -(n-1)\cdot ln\sqrt{(Var(p))} - \sum_{i=2}^{n}\left[\frac{(p_i - E(p))^2}{2\cdot Var(p)}\right] \tag{3.11}$$

Inserting the appropriate expressions for the expected value and variance into Equation 3.11, I obtain the following expression for the log-likelihood function:

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = -(n-1)\cdot ln\zeta - \sum_{i=2}^{n}\left[\frac{(p_t - (p_{t-1}e^{-k} + \mu(1-e^{-k})))^2}{2\zeta^2}\right] \tag{3.12}$$

where $\zeta$ is used to denote the standard error of dark spread prices, i.e.

$$\zeta = \sqrt{Var[p_t]} = \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}}$$

The values of the estimated parameters are given in Table 3.6.

| | Weekly values |
|---|---|
| Mean reversion ($k$) | 0.28 (0.043) |
| Long-run price ($\mu$) | 24.05 (1.85) |
| Volatility ($\sigma$) | 10.62 (0.43) |
| Standard error ($\zeta$) | 9.27 (0.325) |

TABLE 3.6: Estimated weekly parameters for dark spread prices following OU process

To determine how well the OU process fits the data, 10,000 trajectories are first simulated with the following expression:

$$p_t = p_0 e^{-k\Delta t} + \mu(1 - e^{-k\Delta t}) + \sigma\sqrt{\frac{1 - e^{-2k\Delta t}}{2k}} \cdot N(0,1) \tag{3.13}$$

The values of summary statistics for simulated and observed dark spread prices are given in Table 3.7, while the corresponding histograms are given in Figure 3.5. Table 3.7 shows that the first and second moment of the simulated prices are very close to the values for observed prices. Due to the fact that the OU process admits negative values which are not present in the observed prices, the simulated and observed prices differ in terms of skewness. The observed prices exhibit positive skewness which is evident in the histogram of prices shown in Figure 3.5. Furthermore, the OU process is not able to generate as high a kurtosis as is present in the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Simulated | -44.40 | 23.95 | 23.99 | 102.42 | 14.05 | 0.01 | 3.01 |
| Observed | 1.09 | 20.27 | 24.01 | 121.39 | 14.07 | 1.93 | 9.86 |

TABLE 3.7: Summary statistics for observed and 10,000 simulated price trajectories for OU process

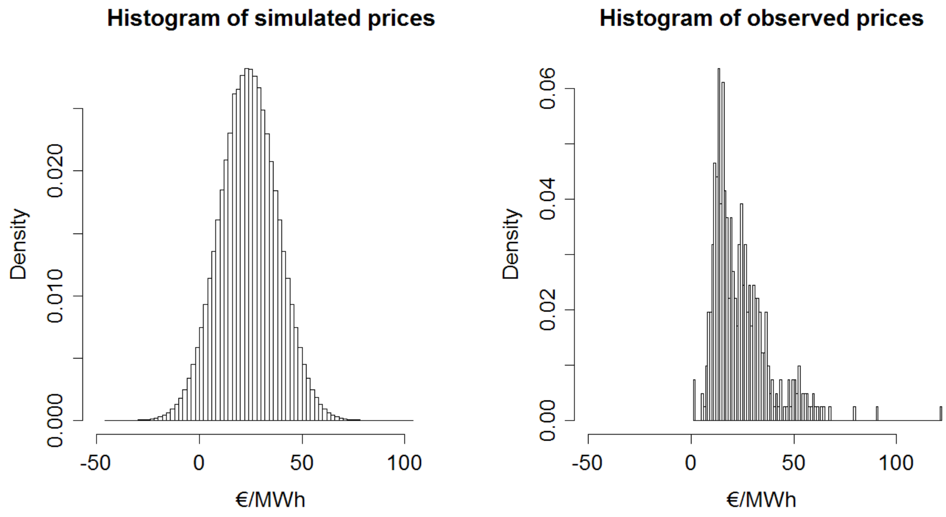



FIGURE 3.5: Histograms of observed and simulated prices for OU process

A sample trajectory of the OU process is given in Figure 3.6. Taking into consideration the summary statistics of Table 3.7 and comparing them with those of ABM, it is apparent that the OU process is much better at describing the evolution of dark spread prices.

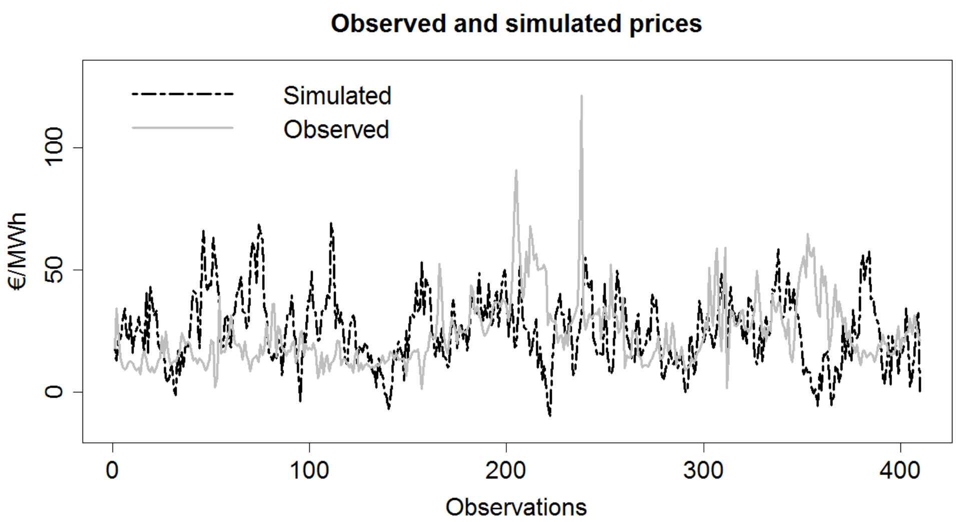



FIGURE 3.6: A simulated path of dark spread prices following OU process

For a second test of goodness of fit, the RMSE is computed and given in Table 3.8. Comparing it to ABM, both the mean and standard error are significantly improved.

| | mean | se. |
|---|---|---|
| RMSE | 19.82 | 1.22 |

TABLE 3.8: RMSE for in sample forecasts based on OU process

## 3.6 Cox-Ingersoll-Ross model

Cox et al. (1985) introduced a model to describe the evolution of interest rates which can also be used to model dark spread prices (henceforth referred to as CIR model). The SDE governing the process is given by the following expression:

$$dp_t = k(\mu - p_t)dt + \sigma\sqrt{p_t}dz_t \tag{3.14}$$

The main difference between the CIR and OU processes lies in the standard error of the process which depends on the square root of price: this feature makes it very interesting for modeling dark spread prices. Another difference between the two processes is in the distribution of prices. Unlike the OU process where prices are normally distributed, Cox et al. (1985) show that the probability density function of the value of dark spread at time $s$, given the initial value at $t$, is given by the following expression:

$$f(p_s|p_t) = ce^{-u-v}\left(\frac{v}{u}\right)^{q/2} I_q(2(uv)^{1/2}) \tag{3.15}$$

where:

$$c = \frac{2k}{\sigma^2(1 - e^{-k(s-t)})} \tag{3.16a}$$

$$u = cp_t e^{-k(s-t)} \tag{3.16b}$$

$$v = cp_s \tag{3.16c}$$

$$q = \frac{2k\mu}{\sigma^2} - 1 \tag{3.16d}$$

$I_q$ is the modified Bessel function of the first kind of order q (3.16e)

The expected value and variance of a variable $p$ following the CIR process at time $s$ given initial time $t = 0$ are given by the following two equations.

$$E[p_s|p_t] = p_t e^{-ks} + \mu(1 - e^{-ks}) \tag{3.17a}$$

$$Var(p_s|p_t) = p_t\frac{\sigma^2}{k}(e^{-ks} - e^{-2ks}) + \mu\frac{\sigma^2}{2k}(1 - e^{-ks})^2 \tag{3.17b}$$

Looking at the above equations, we see that the expected value of the CIR process is the same as for the OU process, while the variance of the process is non-constant and depends on the price level $p$. A useful property is that a variable following the CIR process cannot become negative. More formally, if $2k\mu/\sigma^2 \geq 1$ the process never reaches zero; if $0 < 2k\mu/\sigma^2 < 1$, zero serves as a reflecting barrier; and if $k\mu = 0$ zero is an absorbing barrier and it is reached surely in finite time (Overbeck and Ryden (1997)).

### 3.6.1 Parameter estimation

To estimate the parameters, the ML approach is adopted. The log-likelihood function is given by:

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = (n-1)lnc + \sum_{i=2}^{n} \left\{ -u - v + 0.5qln\left(\frac{v}{u}\right) + ln\{I_q(2\sqrt{uv})\} \right\} \quad (3.18)$$

Using the above log-likelihood function, the following parameters are estimated:

| | Weekly values |
|---|---|
| Mean reversion ($k$) | 0.287 (0.044) |
| Long-run price ($\mu$) | 24.05 (1.644) |
| Volatility ($\sigma$) | 1.94 (0.079) |

TABLE 3.9: Estimated weekly parameters for dark spread prices following CIR process

In order to test the goodness of fit of the model, the price trajectories are simulated with the estimated parameters. To simulate the trajectories, it is not possible to use Euler discretization[7] such as in the case of ABM or OU processes because of the possibility of obtaining negative values for $p$. Formally, the following expression will not guarantee positive values of $p$, in which case the value under the square root will be undefined:

$$p_{t+1} = p_t + k(\mu - p_t)dt + \sigma\sqrt{p_t}dz_{t+1} \quad (3.19)$$

Nevertheless, the transition density for $p_s$ is known (Glasserman (2003)) and it can be used to simulate the price trajectories following the CIR process.

[7]Euler discretization (also called Euler approximation, Euler method or Euler scheme) refers to approximation of continuous time expressions with discrete time equivalents. This scheme is commonly used in simulating trajectories of stochastic processes (e.g. Wilmott (1994) page 1264, Glasserman (2003) page 7).

The transition density is given by the following expression:

$$p_s = \frac{\sigma^2(1-e^{-k(s-t)})}{4k}\chi^2_d\left(\frac{4ke^{-k(s-t)}}{\sigma^2(1-e^{-k(s-t)})}p_t\right) \qquad s > t \tag{3.20}$$

where

$$d = \frac{4\mu k}{\sigma^2} \tag{3.21}$$

Equation 3.20 says that given $p_t$, $p_s$ is distributed as $\sigma^2(1-e^{-k(t-s)})/4k$ times a non central chi-square random variable with $d$ degrees of freedom and non centrality parameter $\lambda$ given by:

$$\lambda = \frac{4ke^{-k(t-s)}}{\sigma^2(1-e^{-k(t-s)})}p_s \tag{3.22}$$

Using Equation 3.20, 10,000 simulation are run and the summary statistics reported in Table 3.10 are obtained.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Simulated | 0.19 | 21.83 | 23.99 | 158.06 | 12.51 | 1.05 | 4.66 |
| Observed | 1.09 | 20.27 | 24.01 | 121.39 | 14.07 | 1.93 | 9.86 |

TABLE 3.10: Summary statistics for observed and simulated prices for CIR process

Histograms of observed versus simulated prices are given in Figure 3.7.

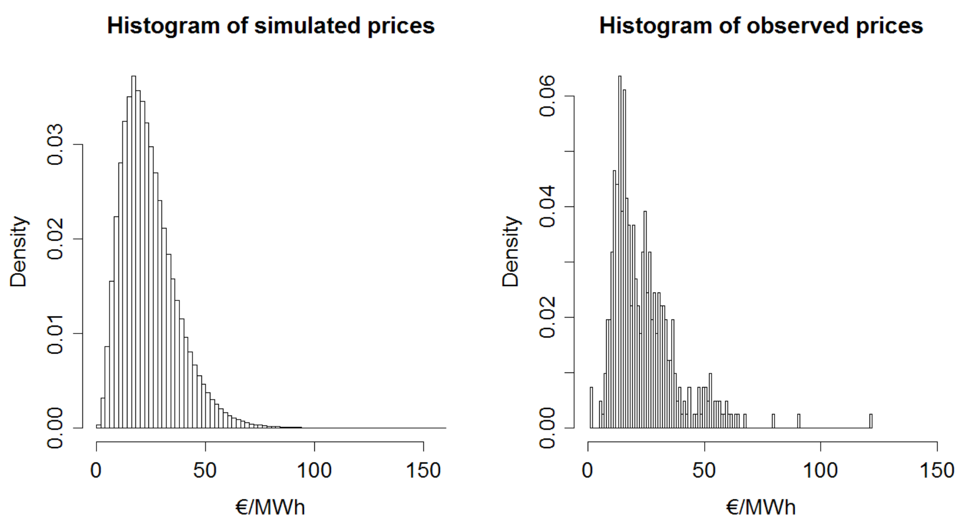



FIGURE 3.7: Histograms of observed and simulated prices for CIR process

Table 3.10 shows that the simulated data come very close to the observed data in terms of mean and standard error, as was the case with the OU process. Unlike the OU process, CIR comes significantly closer to the observed data in terms of skewness. A most likely reason for this improvement lies in the fact that CIR does not admit negative values, unlike OU. In terms of kurtosis, CIR also gives much better results compared with the OU process, i.e. it is capable of generating larger price 'spikes'. Nevertheless, CIR is not able to come very close to the kurtosis in the observed prices.

As an illustration, a plot of the simulated trajectory versus the observed prices is given in Figure 3.8: one can see that a random CIR trajectory resembles the observed one better than achieved with ABM or OU process.

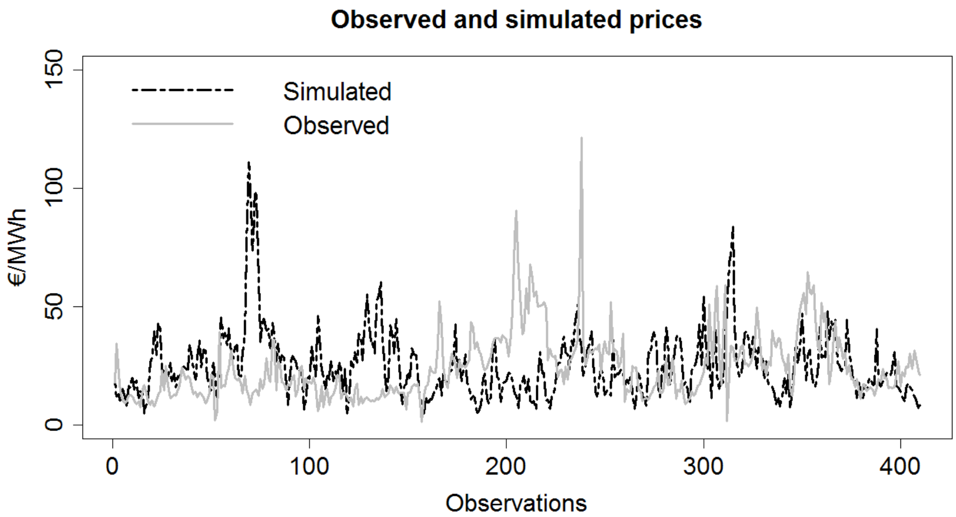



FIGURE 3.8: A simulated path of dark spread prices following the CIR process

Furthermore, the goodness of fit is tested using the RMSE calculated from in sample forecast: the values are reported in Table 3.11.

| | mean | se. |
|---|---|---|
| RMSE | 18.73 | 1.27 |

TABLE 3.11: RMSE for in sample forecasts based on CIR process

## 3.7 Schwartz one-factor model

Schwartz (1997) introduced a one-factor mean reverting model for valuing commodities which is given by:

$$dp_t = k(a - lnp_t)p_t dt + \sigma p_t dz_t \tag{3.23}$$

The symbols have the same meaning as for the CIR and OU processes except that instead of $\mu$, $a$ is used. Advantages of this model are a non-constant variance and the generation of strictly positive values, qualities shared by the CIR process as well. As a first step in the estimation process, the model is transformed to a logarithmic scale by defining $x = lnp$ and by using the Ito lemma the following expression is obtained (where $m = a - \frac{\sigma^2}{2k}$):

$$dx_t = k(m - x_t)dt + \sigma dz_t \tag{3.24}$$

To estimate the parameters, the log of the price is used, a plot of which is given in Figure 3.9.

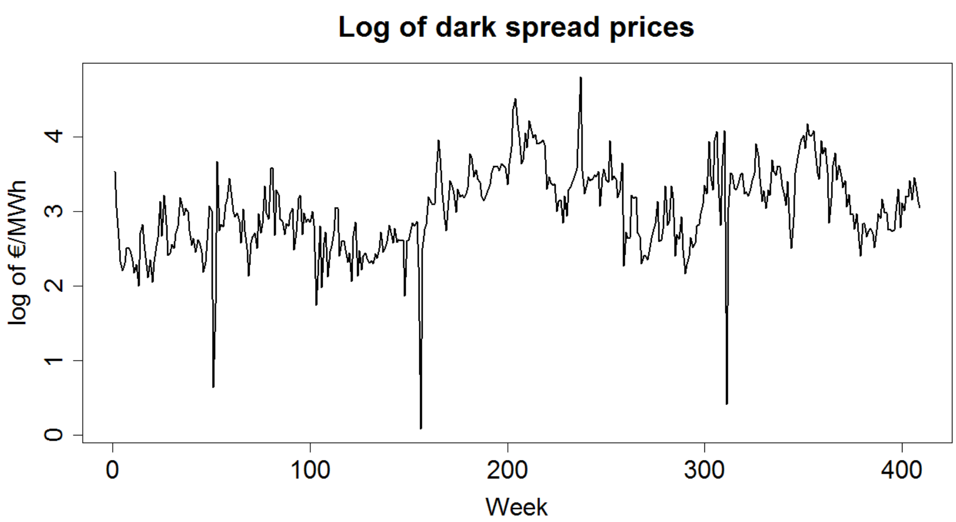



FIGURE 3.9: Logarithm of dark spread

The logarithm of dark spread given in Equation 3.24 is normally distributed[8]. The expected value and variance of $x_t$ are given by the following two expressions (to derive the expected value and variance, I use the same approach as in the case of the OU process):

$$E[x_t] = x_0 e^{-kt} + m(1 - e^{-kt}) \tag{3.25a}$$
$$V[x_t] = \frac{\sigma^2}{2k}\left(1 - e^{-2kt}\right) \tag{3.25b}$$

[8] Defining x=ln(p) and using Ito lemma it can be shown that logarithm of price which in level form follows Schwartz process, follows OU process which is normally distributed: see Schwartz (1997).

### 3.7.1 Parameter estimation

To estimate the parameters of the log of dark spread, the ML approach is adopted. The log-likelihood function is given by the following expression:

$$lnL(\boldsymbol{\theta}|\boldsymbol{x}) = -(n-1)\cdot ln\sqrt{(Var(x))} - \sum_{i=2}^{n}\left[\frac{(x_i - E(x))^2}{2\cdot Var(x)}\right] \tag{3.26}$$

Upon inserting values for the expected value and standard error of the log of dark spread from Equations 3.25a and 3.25b, the log-likelihood function becomes:

$$lnL(\boldsymbol{\theta}|\boldsymbol{x}) = -(n-1)\cdot ln\zeta - \sum_{i=2}^{n}\left[\frac{(x_t - (x_{t-1}e^{-kt} + m(1-e^{-kt})))^2}{2\zeta^2}\right] \tag{3.27}$$

In Equation 3.27, the following short-hand notation is used:

$$\zeta = \sqrt{Var[x_t]} = \sigma\sqrt{\frac{1-e^{-2kt}}{2k}} \tag{3.28}$$

The values of the estimated parameters are given in Table 3.12.

| | Weekly values |
|---|---|
| Mean reversion ($k$) | 0.367 (0.051) |
| Long-run price ($m$) | 3.029 (0.0656) |
| Volatility ($\sigma$) | 0.484 (0.02) |
| Standard error ($\zeta$) | 0.407 (0.0143) |

TABLE 3.12: Weekly and annual parameters for log of dark spread

To asses the goodness of fit, the expression from Equation 3.29 is used to simulate 10,000 trajectories of the logarithm of prices following the Schwartz model. Once the simulation has been performed, the exponential of the simulated log prices is taken to obtain the level prices following the Schwartz one-factor model. Summary statistics are given in Table 3.13 and histograms of the observed and simulated prices are shown in Figure 3.10.

$$x_t = x_0e^{-k\Delta t} + (a - \frac{\sigma^2}{2k})(1-e^{-k\Delta t}) + \sigma\sqrt{\frac{1-e^{-2k\Delta t}}{2k}}\cdot N(0,1) \tag{3.29}$$

What can be noticed is that the Schwartz model comes close to the observed data in terms of the first three moments of the distribution based on 10,000 simulations. On the other hand, unlike the CIR process, the Schwartz process overestimates the kurtosis.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Simulated | 1.26 | 20.61 | 24.21 | 465.07 | 14.86 | 2.10 | 11.97 |
| Observed | 1.09 | 20.27 | 24.01 | 121.39 | 14.07 | 1.93 | 9.86 |

TABLE 3.13: Summary statistics for observed and simulated prices following Schwartz one-factor model

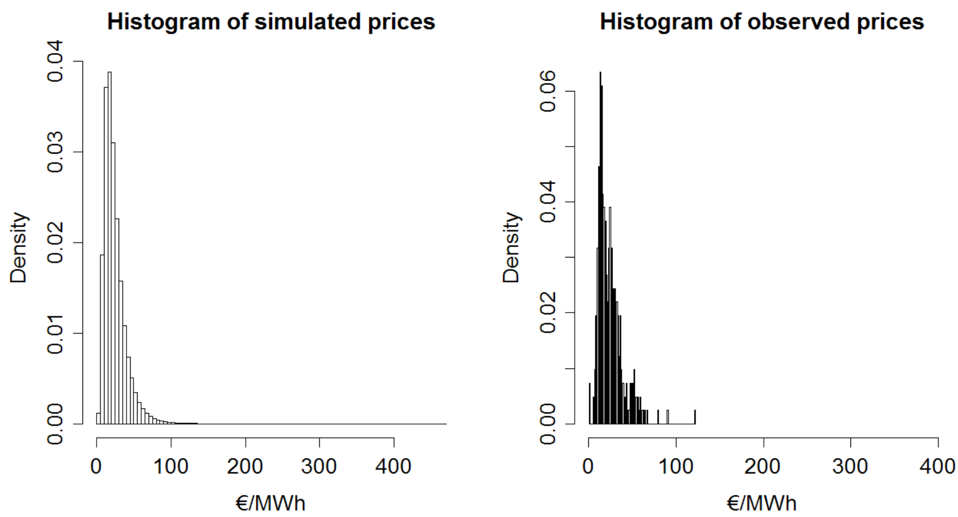



FIGURE 3.10: Histograms of observed and simulated prices for Schwartz process

As an illustration, Figure 3.11 shows a sample trajectory of the Schwartz one-factor model.

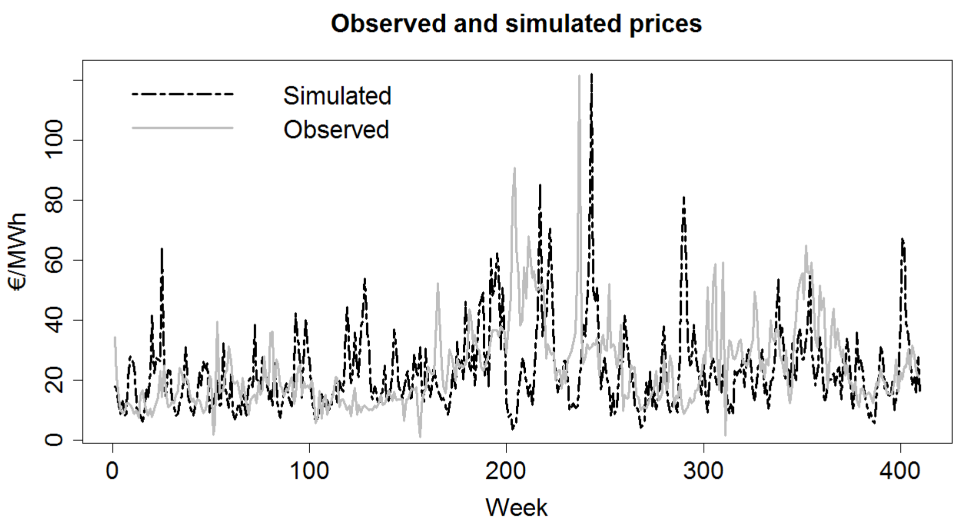



FIGURE 3.11: A simulated path of dark spread prices following Schwartz process

Next, to further asses the goodness of fit, sample forecasts are performed and the RMSE is calculated. The value of RMSE is given in Table 3.14.

| | mean | se. |
|---|---|---|
| RMSE | 20.38 | 1.69 |

TABLE 3.14: RMSE for in sample forecasts based on Schwartz model

## 3.8 Model selection

Having estimated the parameters of the stochastic processes, I choose the process which is most appropriate for modeling dark spread prices. For this, two criteria are used: distributional properties and RMSE.

### 3.8.1 Comparison based on distributional properties

Table 3.15 gives a summary of the distributional statistics of the selected processes. The purpose is to identify the stochastic process that comes closest to the observed data in terms of distributional statistics. Of the four stochastic processes, ABM performs the worst: it only manages to capture the first moment of the observed prices. As a consequence of the fact that it is an unbounded process, the maximum and minimum values attained by it are far from the observed data. Furthermore, it performs rather poorly in terms of the second moment as well. On the other hand, it does not perform so badly in terms of skewness and kurtosis relative to the other processes, but this was expected: skewness and kurtosis are what would be expected from a normally distributed random variable.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 1.09 | 20.27 | 24.01 | 121.39 | 14.07 | 1.93 | 9.86 |
| ABM | -819.66 | 18.64 | 19.48 | 830.51 | 142.37 | -0.01 | 4.03 |
| OU | -44.40 | 23.95 | 23.99 | 102.42 | 14.05 | 0.01 | 3.01 |
| CIR | 0.19 | 21.83 | 23.99 | 158.06 | 12.51 | 1.05 | 4.66 |
| Schwartz | 1.26 | 20.61 | 24.21 | 465.07 | 14.86 | 2.10 | 11.97 |

TABLE 3.15: Summary statistics for observed and simulated prices

The three mean reverting processes represent significant improvement over ABM in many aspects. In terms of minimum value, OU performs the worst as it allows negative values. All three processes generate mean values which are very close to the observed data. With regard to the standard error, the OU and Schwartz one-factor processes come very close to the observed data, while CIR performs a little less satisfactory. Finally, in terms of higher moments, OU performs the worst. CIR on the one hand underestimates the observed values while the Schwartz model overestimates them.

To facilitate better comparison, Table 3.16 gives the absolute difference between each stochastic process and observed data for all distributional statistics. Absolute values are computed according to the following expression:

$$|\bar{y}_{p,i} - y_i|$$

$\bar{y}_{p,i}$ represents the value of a statistic $i$ (minimum and maximum value, median, mean, standard error, skewness and kurtosis) for stochastic process $p$ (ABM, OU, CIR and Schwartz) and $y_i$ represents the value of statistic $i$ for the observed prices.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| ABM | 820.75 | 1.63 | 4.53 | 709.12 | 128.3 | 1.94 | 5.83 |
| OU | 45.49 | 3.68 | 0.02 | 18.97 | 0.02 | 1.92 | 6.85 |
| CIR | 0.90 | 1.56 | 0.02 | 36.67 | 1.56 | 0.88 | 5.20 |
| Schwartz | 0.17 | 0.34 | 0.2 | 343.68 | 0.79 | 0.17 | 2.11 |

TABLE 3.16: Absolute difference between simulated and observed data

Using Table 3.16, Table 3.17 shows the ordering of the models according to each statistic, i.e. how closely each model comes to the observed data in terms of absolute difference. The Schwartz model comes out the best in most of the statistics (1st place) and is followed closely by the CIR and OU processes. The ABM process on the other hand fares the worst: it is the worst performing process in terms of all statistics except for median. The last column represents a summary of row values: the lower the value, the better the process relative to others.

| | min. | median | mean | max. | se. | skew. | kurtosis | overall |
|---|---|---|---|---|---|---|---|---|
| ABM | 4 | 3 | 4 | 4 | 4 | 4 | 3 | 26 |
| OU | 3 | 4 | 1 | 1 | 1 | 3 | 4 | 17 |
| CIR | 2 | 2 | 1 | 2 | 3 | 2 | 2 | 14 |
| Schwartz | 1 | 1 | 3 | 3 | 2 | 1 | 1 | 12 |

TABLE 3.17: Ordering of stochastic processes according to absolute differences

Based on the distributional statistics, one can conclude that the ABM process is not an appropriate choice for modeling dark spread prices. If equal weight is given to all statistics, the Schwartz model turns out to be the most appropriate process, followed by the CIR and OU processes.

### 3.8.2 Comparison based on root mean squared error

Using distributional properties is an informal method of choosing among different processes. While it is useful for identifying processes that stand out (such as ABM) it is not very useful in choosing among competing processes. Therefore, as a second measure of goodness of fit, I use the RMSE for in sample forecasts based on 10,000 simulations.

RMSE shows again that ABM is the worst performing of all four processes, which can be seen in the highest mean and standard errors of RMSE. For mean reverting processes, the result is different from that obtained using distributional properties. According to the RMSE, the CIR process has the lowest mean value of RMSE among all three mean reverting processes. It is followed by the OU process which has slightly higher mean RMSE value but it has a lower standard error for the estimate. Finally, the Schwartz process comes last, with the highest mean and standard error of RMSE among the mean reverting processes. Summary statistics for mean and standard error of RMSE are given in Table 3.18.

| | mean | se. |
|---|---|---|
| ABM | 126.07 | 67.68 |
| OU | 19.82 | 1.22 |
| CIR | 18.73 | 1.27 |
| Schwartz | 20.38 | 1.69 |

TABLE 3.18: RMSE for all stochastic processes

The next question addressed is whether the differences in RMSE among stochastic processes are statistically significant. Diebold and Mariano (1995) developed a test for comparing the prediction accuracy of two different models. Denote by $y_{t+1}$ the actual time series and by $\hat{y}_{A,t+1|t}$ and $\hat{y}_{B,t+1|t}$ two different forecasts for the actual series. For $t = 1...T$, the error associated with each forecast is given by:

$$e_{A,t+1|t} = \hat{y}_{A,t+1|t} - y_{t+1} \tag{3.30a}$$

$$e_{B,t+1|t} = \hat{y}_{B,t+1|t} - y_{t+1} \tag{3.30b}$$

The goal is to determine the time $t + 1$ loss associated with each forecast. To calculate the loss, a loss function $g$ is defined which is a direct function of the forecast error, i.e. $g = g(e_{i,t+1|t})$ where $i = A, B$. A quadratic loss function is chosen, i.e. $g = g(e_{i,t+1|t}) = e^2_{i,t+1|t}$. Then a null hypothesis can be stated of equal forecast accuracy for both models, i.e. that the loss associated with each forecast is the same against the alternative hypothesis that the forecasts differ:

$$H_0 : \; e^2_{A,t+1|t} = e^2_{B,t+1|t} \tag{3.31a}$$
$$H_1 : \; e^2_{A,t+1|t} \neq e^2_{B,t+1|t} \tag{3.31b}$$

Defining $d_t = e^2_{A,t+1|t} - e^2_{B,t+1|t}$, the null and alternative hypotheses can be stated as:

$$H_0 : \; d_t = 0 \tag{3.32a}$$
$$H_1 : \; d_t \neq 0 \tag{3.32b}$$

The Diebold-Mariano (DM) test statistic is given by the following expression (Harvey et al. (1997)):

$$DM = \frac{\bar{d}}{\sqrt{Var(d)}} \tag{3.33}$$

where:

$$\bar{d} = T^{-1} \sum_{i=1}^{T} d_i \tag{3.34a}$$
$$Var(d) = T^{-1} \left[ \gamma_0 + 2 \sum_{k=1}^{h-1} \gamma_k \right] \tag{3.34b}$$

And $\gamma_k$ is $k^{th}$ auto-covariance that can be estimated by:

$$\gamma_k = T^{-1} \sum_{t=k+1}^{T} (d_t - \bar{d})(d_{t-k} - \bar{d}) \tag{3.35}$$

The DM test statistic is a pairwise test, thus results are reported for various combinations of stochastic processes. As each simulation of the stochastic process is different from the previous one, the results of the DM test are reported for all 10,000 simulations. Detailed explanation of calculation of DM statistic is given in Appendix A.3.

#### 3.8.2.1 ABM versus OU process

Using the DM statistic, I test the null hypothesis of equal forecast accuracy between ABM and OU processes against the alternative of ABM having a worse forecast accuracy i.e. higher forecast error:

$$H_0 : \ e_{ABM,t+1|t} = e_{OU,t+1|t} \tag{3.36a}$$

$$H_1 : \ e_{ABM,t+1|t} > e_{OU,t+1|t} \tag{3.36b}$$

To test the null hypothesis I compare the values of DM statistic with critical values from standard normal distribution. The critical value for one sided test at $5\%$ significance level is $1.65$. Consequently, if value of DM statistic is greater than $1.65$ I reject the null in favor of the alternative hypothesis.

I calculate the value of DM statistic for all 10,000 sample paths, histogram of which is given in Figure 3.12. I find that for all the sample paths, the OU process has statistically lower forecast error. Therefore, I conclude that the OU process is statistically superior to ABM in terms of forecasting accuracy.

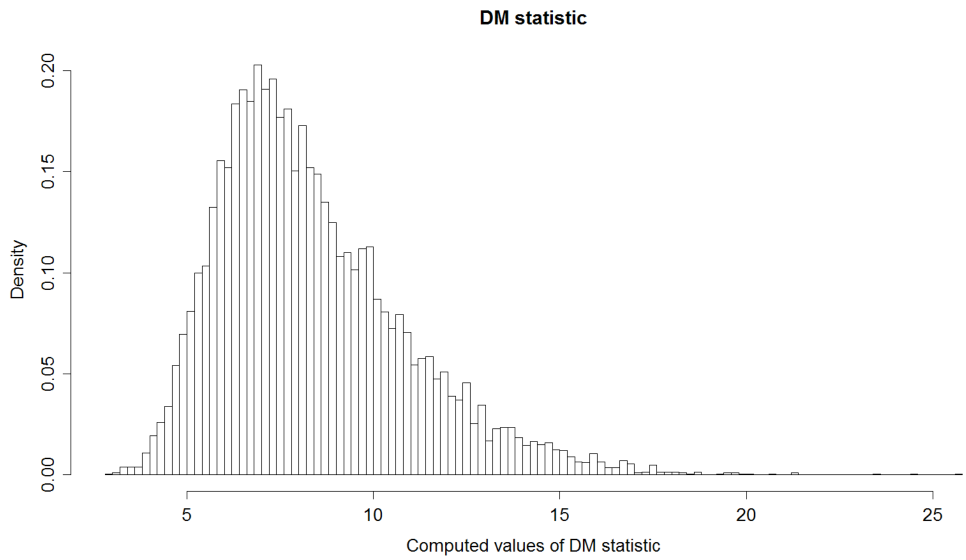



FIGURE 3.12: ABM versus OU - histogram of computed DM statistics

#### 3.8.2.2 ABM versus CIR process

The null hypothesis of equal forecast accuracy between the ABM and the CIR processes is tested against the alternative hypothesis that the CIR process has better forecast accuracy:

$$H_0 : \; e_{ABM,t+1|t} = e_{CIR,t+1|t} \tag{3.37a}$$

$$H_1 : \; e_{ABM,t+1|t} > e_{CIR,t+1|t} \tag{3.37b}$$

The result is the same as in the case of the OU process: the CIR process has statistically better forecasting accuracy in all of the sample paths. Figure 3.13 gives a histogram of computed DM statistics.

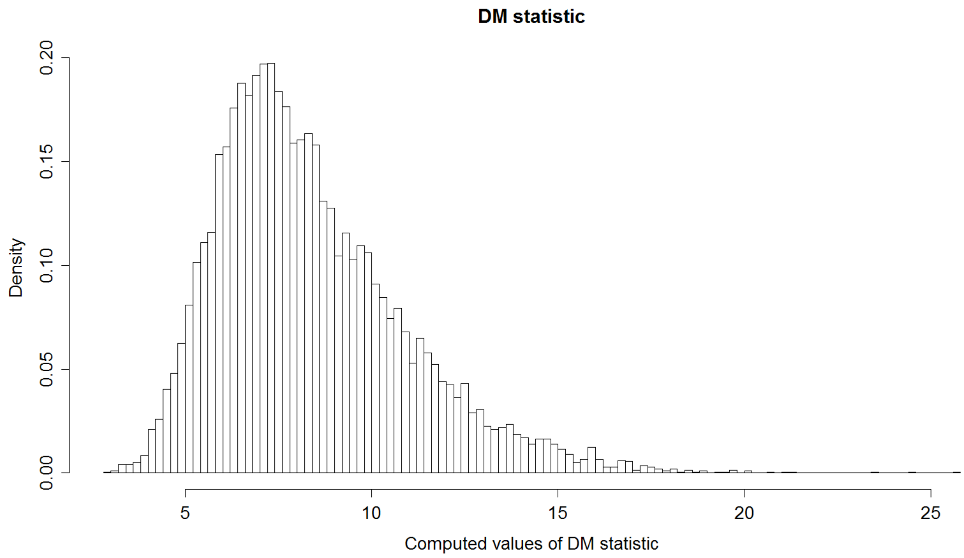



FIGURE 3.13: ABM versus CIR - histogram of computed DM statistics

### 3.8.2.3 ABM versus Schwartz process

Finally, ABM is tested against the Schwartz one-factor model using the same null and alternative hypotheses as in the previous cases:

$$H_0 : \; e_{ABM,t+1|t} = e_{Sch,t+1|t} \tag{3.38a}$$

$$H_1 : \; e_{ABM,t+1|t} > e_{Sch,t+1|t} \tag{3.38b}$$

The results are as expected. For all trajectories, the null hypothesis that the ABM and the Schwartz processes generate forecasts of equal accuracy is rejected in favor of the alternative hypothesis that the Schwartz process generates more accurate forecasts. The histogram of computed values of the DM statistic is given in Figure 3.14.

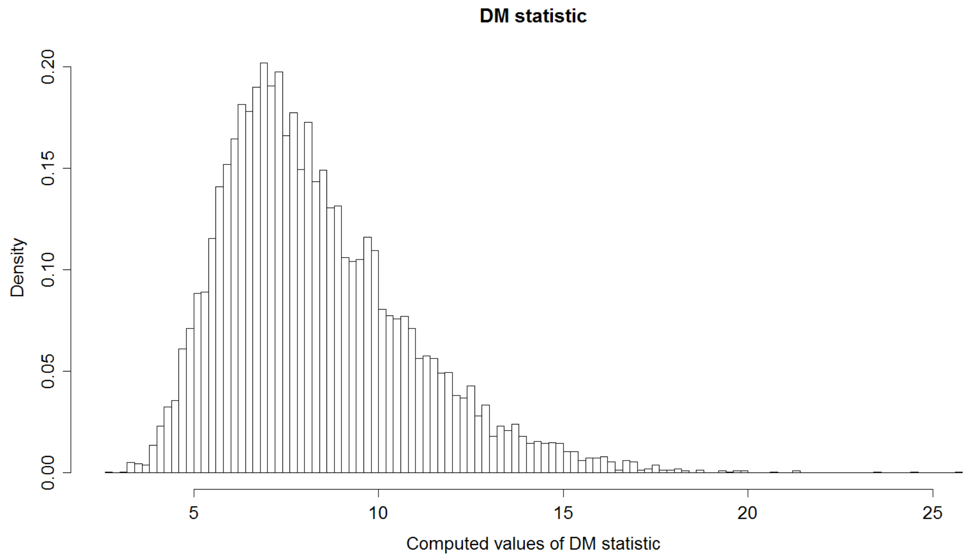



FIGURE 3.14: ABM versus Schwartz - histogram of computed DM statistics

#### 3.8.2.4 OU versus CIR process

Next, the mean reverting processes are tested against each other. First, I test whether the OU and CIR processes have the same forecast accuracy against the two-sided alternative that their forecasts differ. At $5\%$ significance level, the critical value is $1.96$ and the null hypothesis is rejected if the DM statistic is greater than $1.96$ in absolute value. For the data given, the null hypothesis is rejected in $20.07\%$ of all sample paths.

$$H_0 : e_{OU,t+1|t} = e_{CIR,t+1|t} \tag{3.39a}$$

$$H_1 : e_{OU,t+1|t} \neq e_{CIR,t+1|t} \tag{3.39b}$$

Given that a significant proportion of the OU and CIR forecasts differ, I test whether OU and CIR have the same forecast accuracy against the alternative that CIR has better forecast accuracy, i.e. lower forecast error. At $5\%$ significance level, the one-sided critical value is $1.65$, and the null hypothesis is rejected if the value of the DM test exceeds $1.65$. The null hypothesis is rejected in $26.72\%$ of all sample paths.

$$H_0 : e_{OU,t+1|t} = e_{CIR,t+1|t} \tag{3.40a}$$

$$H_1 : e_{OU,t+1|t} > e_{CIR,t+1|t} \tag{3.40b}$$

I also try a different alternative hypothesis where the forecast error of the CIR model exceeds the forecast error of the OU model. Here again, I use a one-sided alternative with $5\%$ significance level: the null hypothesis is rejected if the value of the DM statistic is smaller than $-1.65$. In this case, the null is rejected in only $2.41\%$ of all trajectories.

$$H_0:\ e_{OU,t+1|t} = e_{CIR,t+1|t} \tag{3.41a}$$
$$H_1:\ e_{OU,t+1|t} < e_{CIR,t+1|t} \tag{3.41b}$$

Therefore, I conclude that the CIR process has forecasts of better accuracy. A histogram of the computed values of the DM statistic is given in Figure 3.15.

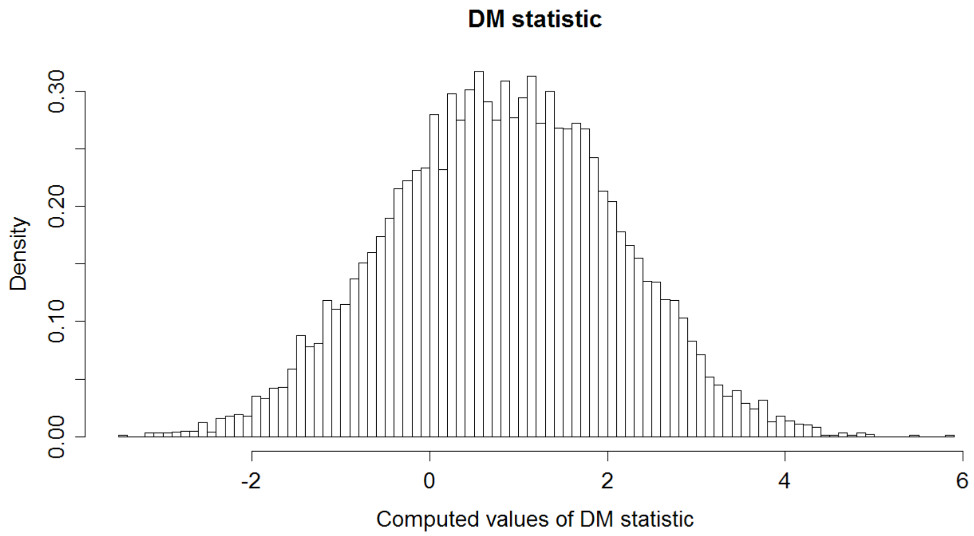



FIGURE 3.15: OU versus CIR - histogram of computed DM statistics

#### 3.8.2.5 OU versus Schwartz process

I test whether the OU and Schwartz models generate forecasts of equal accuracy against the alternative that their forecasts differ. At $5\%$ significance level, the null hypothesis is rejected in $9.3\%$ of all trajectories.

$$H_0:\ e_{OU,t+1|t} = e_{Sch,t+1|t} \tag{3.42a}$$
$$H_1:\ e_{OU,t+1|t} \neq e_{Sch,t+1|t} \tag{3.42b}$$

Following, at $5\%$ significance level I test a one-sided alternative that the OU process has lower forecast error than the Schwartz one-factor model. In this case I reject the null in favor of the alternative in $11.5\%$ of all sample paths.

$$H_0:\ e_{OU,t+1|t} = e_{Sch,t+1|t} \tag{3.43a}$$
$$H_1:\ e_{OU,t+1|t} < e_{Sch,t+1|t} \tag{3.43b}$$

Finally, I test whether Schwartz process has better forecasts than OU process. In $6\%$ of all sample paths, Schwartz one-factor model performs statistically better. Therefore, I conclude that OU and Schwartz model perform rather equally well, but the small advantage should be given to the OU model.

$$H_0 : e_{OU,t+1|t} = e_{Sch,t+1|t} \tag{3.44a}$$
$$H_1 : e_{OU,t+1|t} > e_{Sch,t+1|t} \tag{3.44b}$$

A histogram of computed values of the DM statistic is given in Figure 3.16.

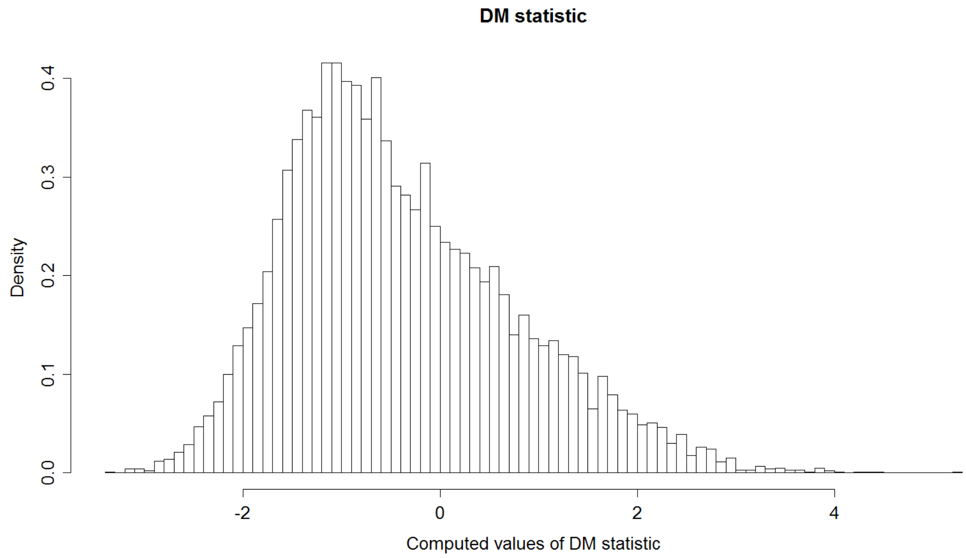



FIGURE 3.16: OU versus Schwartz - histogram of computed DM statistics

#### 3.8.2.6 CIR versus Schwartz process

First, I test the alternative hypothesis that the Schwartz and CIR models generate forecasts of different accuracies. For a two-sided alternative at $5\%$ significance level, the null hypothesis is rejected in favor of the alternative in $18.49\%$ of trajectories.

$$H_0 : e_{CIR,t+1|t} = e_{Sch,t+1|t} \tag{3.45a}$$
$$H_1 : e_{CIR,t+1|t} \neq e_{Sch,t+1|t} \tag{3.45b}$$

Next, I test whether the CIR model performs better than the Schwartz one-factor model. For a one-sided alternative and $5\%$ significance level, this is found to be the case in $26.95\%$ of all trajectories.

$$H_0 : \; e_{CIR,t+1|t} = e_{Sch,t+1|t} \tag{3.46a}$$
$$H_1 : \; e_{CIR,t+1|t} < e_{Sch,t+1|t} \tag{3.46b}$$

In the following, I test whether the CIR process performs worse than the Schwartz process, i.e. whether the CIR process has higher forecast error. Again, for a one-sided alternative at $5\%$ significance level, this is the case in $1.95\%$ of all trajectories.

$$H_0 : \; e_{CIR,t+1|t} = e_{Sch,t+1|t} \tag{3.47a}$$
$$H_1 : \; e_{CIR,t+1|t} > e_{Sch,t+1|t} \tag{3.47b}$$

I conclude that the CIR process performs better: the CIR model appears to be a better candidate for the modeling of dark spread prices. Figure 3.17 gives a histogram of computed DM statistics.

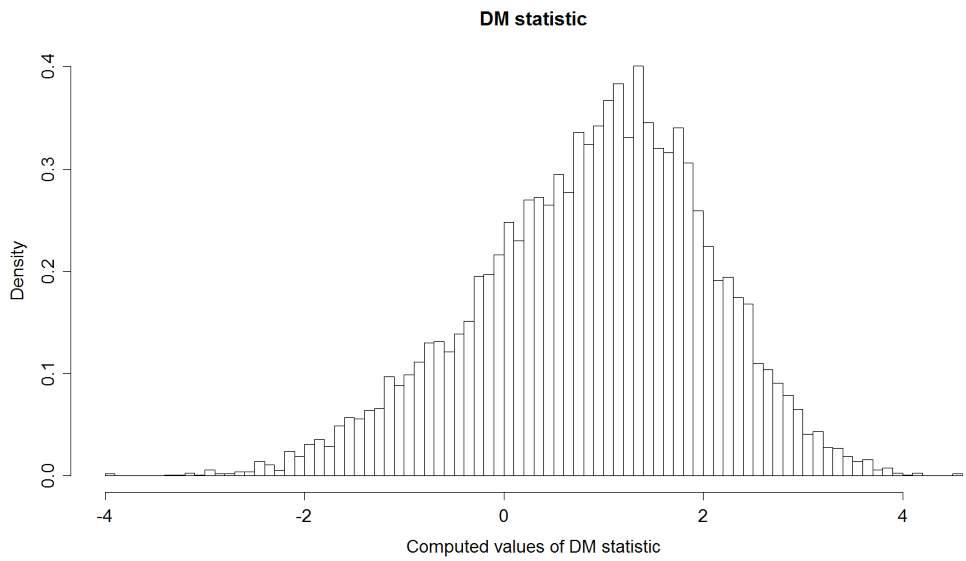



FIGURE 3.17: CIR versus Schwartz - histogram of computed DM statistics

## 3.9 Conclusion

The purpose of this chapter was to find an appropriate stochastic process to fit observed dark spread prices. Four different processes were used: ABM, OU, CIR and Schwartz one-factor model. After estimating the parameters of each process, 10,000 simulations were run. As a first measure of the goodness of fit, the distributional statistics were computed for each process and compared with the distributional statistics of the observed data. If the same weight is given to all statistics, the Schwartz one-factor process ranked the highest, followed by the CIR and OU processes, while ABM performed the worst.

As a second test of the goodness of fit, the RMSE was computed. As far as ABM is concerned, the results were the same as in the case of distributional statistics: it performed the worst. On the other hand, the CIR process performed the best followed closely by the OU process and then the Schwartz process.

Finally, the DM statistic was used to test whether the difference in forecast accuracy between different models was statistically significant. Again, ABM performed the worst: the null hypothesis of equal forecast accuracy between ABM and each individual mean reverting process was rejected in all of the 10,000 sample paths. When comparing the mean reverting processes among each other, the difference between them is not so significant. When comparing OU against the CIR process, the advantage should be given to the CIR process. Comparing the CIR and Schwartz processes, the advantage should again be given to the CIR process. Finally, when comparing the OU and Schwartz one-factor processes, one concludes that these two processes have forecasts of very similar accuracy. Therefore, based on the DM test, the CIR process seems to be statistically the most accurate in terms of in sample forecasts.

The conclusion of this chapter is that the CIR process is the most appropriate process to be used in dark spread modeling. The OU and Schwartz one-factor process perform equally well while the ABM process completely under-performs compared to the mean reverting processes.

Chapter 4

# One factor real options model

## 4.1 Introduction

In the previous chapter, I used four different stochastic processes to model the evolution of dark spread prices. Of the four models, the CIR process turned out to be the best. The Schwartz one-factor and OU processes performed comparably, while the ABM process performed the worst in terms of distributional properties as well as RMSE.

The purpose of this chapter is to asses how the choice of each of the four stochastic processes affects an investment decision within the real options (RO) framework. More particularly, do the more sophisticated models such as CIR or Schwartz one-factor model give significantly different results from less sophisticated models such as OU or ABM processes. The analysis is done assuming a perpetual investment option.

## 4.2 Literature review

Takizawa et al. (2001) analyze the investment in a nuclear power plant in Japan and calculate the threshold electricity price at which one should invest. They assume the value of the power plant depends on the price of electricity and uranium, both of which are modeled using geometric Brownian motion (GBM). They find that the optimal investment threshold calculated using the RO approach is higher than that calculated using a traditional capital budgeting approach such as a net present value (NPV) analysis. They also conclude that the optimal investment threshold is significantly affected by the volatility of state variables: a relation that cannot be observed in an NPV analysis.

Venetsanos et al. (2002) analyze the investment in a wind farm in the Greek electricity market with an emphasis on modularity and flexibility of wind power plants. The authors analyze an actual project using NPV which gives negative project value, suggesting the project should not be undertaken. Using RO analysis, the authors reach a different conclusion, namely that investment in the wind farm should be undertaken but in stages. Trying to make the valuation process as simple as possible so that it is accessible to practitioners, the authors assume that electricity prices follow GBM and thus use the classical Black-Scholes option valuation equation.

Gollier et al. (2005) investigate the impact of market liberalization on investments in large nuclear power plants. With the liberalization of electricity markets, electricity producers are faced with a plethora of risks, such as demand, price, fuel and regulatory risk. The authors ask whether this increase in overall riskiness of the business environment favors large and cost-effective investments or smaller and modular ones. Thus, the paper investigates the investment in a large 1200 MW, and a small, modular $4\times300$ MW nuclear power plant (in a non-liberalized market the choice would be to build the larger unit, because there is practically no uncertainty and the goal of the utilities is to minimize investment costs per unit of power installed). It is assumed that the electricity price follows GBM (parameters are also assumed). The authors conclude that, in a risky environment, investors should be more inclined to invest in smaller though more expensive units which offer more flexibility.

Kjaerland (2007) evaluates the investment in a hydro-power plant in Norway. Given that electricity is a non-storable commodity which implies that traditional no arbitrage arguments cannot be used, the author uses forward contracts to determine parameters for the evolution of electricity prices. As a simplification the author assumes that forward prices follow GBM: this allows for a closed form solution to the investment problem. As a final result, the author obtains the investment threshold in the form of a critical forward price. In a similar paper, Bockman et al. (2008) also evaluate the investment in a hydro-power plant in Norway. Unlike the previous paper, they model the technical characteristics of small hydro-power plants (sHPPs). The authors assume that a sHPP incurs some maintenance costs, which are divided into fixed maintenance costs (those that do not vary with production) and variable maintenance costs (those that vary with production). Fixed maintenance costs are lumped together with investment costs while variable maintenance costs are subtracted from the price of electricity, giving rise to a contribution margin $\theta$ (the variable maintenance costs are assumed to be constant per unit of electricity produced). It is assumed that the contribution margin follows GBM for which the parameters are obtained using data on forward contracts. The authors then present three potential projects: to value these projects, they apply the methodology presented in the paper.

Abadie and Chamorro (2008) investigate the choice of investment between a Natural Gas Combined Cycle (NGCC) plant (also known as a Combined Cycle Gas Turbine - CCGT) and an Integrated Gasification Combined Cycle plant (IGCC). While NGCC can only burn gas, IGCC can burn either gas or coal. As both of these technologies produce base load electricity, electricity price risk affects both of them in the same manner. Thus, the authors decide not to focus on electricity price uncertainty and assume constant electricity price. As for fuel prices, the authors assume that they follow an inhomogeneous mean reverting process. They derive the value of the option of investing in each of the two technologies, as well as the value of the option of investing in either of the two technologies.

## 4.3 Plant characteristics

I investigate the optimal investment decision in a coal-fired power plant undertaken by a merchant producer. Unlike state-owned utilities that can have different goals, a merchant producer has the single goal of profit maximization. I assume the plant owner is not faced with any emission allowance risk but only with the risk stemming from the uncertainty in coal and electricity prices: they are introduced into the model via dark spread prices. I also assume the power plant is of a base load type, thus I do not include any operational options (such as the option of suspending production in case prices fall too low).

The purpose of this chapter is not to create a detailed investment study but rather to assess the impact of uncertainty on investment decisions and to determine the difference in results obtained by traditional capital budgeting technique (NPV) and by RO analysis. Thus, I use generic technical assumptions for a coal-fired power plant. Plant characteristics are presented in Table 4.1.

| Parameter | Value |
|---|---|
| Project life (T) | 40 years |
| Capacity | 500 MW |
| Capacity factor | 85% |
| Annual production (q) | 3,723,000 MWh |
| Investment cost per kW | 1,500 € |
| Discount rate (r) | 10% |

TABLE 4.1: Investment parameters

Most of the technical assumptions were taken from a publication by the International Energy Agency (IEA) on the production costs of electricity (IEA (2005)). The report is based on interviews / questionnaires with actual owners of power plants around the world. In the report, it was assumed that the economic life of a coal-fired plant (amortization period) is 40 years, even though technically coal power plants can operate for much longer. The report also assumed a capacity factor of 85% [1]. This capacity factor translates into an annual production of 3,723,000 MWh of electricity.

Investment costs vary significantly and they are affected, among other factors, by the demand for coal plants, the cost of steel and concrete, and location (MIT (2007)). IEA (2005) cites an investment cost of approximately US$1,500 per kW, but this number seems too low. More recent numbers on the cost of coal-fired plants in the US, such as those presented in JSOnline (2008), estimate a cost per kW that exceeds US$3,000, which seems to be a bit on the high side. Therefore, I assume an investment cost of €1,500 per kW. Assuming 1.45 US$/€ exchange rate, this translates into US$2,175 per kW, which seems a reasonable number. In terms of operating and maintenance costs (O&M), I include only fuel costs. The reason is that other O&M costs are relatively small and, given that they can be predicted with reasonable accuracy, they affect investment decisions equally regardless of what approach is adopted. For the discount rate, I use a rate of 10%. Finally, as a simplifying assumption, I assume the power plant is build instantaneously.

## 4.4 NPV calculations

As a benchmark for investment analysis, I calculate the optimal investment threshold using NPV, an industry standard. It is well known that NPV

[1] The capacity factor is equal to the ratio of actual output during one year period divided by the output that would be produced if the plant operated at maximum (rated) capacity during the same period.

underestimates critical values at which one should invest because it does not take into account the opportunity cost of immediate action (option value), nor does it take into account the stochastic behavior of state variables. In other words, NPV does not account for uncertainty but rather assumes that investment is made in a perfectly certain environment.

The NPV value of the investment is obtained from the following expression:

$$NPV = \sum_{t=0}^{40} \frac{E(p_t \cdot q_t)}{(1+r)^t} - I \tag{4.1}$$

where $p_t$ stands for the dark spread price at year $t$ (price is assumed constant), $q_t$ is the annual production in MWh and $I$ is the investment cost. According to the NPV approach, an investment should be undertaken when NPV>0.

According to NPV calculations, it is optimal to invest in the coal power plant if one believes the price will on average be 20.6 €/MWh for each year during the life of the project. Figure 4.1 shows the value of NPV as a function of the underlying dark spread price.

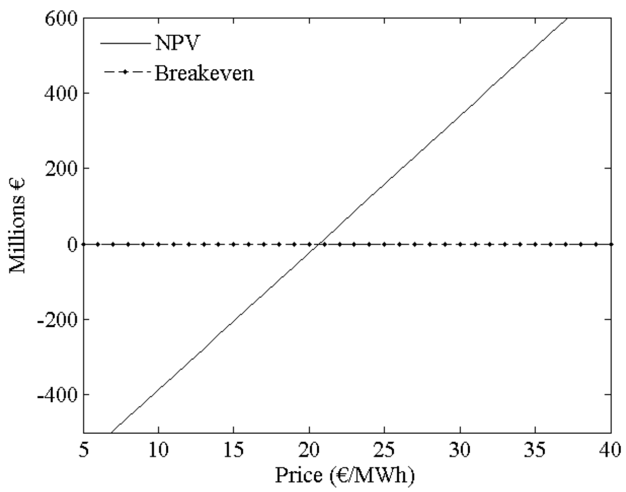



FIGURE 4.1: NPV - project value as a function of dark spread price

## 4.5 Real options valuation

In performing RO calculations, one can take two approaches: contingent claims (CC) or dynamic programming (DP). The CC approach rests on the assumption that one can use no-arbitrage principle to derive the optimal investment policy. The benefit of using the CC approach is the ability to use a risk-free interest rate. The DP approach on the other hand uses the Bellman equation to determine the optimal investment timing using an arbitrary discount rate. Because electricity cannot be stored which implies that traditional no-arbitrage principle breaks down, one cannot use the CC approach[2]. Therefore, I will use DP to solve for the optimal investment threshold.

### 4.5.1 Arithmetic Brownian motion

Because I assume no operational options once the power plant is in operation (such as option to temporarily shut down or mothball), the value of the project is a simple expected value of the discounted future net revenues (dark spread) given by the following expression (where the parameters have the same meaning as in Table 4.1):

$$V(p) = \int_0^T qE(p_t)e^{-rt}dt = q\int_0^T (at+p)e^{-rt}dt \tag{4.2a}$$

$$V(p) = \frac{q\left(a + pr - e^{-rT}(a + pr + arT)\right)}{r^2} \tag{4.2b}$$

In the previous equation, I used the definition for $E(p_t)$ given in Equation 3.2a. Following Dixit and Pindyck (1994) (Chapter 4), the option value ($f$) is given by:

$$E[df] = rf dt \tag{4.3a}$$

$$\frac{1}{2}f_{pp}\sigma^2 + af_p - rf = 0 \tag{4.3b}$$

Equation 4.3b is an ordinary differential equation (ODE) which has the following general solution:

$$f(p) = Ae^{\beta_1 p} + Be^{\beta_2 p} \tag{4.4}$$

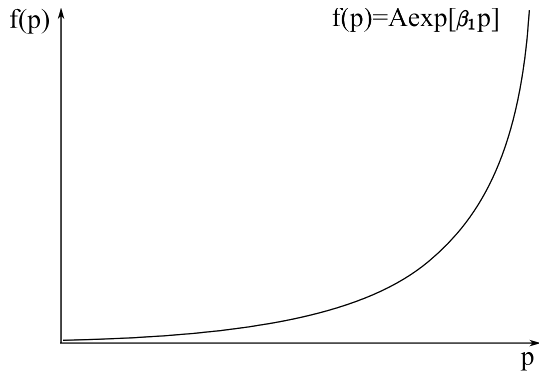
where $A$ and $B$ are constants of integration and $\beta$ is given by:

$$\beta_1 = -\frac{a}{\sigma^2} + \frac{\sqrt{a^2 + 2r\sigma^2}}{\sigma^2} > 0 \tag{4.5}$$

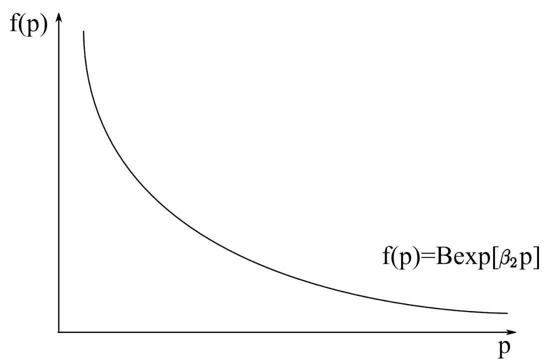
$$\beta_2 = -\frac{a}{\sigma^2} - \frac{\sqrt{a^2 + 2r\sigma^2}}{\sigma^2} < 0 \tag{4.6}$$

[2]One potential way around this problem is to use futures contracts which can actually be stored. Unfortunately, futures contracts on EEX have a maturity of only a few years and they are not very liquid, which rules out this option as well.

Because $\beta_2$ is negative, the second part of Equation 4.4 ($Be^{\beta_2 p}$) implies that the value of the option decreases as the value of dark spread increases. Intuition tells us this is incorrect and that the option value should increase with the increase in value of the underlying variable: the option to invest should be worth more, the higher the value of dark spread. Thus, I can eliminate the second part of Equation 4.4 from the general solution (a graphical interpretation of the meaning of the second part of Equation 4.4 is given in Figure 4.2b). The first part of Equation 4.4 incorporates $\beta_1$ which is positive, implying that the value of the option increases as the value of dark spread increases (Figure 4.2a graphically depicts the meaning of the first part of Equation 4.4). This is in line with economic logic, therefore, the solution to option value given in Equation 4.3b is given by Equation 4.7.




(A) Increasing function




(B) Decreasing function

FIGURE 4.2: Option values

$$f(p) = Ae^{\beta_1 p} \tag{4.7}$$

To find the optimal investment threshold and to determine the value of the constant $A$, I use the value matching and smooth pasting conditions which hold only at the optimal exercise price $p^*$ at which one should invest (Dixit (1993), Dixit and Pindyck (1994), Chapter 4). The expressions for value matching and smooth pasting conditions are given in Equations 4.8a and 4.8b respectively.

$$f(p) = V(p) - I \tag{4.8a}$$

$$f_p = V_p \tag{4.8b}$$

From the smooth pasting condition I find the value of the constant $A$:

$$A = \frac{qe^{-\beta_1 p}(1 - e^{-rT})}{r\beta_1} \tag{4.9}$$

By inserting the expression for $A$ and $V(p)$ into the value matching Equation 4.8a, I obtain that price at which it becomes optimal to invest equals:

$$p = \frac{q(a\beta_1 - a\beta_1 e^{rT} - r + e^{rT}r + a\beta_1 rT) + \beta_1 e^{rT} I r^2}{\beta_1(e^{rT} - 1)qr} \tag{4.10}$$

With the given parameters, it becomes optimal to invest when the dark spread price reaches €178.3 per MWh. The critical price is extremely high and it has never been recorded in observed data. Thus, if one assumes ABM, it is very unlikely that a coal power plant would ever be built. This is evidence that shows that the use of ABM in valuing investments in electricity generation cannot serve as a proxy for a more realistic mean reverting process. A plot of the threshold price is given in Figure 4.3.

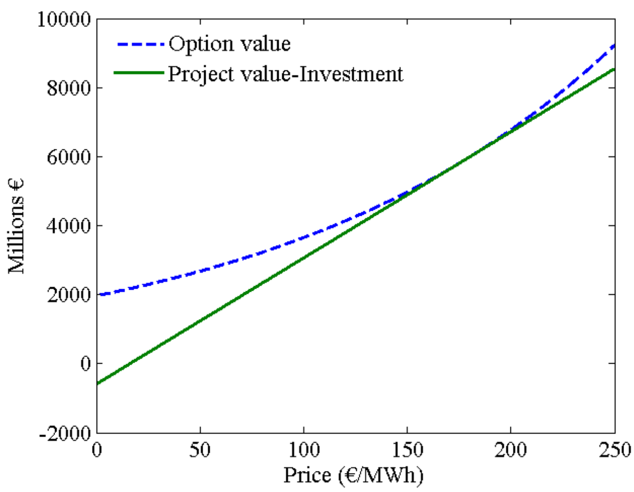



FIGURE 4.3: Optimal exercise price for ABM

The major culprit for such a high investment threshold is the extremely high volatility of dark spread. Figure 4.4 shows the dependence of the investment threshold price on the standard error of dark spread. As can be seen from

Figure 4.4, for lower values of the standard error, the investment threshold is also lower.

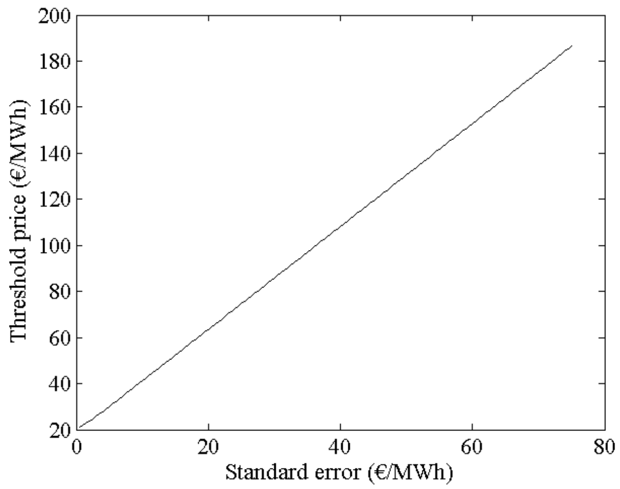



FIGURE 4.4: Threshold price as a function of standard error of dark spread

### 4.5.2 Ornstein-Uhlenbeck model

Assuming ABM for the evolution of dark spread prices resulted in a threshold price which was more than eight times higher than the threshold price obtained using NPV analysis. The reason for this is that ABM is an unbounded process. Thus, by assuming ABM, we implicitly assume that dark spread prices can move downward without bound, resulting in a high threshold price. Therefore, a high threshold price provides protection against a large drop in prices.

In this section, I assume an OU process for the evolution of dark spread prices. The value of a project where the dark spread follows an OU process can be calculated analytically and is given by:

$$V(p) = \int_0^T qE(p_t)e^{-rt}dt \tag{4.11a}$$

$$V(p) = q\left[\frac{p_0(1-e^{-(k+r)T})}{k+r} + \frac{m(1-e^{-rT})}{r} + \frac{m(e^{-(k+r)T}-1)}{k+r}\right] \tag{4.11b}$$

where I use the definition of $E(p_t)$ from Equation 3.10a. The option value is given by:

$$\frac{1}{2}f_{pp}\sigma^2 + f_p k(m-p) - rf = 0 \tag{4.12}$$

Unfortunately, Equation 4.12 does not have an analytical solution, therefore

I have to resort to numerical procedures (a detailed description of how to solve differential equations by numerical methods is included in Appendix B.1). To calculate the optimal investment threshold, I first add the time derivative ($f_t$) to Equation 4.12 and then use an implicit finite difference scheme (Wilmott (2006), Chapter 78). The derivatives in Equation 4.12 are approximated in the following way:

$$f_t = \frac{f_{i,j+1} - f_{i,j}}{dt} \qquad f_{pp} = \frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{dp^2} \qquad f_p = \frac{f_{i+1,j} - f_{i-1,j}}{2dp} \tag{4.13}$$

Inserting derivative approximations from Equation 4.13 into Equation 4.12, I obtain the following expression for the value of the option:

$$\frac{1}{2}\frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{dp^2}\sigma^2 + \frac{f_{i+1,j} - f_{i-1,j}}{2dp}k(m-p_i) + \frac{f_{i,j+1} - f_{i,j}}{dt} - rf_{i,j} = 0 \tag{4.14}$$

After some manipulation, Equation 4.14 can be written more compactly as:

$$f_{i,j+1} = A_i f_{i-1,j} + B_i f_{i,j} + C_i f_{i+1,j} \tag{4.15}$$

where $A$, $B$ and $C$ are defined as:

$$\begin{aligned} A_i &= \frac{dt}{2dp^2}\left((m - p_i)dpk - \sigma^2\right) \quad B_i = \frac{1}{dp^2}\left(dt\sigma^2 + dp^2(1 + dtr)\right) \\ C_i &= \frac{dt}{2dp^2}\left((p_i - m)dpk - \sigma^2\right) \end{aligned} \tag{4.16}$$

For the boundary conditions, I assume *Dirichlet* boundary conditions[3], i.e. for each time step $j$, the value of the option remains constant at $p_{min}$ and $p_{max}$ and is equal to the option value at expiration ($T$), i.e.:

$$f(p_{min}, j) = f(p_{min}, T) \qquad\qquad f(p_{max,j}) = f(p_{max}, T) \tag{4.17}$$

At expiration, the following terminal condition holds:

$$f(idp, T) = max[V(idp) - I, 0] \tag{4.18}$$

where $I$ denotes the investment cost and $V$ value of the project. Equation 4.18 implies that value of the option at expiration ($T$) is equal to the greater of exercise value or zero. Exercise value is determined as a difference between value of the project, which depends on the value of dark spread

[3]Please refer to Section B.1.1, and paragraph preceding Equation B.17

price, and investment cost.

Now, I solve the following set of simultaneous equations for each time step $j$ and price step $i = 0..M$:

$$\begin{aligned} f_{1,j+1} &= A_1 f_{0,j} + B_1 f_{1,j} + C_1 f_{2,j} \\ f_{2,j+1} &= A_2 f_{1,j} + B_2 f_{2,j} + C_2 f_{3,j} \\ &\dots \\ f_{M-1,j+1} &= A_{M-1} f_{M-2,j} + B_{M-1} f_{M-1,j} + C_{M-1} f_{M,j} \end{aligned} \tag{4.19}$$

Equations 4.19 can be written in matrix form as:

$$\begin{pmatrix} f_{1,j+1} \\ f_{2,j+1} \\ \dots \\ f_{M-1,j+1} \end{pmatrix} = \begin{pmatrix} B_1 & C_1 & \dots & \dots \\ A_2 & B_2 & C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & A_{M-1} & B_{M-1} \end{pmatrix} \cdot \begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \dots\dots \\ f_{M-1,j} \end{pmatrix} + \begin{pmatrix} A_1 \cdot f_{0,j} \\ 0 \\ \dots \\ C_{M-1} \cdot f_{M,j} \end{pmatrix} \tag{4.20}$$

Finally, starting at the terminal condition and moving backwards in time, I solve for the option value at time $j$:

$$\begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \dots\dots \\ f_{M-1,j} \end{pmatrix} = \begin{pmatrix} B_1 & C_1 & \dots & \dots \\ A_2 & B_2 & C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & A_{M-1} & B_{M-1} \end{pmatrix}^{-1} \cdot \left[ \begin{pmatrix} f_{1,j+1} \\ f_{2,j+1} \\ \dots \\ f_{M-1,j+1} \end{pmatrix} - \begin{pmatrix} A_1 \cdot f_{0,j} \\ 0 \\ \dots \\ C_{M-1} \cdot f_{M,j} \end{pmatrix} \right] \tag{4.21}$$

To approximate a perpetual American option, I evaluate the impact of option maturity on optimal exercise price. Basically, I let *time* increase and observe what happens to the optimal exercise price at the start of the option life. By making the option life sufficiently long, I try to mimic a perpetual option.

It turns out that I need to make the option life only a couple of years, as with increasing maturity the optimal exercise price does not change. The optimal investment threshold is found at the point where:

$$f(p) = V(p) - I$$

For the parameters shown in Table 4.1 and for the estimated parameters for the OU process from the previous chapter, and using a price increment of $dp = 0.2$ with $p_{min} = -70$ and $p_{max} = 200$ and time increment of $dt = 1/8760$, I obtain an optimal exercise price of €39 per MWh. A plot of the exercise boundary as a function of time to maturity is given in Figure 4.5. The reason why an option with such a short time to maturity can be used to approximate a perpetual option lies in the fact that OU is a mean reverting process, implying that if the price deviates from the long-run average, it will be pulled back. The stronger the mean reversion factor, the shorter one has to make the option life to mimic a perpetual option. On the other hand, the speed of convergence is not affected by the choice of time and price increment: they simply affect the level of threshold price.

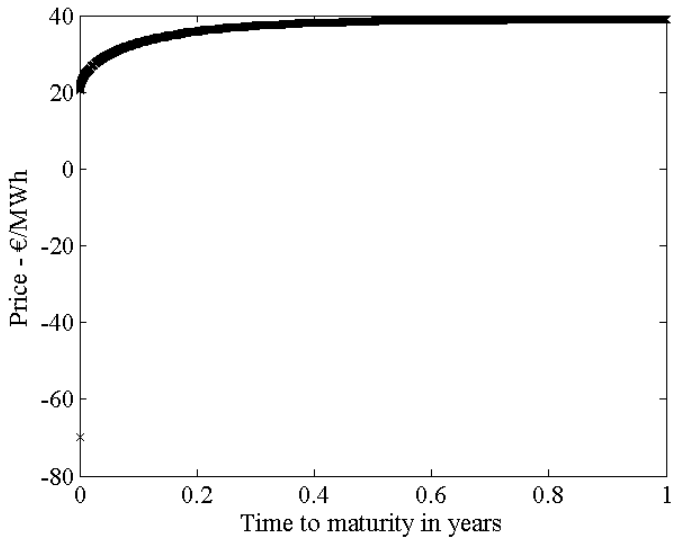



FIGURE 4.5: Exercise boundary for Ornstein-Uhlenbeck process as a function of option life

The threshold price obtained under OU is about twice as large as that obtained under NPV analysis. This higher threshold is a consequence of the stochastic behavior of dark spread price and of the irreversible nature of the investment: factors that traditional capital budgeting analysis does not take into account. Nevertheless, the threshold price is significantly lower than that obtained when one assumes ABM for the evolution of dark spread prices. This is because a project whose state variable follows an OU process as opposed to ABM is less risky: the price always reverts back to the long-run average.

### 4.5.3 Cox-Ingersoll-Ross model

If the dark spread prices are assumed to be strictly non-negative, than one option is to assume that the dark spread prices follow a CIR process.

Because the expected value of the variable following CIR process is the same as the expected value of the variable following OU process, the project value ($V(p)$) is also the same as in the case of OU process and is given by:

$$V(p) = q\left[\frac{p_0(1-e^{-(k+r)T})}{k+r} + \frac{m(1-e^{-rT})}{r} + \frac{m(e^{-(k+r)T}-1)}{k+r}\right] \tag{4.22}$$

To get the value of the option, I need to solve the following differential equation:

$$\frac{1}{2}f_{pp}\sigma^2 p + f_p k(m-p) - rf = 0 \tag{4.23}$$

Unlike the OU process, it is possible to obtain an analytical solution to the investment problem. Following Ewald and Wang (2010), I first divide Equation 4.23 by k. Then I define:

$$f(p) = w(z) \tag{4.24a}$$
$$z = \frac{2kp}{\sigma^2} \tag{4.24b}$$
$$f_p = w_z\frac{2k}{\sigma^2} \tag{4.24c}$$
$$f_{pp} = w_{zz}\frac{4k^2}{\sigma^4} \tag{4.24d}$$

Using the above expressions, Equation 4.23 becomes Kummer's differential equation:

$$zw_{zz} + w_z(b-z) - aw = 0 \tag{4.25}$$

where I define: $a = \frac{r}{k}$ and $b = \frac{2km}{\sigma^2}$. The solutions to Equation 4.25 are Kummer's M and U functions[4] given by:

$$f(p) = A_1 KummerM\left(\frac{r}{k}, \frac{2km}{\sigma^2}, \frac{2kp}{\sigma^2}\right) + A_2 KummerU\left(\frac{r}{k}, \frac{2km}{\sigma^2}, \frac{2kp}{\sigma^2}\right) \tag{4.26}$$

Because $\lim_{p\to 0} f(p) << \infty$, I can eliminate the KummerU function by assuming $A_2$ to be zero. Thus, the solution is:

$$f(p) = A_1 KummerM\left(\frac{r}{k}, \frac{2km}{\sigma^2}, \frac{2kp}{\sigma^2}\right) \tag{4.27}$$

[4]The confluent hypergeometric function $F_1(a;b;z)$ also known as Kummer's function of the first kind, is a function which arises as a solution to confluent hypergeometric differential equation sometimes called Kummer's differential equation (Abramowitz and Stegun (1972) page 504, Weisstein (n.d.)).

To determine the threshold price at which the investment should be made, I use value matching and smooth pasting conditions, namely:

$$f(p) = V(p) - I \tag{4.28a}$$
$$f_p = V_p \tag{4.28b}$$

The smooth pasting condition gives the value of the constant $A_1$ equal to[5]:

$$A_1 = \frac{e^{(-k-r)T}\left(-1+e^{(k+r)T}\right)kmq}{r(k+r)\text{KummerM}\left(1+\frac{r}{k}, 1+\frac{2km}{\sigma^2}, \frac{2kp}{\sigma^2}\right)} \tag{4.29}$$

Using the expression for $A_1$ and the value matching condition, I determine the optimal investment threshold. The investment threshold is not given in analytical form but rather has to be found numerically. The value at which the investment should be undertaken is €38.3 per MWh. Plots of the option value and project value are given in Figure 4.6.

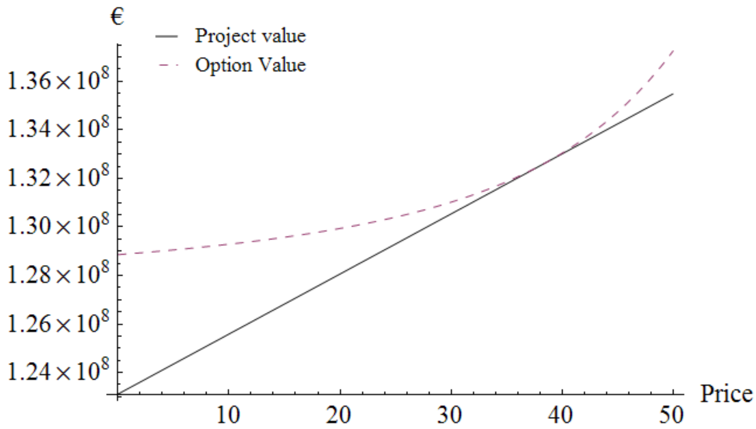



FIGURE 4.6: Optimal exercise price for CIR process

### 4.5.4 Schwartz one-factor model

Another process where dark spread prices can only attain positive values is the Schwartz one-factor model. The value of the plant under this process is given by:

$$V(p) = q \cdot \int_0^T E(p_t)e^{-rt}dt \tag{4.30}$$

[5]The calculations have been performed in *Mathematica* where KummerM function is denoted as HypergeometricF1 function.

To calculate the value of the project, I need an expression for the expected dark spread price. According to Schwartz (1997), the level of dark spread is log normally distributed. Thus, if $x$ is a logarithm of dark spread having mean and variance given by equations 3.25a and 3.25b, then the expected value of the level of dark spread is given by (Aitchison and Brown (1957)):

$$E[p_t] = exp[\mu_x + \frac{1}{2}Var_x] \tag{4.31a}$$

$$E[p_t] = exp[x_0 e^{-kt} + m(1 - e^{-kt}) + \frac{\sigma_p^2}{4k}(1 - e^{-2kt})] \tag{4.31b}$$

Therefore, the value of the project is given by:

$$V(p) = q \cdot \int_0^T exp\left[x_0 e^{-kt} + m(1 - e^{-kt}) - rt + \frac{\sigma_p^2}{4k}(1 - e^{-2kt})\right] \tag{4.32}$$

Nevertheless, the integral in Equation 4.32 does not have an analytical solutions, hence its value has to be calculated numerically. To determine the value of the option, I need to solve the following differential equation:

$$\frac{1}{2}f_{pp}\sigma^2 p^2 + f_p k(a - lnp)p + f_t - rf = 0 \tag{4.33}$$

Just like in the case of the OU process, the equation cannot be solved analytically, thus I resort to numerical procedures and use an implicit finite difference scheme. Again, derivatives are approximated just like for the OU process shown in Equation 4.13. For each time step $j$ and price step $i$, I solve the following set of equations:

$$f_{i,j+1} = A_i f_{i-1,j} + B_i f_{i,j} + C_i f_{i+1,j} \tag{4.34}$$

Here, the values of parameters $A$, $B$ and $C$ are defined as:

$$\begin{aligned} A_i &= \frac{dt}{2}\left((a - lnp_i)ik - \sigma^2 i^2\right) \quad B_i = 1 + dt(\sigma^2 i^2 + r) \\ C_i &= \frac{dt}{2}\left((lnp_i - a)ik - \sigma^2 i^2\right) \end{aligned} \tag{4.35}$$

Thus, using an implicit finite difference scheme and the same approach as in the OU case, I calculate that the optimal exercise price is equal to €41.4 per MWh. A plot of the exercise boundary is shown in Figure 4.7.

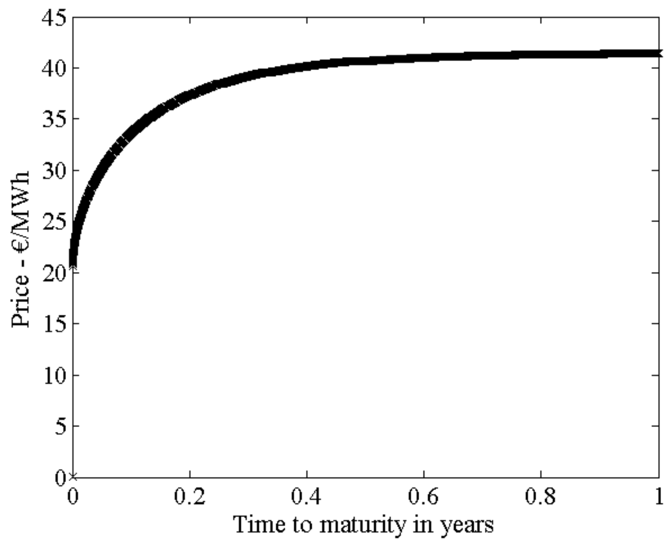



FIGURE 4.7: Exercise boundary for Schwartz process as a function of option life

## 4.6 Selection of stochastic process

So far I have analyzed the impact of four stochastic processes on the investment threshold. I find that of the four processes, ABM is the least useful as it suggests investing at an extremely high price, which never occurs in the observed data. Therefore, assuming that the dark spread follows ABM, one would be very unlikely to ever invest in a new coal-fired power plant. The other three processes are mean reverting and they give rather comparable results. Table 4.2 shows the investment thresholds for four stochastic processes. From the table, it can be seen that the choice of mean reverting process does not fundamentally affect the investment threshold.

| Process | Value (€/MWh) |
|---|---|
| NPV | 20.6 |
| ABM | 178.3 |
| OU | 39 |
| CIR | 38.3 |
| Schwartz | 41.4 |

TABLE 4.2: Threshold prices for different stochastic processes

The reason why the investment threshold does not vary significantly among mean reverting processes is because they revert to the long-run price level soon after they move away from it. Table 4.3 shows values of half life for mean reverting stochastic processes (for details on calculating the half life, see Section B.2). As can be seen, it takes approximately two weeks for the prices to revert half way back to the long-run level from the starting price if there are no stochastic disturbances.

| Process | half-life |
|---|---|
| OU | 17 days |
| CIR | 16.9 days |
| Schwartz | 13.2 days |

TABLE 4.3: Half lives of mean reverting processes

With regard for the distributional properties and RMSE discussed in Chapter 3, I conclude that the CIR process is the most appropriate process for use in the RO valuation of coal-fired plants where the major source of uncertainty is the dark spread price. Nevertheless, as shown in this chapter, the choice of another mean reverting process is also acceptable as all mean reverting processes give very similar investment threshold values.

Next, I perform a sensitivity analysis to investigate the most crucial parameters affecting the threshold price.

### 4.6.1 Sensitivity analysis

Section 4.5.3 shows that if the dark spread prices are assumed to follow a CIR process, it is optimal to invest in a new coal-fired power plant when the dark spread price reaches €38.3 per MWh. Next, I perform a sensitivity analysis in which I investigate the impact of three crucial variables on the threshold price: discount rate, volatility and mean reverting factor.

In the traditional NPV approach, a higher discount rate implies a lower NPV, i.e. higher break-even price. In terms of the CIR process, this relationship is not so straightforward, as can be seen from Figure 4.8. For values of discount rate from 4% and up, a higher discount rate implies a higher threshold price, although the relationship is exponential. For discount rates below 4%, a lower discount rate also implies a higher threshold price. As Dixit and Pindyck (1994) explain on page 156, this phenomenon is a manifestation of option value whereby lower interest rate makes the future more important thus increases the cost of exercising the option and in effect increases the threshold value.

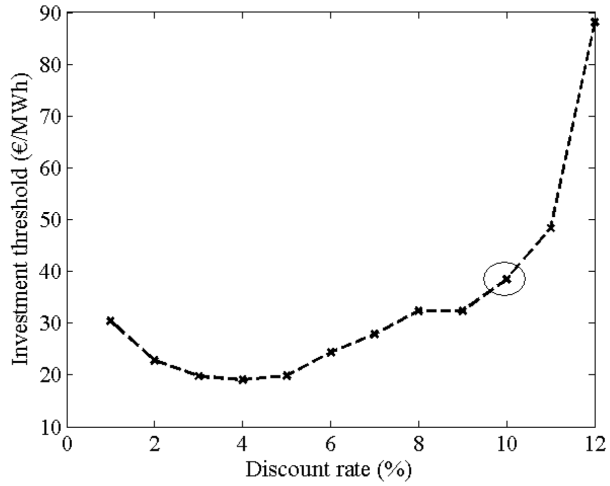



FIGURE 4.8: Impact of discount factor on investment threshold

In the following, I test the impact of volatility on investment threshold. The RO literature states that higher volatility leads to higher investment thresholds. Varying the standard error from 1 €/MWh to 25 €/MWh, a clear positive link emerges between volatility and investment threshold as shown in Figure 4.9. Therefore, I conclude that the volatility of the dark spread price dramatically affects the value of investment threshold.

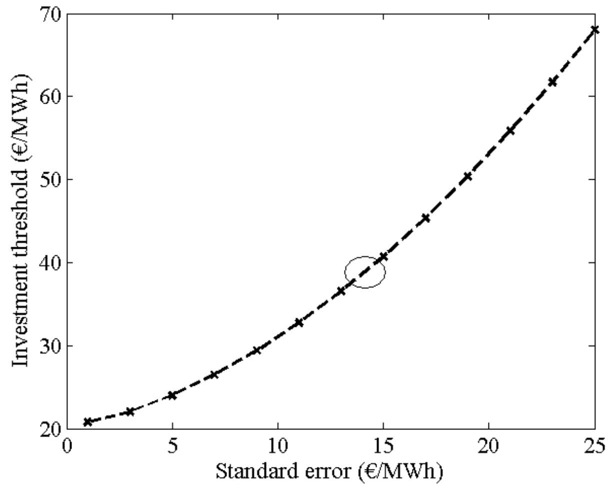



FIGURE 4.9: Impact of standard error on investment threshold

Finally, I investigate the impact of mean reversion on investment threshold. As expected, the relation between investment threshold and mean reversion factor is inverse. The investment threshold reduces exponentially as the mean reversion factor increases, as shown in Figure 4.10.

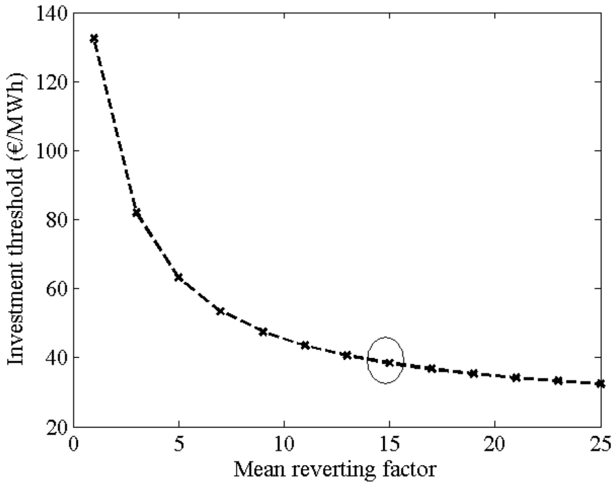



FIGURE 4.10: Impact of mean reversion on investment threshold

## 4.7 Conclusion

In this chapter, I analyzed how the choice among different stochastic processes affects the investment decision. Of the four stochastic processes, ABM turned out to be the least useful - it predicts that investment should only be undertaken at extreme prices, which are very unlikely to ever occur. Furthermore, as shown in Chapter 3, ABM performs badly in terms of distributional properties and RMSE. Therefore, I conclude that it is not appropriate to use ABM to model investment decision.

The other three processes are all mean reverting and, prior to the analysis, one would expect that the OU process results in the highest investment threshold while the CIR and Schwartz models result in lower and perhaps rather comparable investment thresholds. The reason for this conjecture lies in the fact that the OU process admits negative prices, therefore one would require a higher threshold price at which to invest.

Actual calculations partially confirm this conjecture. The lowest threshold price was obtained under the CIR process: €38.3 per MWh. Assuming the dark spread follows an OU process, the investment trigger is at €39 per MWh, while assuming a Schwartz one-factor process, the investment trigger is at €41.4 per MWh. I explain the higher threshold price of the Schwartz model compared to the OU process by noting that it is necessary to carry out numerical calculations of the value of the project, which in the case of the OU process can be done analytically. Furthermore, another reason might lay in the fact that the estimation of parameters in the Schwartz process is done using log prices.

Consequently, the choice of mean reverting process does not bear too much weight on the investment trigger: investment thresholds for all three mean reverting processes are within a few €. From a practitioner's point of view, the CIR process seems the best candidate to use in RO analysis if dark spread is used as a single source of uncertainty. There are several reasons for this statement. First, the CIR process comes very close to observed data in terms of distributional statistics. Second, it performs best in terms of RMSE when compared to other stochastic processes. Finally, CIR precludes negative dark spread values, which is an important characteristic if one assumes that dark spread prices cannot become negative [6].

Comparing the results obtained under RO and NPV analysis, one might wonder why the threshold dark spread price is so much higher under RO analysis compared to the level required to make the project profitable and as obtained under NPV analysis. Explanation for this phenomenon can be found in the workings of two factors:

- In the RO analysis dark spread prices evolve stochastically. The fact that they are mean reverting does not say much about the proportion of time the price will actually drop below the level required to make the project profitable. In traditional project valuation, dark spread is assumed to be constant.
- Investment is irreversible i.e. there is an option value of investment which is very high. This intuitively makes sense as once the investment in coal fired plant is undertaken there is no way back. In the RO framework, the investment will be undertaken when the value of the project is equal to the investment cost plus the value of the option. In traditional capital budgeting, the investment will be undertaken when the value of the project exceeds only investment cost. Therefore, the greater the irreversibility, the greater the option value, and the greater the difference between RO and NPV threshold.

Combination of these two factors results in investment threshold under the RO framework which is substantially higher than in the case of traditional capital budgeting approach.

[6]While theoretically possible, it is practically extremely unlikely for weekly dark spread prices to become negative as generators will suspend the operations in case of prolonged period of negative prices. Furthermore, prolonged periods of negative prices, such as one week, would imply there is extra capacity in the market, which would induce less efficient generators to mothball their facilities, reducing supply and hence increasing prices. Thus one can claim negative dark spread prices can occur, but only as very rare events (again, in my data set I do not observe any negative weekly prices during the eight year period).

Chapter 5

# Two factor real options model

## 5.1 Introduction

In the previous chapter, I calculated the optimal investment threshold for a coal-fired power plant using dark spread prices as the only source of uncertainty. The question that arises is whether by using dark spread prices I am losing valuable information contained in individual coal and electricity prices. Therefore, in this chapter, I model electricity and coal prices separately. This will allow me to obtain the optimal investment threshold as a function of the individual electricity and coal prices. Unlike the previous chapter where the investment decision was evaluated using the finite difference method, in this chapter I use the Monte Carlo least squares (MCLS) approach proposed by Longstaff and Schwartz (2001).

The chapter is organized as follows. In section 5.2, I analyze electricity prices, while in section 5.3 I analyze coal prices. In section 5.4, I perform an investment analysis and the last section concludes.

## 5.2 Electricity prices

Researchers modeling electricity prices have primarily been focused on applying electricity models to derivative pricing and not to real options valuation. Modeling electricity prices for use in real options valuation differs from modeling used for derivative pricing. While for derivative pricing short-term fluctuations are very important, they are less significant for long-term investment analysis such as real options. Investors looking to invest in large projects with useful lives of forty years are interested less in daily and seasonal fluctuations and more in the long-term evolution of electricity prices.

### 5.2.1 Literature review

One of the earliest attempts to model electricity prices in liberalized markets is the work by Pilipovic (1997). In her book, she does not deal solely with electricity prices but discusses some of the important issues in electricity price modeling, including market characteristics, price behavior and the statistical properties of electricity prices. In terms of the models used, she analyzes log-normal, mean reverting and cost-based models. In addition, she also introduces seasonality in the analysis through the use of trigonometric functions.

Clewlow and Strickland (2000) have written a book on energy derivatives which focuses a little more on electricity price modeling than Pilipovic (1997). Also, unlike Pilipovic (1997), the authors acknowledge the need to incorporate jumps in models of electricity prices. Both of these books represent general reference material and do not go deep into the analysis of a particular electricity market.

Lucia and Schwartz (2002) model electricity prices in the NordPool market for the purpose of derivative pricing. They describe the behavior of spot prices in terms of a two-component mean-reverting process: a deterministic part which explains predictable movements in prices and a stochastic component following a certain stochastic process. They propose one- and two-factor models for the stochastic component, each based on the level and log of price. They assess the models by comparing the theoretical derivative prices obtained under estimated models with actual data. As might have been predicted, they conclude that two-factor models give better results than one-factor models.

Escribano et al. (2002) analyze electricity prices in Argentina, NordPool, Australia (Victoria), New Zealand (Hayward), US (Pennsylvania-New Jersey-Maryland: PJM) and Spain. They try to develop a general model that captures the major characteristics of these markets: a model that is flexible enough to be applied to each of them. They propose a model that includes mean reversion, seasonality, non-constant volatility (GARCH) and time-dependent jumps. With regard to the seasonal component, the authors model it in two ways: using sinusoidal functions and dummy variables. An empirical analysis revealed that there is not much difference between these two approaches. They conclude that mean-reverting behavior is obvious and strong in all markets except Scandinavia; the reason for this lies in the abundance of hydro potential which allows for smoothing of supply/demand imbalances (hydro potential serves as storage). Furthermore, they find that electricity prices have non-constant volatility and time-dependent jumps.

Knittel and Roberts (2005) analyze the level and log of hourly electricity prices in California. They use the following models: mean-reverting with constant mean, mean-reverting with time-varying mean, mean-reverting jump diffusion model, mean-reverting jump diffusion model with time-varying jump intensity, ARMAX, EGARCH and ARMAX incorporating weather data. The adequacy of each of the models is assessed by performing out-of-sample forecasts and calculating the RMSE. The authors reach a few conclusions. First, there is no significant difference between using either the level or log of prices in the analysis. The reason for this is that forecasting performance is mostly affected by the large kurtosis of electricity prices, which cannot be generated by the selected models. Second, they find that the ARMAX model is the best model in terms of RMSE.

Seifert and Uhrig-Homburg (2007) analyze prices from EEX with an emphasis on understanding and modeling jumps in electricity prices. The authors decompose the logarithm of electricity prices into deterministic and stochastic components. The deterministic component is modeled using a standard approach in the literature, i.e. using sinusoidal and dummy variables. For the stochastic component, they use a two-factor model representing the short-term and long-term price deviations. For the long-term component, they assume a Brownian motion process while for the short-term component they use an Ornstein-Uhlenbeck process. As the jumps are the consequence of short-term variations, they experiment with different jump specifications in the short-term component. They find that the best model for describing the observed prices is an exponential jump model with stochastic jump intensity.

### 5.2.2 Statistical properties of electricity prices

Electricity is different from most other commodities. In particular, electricity prices are characterized by extreme volatility and outliers (also known as 'spikes'). There are several reasons that can explain this behavior. The major reason is that electricity cannot be stored: demand has to match supply at all instances. Other reasons include:

- Demand is extremely inelastic as electricity is an essential commodity for households and industry.
- Electricity consumption is not charged on a real-time basis, implying that consumers do not have an incentive to adjust their consumption in response to temporary shortages.
- Electricity production and consumption are extremely weather-dependent. The amount of rainfall affects the amount of electricity generated. On the consumption side, extreme temperatures induce consumers to use air conditioning in summer and electric heaters in winter.
- The shape of the supply curve plays an important role. At the lower levels of production, electricity is generated using low marginal cost technologies such as nuclear, large hydro and coal. But as the system reaches its peak, the supply curve becomes more vertical. Combining a vertical supply curve with a vertical demand curve (because electricity consumption is extremely inelastic), even the smallest changes in demand can result in significant price increases. This is illustrated in Figure 5.1. The figure shows that if demand is at the level represented by curve D2, then even the smallest changes in demand or supply will result in large price changes.

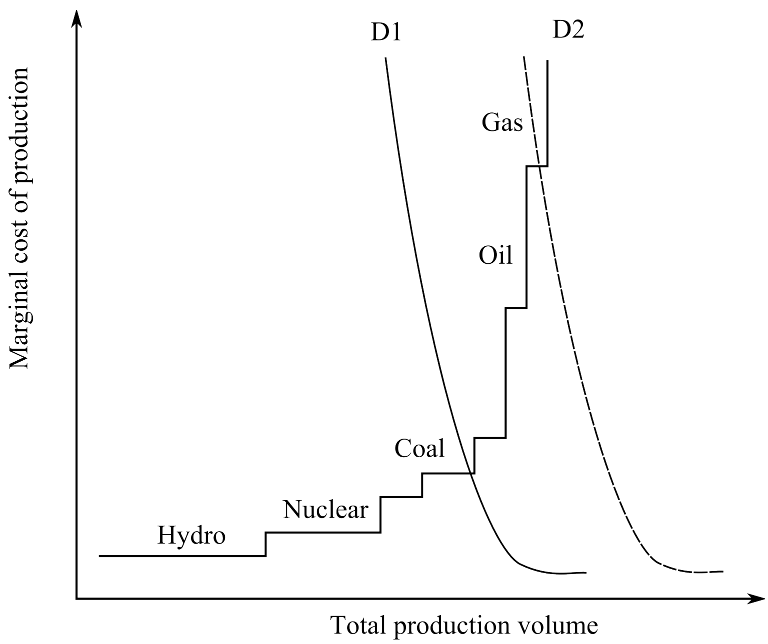



FIGURE 5.1: Stylized supply and demand curve

I use data from the European Energy Exchange in Leipzig to model electricity prices. The data sample consists of the arithmetic averages of the weekly base load electricity prices going from January 1, 2002 to February 26, 2010. A plot of the prices together with the corresponding histogram is given in Figure 5.2 and summary statistics are given in Table 5.1.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed prices | 10.14 | 35.61 | 39.82 | 135.69 | 17.25 | 1.35 | 5.53 |

TABLE 5.1: Summary statistics for weekly electricity prices

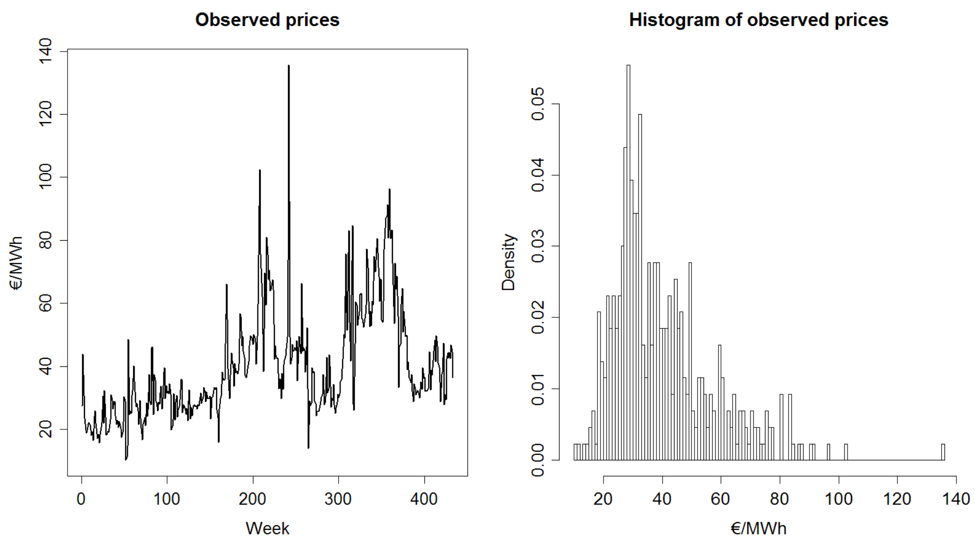



FIGURE 5.2: Weekly electricity prices

### 5.2.3 Modeling electricity prices

In most of the literature on electricity price modeling, the authors try to decompose electricity prices into stochastic and seasonal components. But most of this analysis is done with the purpose of derivative pricing (for instance Schwartz (1997), Schwartz and Smith (2000), Seifert and Uhrig-Homburg (2007)). I argue that such an approach is not necessary when valuing long-term investments. For a potential investor what matters is long term price evolution, which means that investment decisions are not affected by seasonal variations.

It is well established now that electricity prices are mean-reverting (e.g. Pilipovic (1997), Clewlow and Strickland (2000), Lucia and Schwartz (2002), Escribano et al. (2002), Knittel and Roberts (2005)). Therefore, in this chapter I will focus on three mean-reverting processes to model electricity prices: Ornstein-Uhlenbeck (OU), Cox-Ingersoll-Ross (CIR) and Schwartz one-factor model.

#### 5.2.3.1 Ornstein-Uhlenbeck model

The first process I use to model electricity prices is the OU model. The stochastic differential equation (SDE) describing the evolution of prices following the OU process is given in Equation 5.1, where the variables have the following interpretation: $p$ represents the price level, $\mu$ the long-run electricity price, $k$ the mean-reverting factor, $\sigma$ the volatility of the process, $dz$ is equal to $\epsilon\sqrt{dt}$ where $\epsilon \sim N(0,1)$ and $Cov(\epsilon_t, \epsilon_s) = 0$ for $t \neq s$, and $dt$ represents a time increment.

$$dp_t = k(\mu - p_t)dt + \sigma dz \tag{5.1}$$

Following the same approach as in Chapter 3, I find that the solution to Equation 5.1 is given by the following expression:

$$p_t = p_0 e^{-kt} + \mu(1 - e^{-kt}) + \sigma \int_0^t e^{k(s-t)} dz_s \tag{5.2}$$

The expected value and variance of the OU process are given by Equations 5.3a and 5.3b.

$$E[p_t] = p_0 e^{-kt} + \mu(1 - e^{-kt}) \tag{5.3a}$$

$$V[p_t] = \frac{\sigma^2}{2k}\left(1 - e^{-2kt}\right) \tag{5.3b}$$

Given the vector of observed prices $\boldsymbol{p}$, I estimate the values of the long-run electricity price level ($\mu$), mean-reverting factor ($k$) and volatility ($\sigma$). I assume that these parameters are contained in a parameter vector $\boldsymbol{\theta}$. To estimate the values of the parameters, I use the maximum likelihood (ML) approach. For an OU process, the variables are distributed normally, therefore the log-likelihood function is given by the following expression:

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = -(n-1) \cdot ln\sqrt{(Var(p))} - \sum_{i=2}^{n} \left[ \frac{(p_i - E(p))^2}{2 \cdot Var(p)} \right] \tag{5.4}$$

Inserting the expressions for the expected value and variance from Equations 5.3a and 5.3b into Equation 5.4, I obtain a complete specification for the log-likelihood function given in Equation 5.5.

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = -(n-1) \cdot ln\zeta - \sum_{i=2}^{n} \left[ \frac{(p_i - (p_{i-1}e^{-kt} + \mu(1 - e^{-kt})))^2}{2\zeta^2} \right] \tag{5.5}$$

where I define:

$$\zeta = \sqrt{Var(p)} = \sigma\sqrt{\frac{1-e^{-2kt}}{2k}} \tag{5.6}$$

Using the likelihood function from Equation 5.5, I estimate values for the parameters and these are given in Table 5.2.

| | Weekly values |
|---|---|
| Long-run price ($\mu$) | 39.9663<br>(2.77) |
| Mean reversion ($k$) | 0.1793<br>(0.032) |
| Standard deviation ($\zeta$) | 8.6752<br>(0.32) |
| Volatility ($\sigma$) | 10.3230<br>(0.38) |

TABLE 5.2: Estimated parameters for OU process

To assess the goodness of fit of the OU process, I first simulate 10,000 trajectories using the estimated parameters. I use the following expression to perform the simulation:

$$p_{t+1} = p_t e^{-k\Delta t} + \mu(1 - e^{-k\Delta t}) + \sigma\sqrt{\frac{1-e^{-2k\Delta t}}{2k}} \cdot N(0,1) \tag{5.7}$$

Given that I simulate weekly prices, I set $\Delta t = 1$. Table 5.3 shows the statistical properties of 10,000 simulations while Figure 5.3 shows a histogram of simulated and observed prices. Two properties of the simulated prices stand out. First, the OU process generates negative values which are not present in the observed prices and are generally very unlikely to occur[1]. Second, the OU process cannot generate sufficiently large spikes, therefore the kurtosis of simulated prices is lower than in the observed data. On the other hand, the OU process captures relatively well the mean value and standard error of the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 10.14 | 35.61 | 39.82 | 135.69 | 17.25 | 1.35 | 5.53 |
| Simulated | -26.89 | 37.90 | 38.12 | 110.43 | 14.99 | 0.11 | 2.97 |

TABLE 5.3: Summary statistics for level of observed and simulated prices of OU process

[1]While negative values are possible for hourly prices and rarely for average daily prices, observing negative weekly prices seems very unlikely.

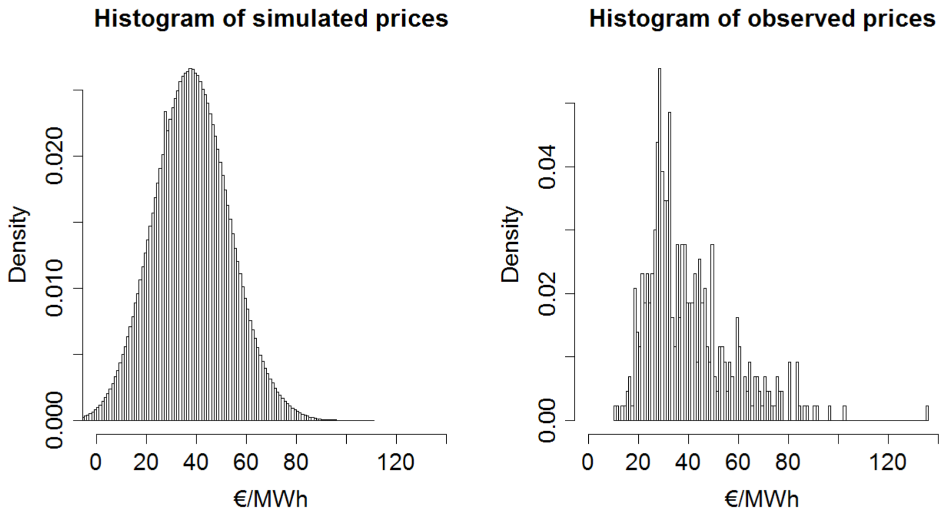



FIGURE 5.3: Histograms of simulated and observed prices of OU process

As an illustration of how well the OU process describes the observed data, Figure 5.4 shows a sample path generated using the estimated parameters.

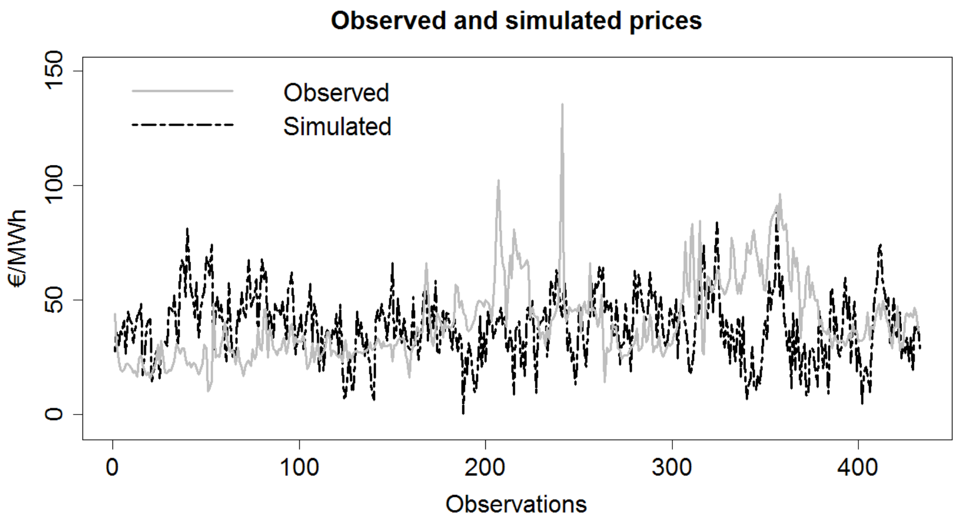



FIGURE 5.4: A simulated path of electricity prices following OU process

As a second measure of the goodness of fit, I calculate the RMSE which is displayed in Table 5.4.

| | mean | se. |
|---|---|---|
| RMSE | 22.93 | 0.40 |

TABLE 5.4: RMSE for in sample forecasts based on OU process

#### 5.2.3.2 Cox-Ingersoll-Ross model

The SDE describing the evolution of electricity prices following the CIR process is given in Equation 5.8, where the variables have the same interpretation as in Equation 5.1.

$$dp_t = k(\mu - p_t)dt + \sigma\sqrt{p_t}dz_t \tag{5.8}$$

The CIR process is capable of solving some of the shortcomings of the OU process. First, the CIR process generates strictly non-negative prices. Formally, if $2k\mu/\sigma^2 \geq 1$, the process never reaches zeros; if $0 < 2k\mu/\sigma^2 < 1$, zero serves as a reflecting barrier; and if $k\mu = 0$, zero is an absorbing barrier and it is reached surely in finite time (Overbeck and Ryden (1997)). Second, the volatility of the CIR process is non-constant and depends on the price level.

Given a vector of observed prices $\boldsymbol{p}$, I estimate a parameter vector $\boldsymbol{\theta}$ consisting of the long-run price level ($\mu$), volatility ($\sigma$) and mean-reverting factor ($k$) based on the ML approach. The log-likelihood function used to estimate the parameters is given in Equation 5.9 where $s > t$.

$$lnL(\boldsymbol{\theta}|\boldsymbol{p}) = (n-1)lnc + \sum_{i=2}^{n}\left\{-u - v + 0.5qln\left(\frac{v}{u}\right) + ln\{I_q(2\sqrt{uv})\}\right\} \tag{5.9}$$

where:

$$c = \frac{2k}{\sigma^2(1-e^{-k(s-t)})} \tag{5.10a}$$

$$u = cp_te^{-k(s-t)} \tag{5.10b}$$

$$v = cp_s \tag{5.10c}$$

$$q = \frac{2k\mu}{\sigma^2} - 1 \tag{5.10d}$$

$I_q$ is the modified Bessel function of the first kind of order q (5.10e)

The values of the estimated parameters are reported in Table 5.5.

| | Weekly values |
|---|---|
| Long-run price ($\mu$) | 39.9594 (2.56) |
| Mean reversion ($k$) | 0.1721 (0.031) |
| Volatility ($\sigma$) | 1.4440 (0.054) |

TABLE 5.5: Estimated parameters for CIR process

To test the goodness of fit of the CIR model, I first simulate 10,000 random trajectories. The transition density of $p_s$ conditional on $p_t$ is given by the following expression (Glasserman (2003)):

$$p_s = \frac{\sigma^2(1-e^{-k(s-t)})}{4k}\chi^2_d\left(\frac{4ke^{-k(s-t)}}{\sigma^2(1-e^{-k(s-t)})}p_t\right) \quad s > t \tag{5.11}$$

where

$$d = \frac{4\mu k}{\sigma^2} \tag{5.12}$$

Summary statistics for the simulated trajectories are given in Table 5.6 and histograms of simulated and observed trajectories are given in Figure 5.5. As was the case with the OU process, CIR comes very close to the observed data in terms of mean and standard error. Moreover, it improves over the OU process in several aspects. First, it generates strictly positive electricity prices which results in positive skewness. Also, due to the non-constant volatility, the CIR process has higher kurtosis than is the case with the OU process, albeit lower than in the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 10.14 | 35.61 | 39.82 | 135.69 | 17.25 | 1.35 | 5.53 |
| Simulated | 1.13 | 37.76 | 39.77 | 174.59 | 15.50 | 0.79 | 3.93 |

TABLE 5.6: Summary statistics for observed and simulated prices following CIR process

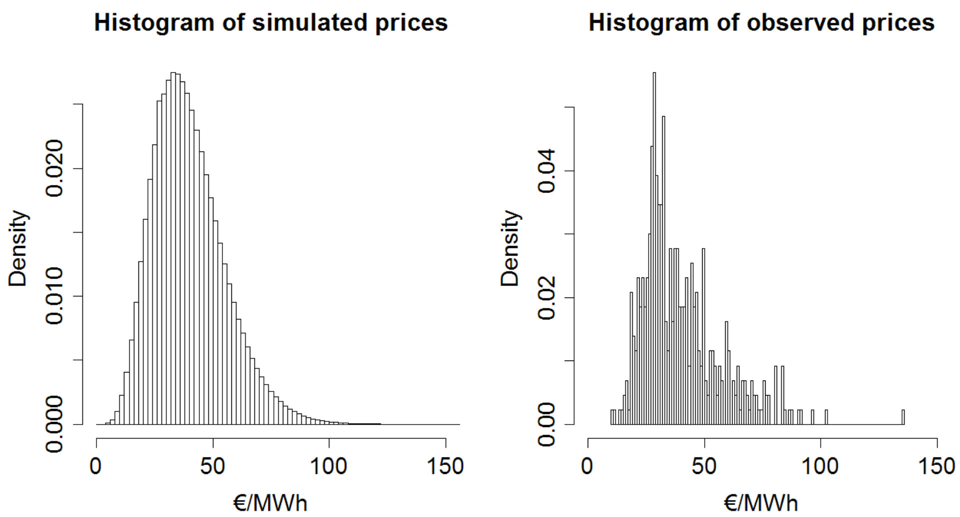



FIGURE 5.5: Histograms of simulated and observed prices for CIR process

For illustration purposes, in Figure 5.6 I show a sample trajectory of the simulated path compared to an observed path.

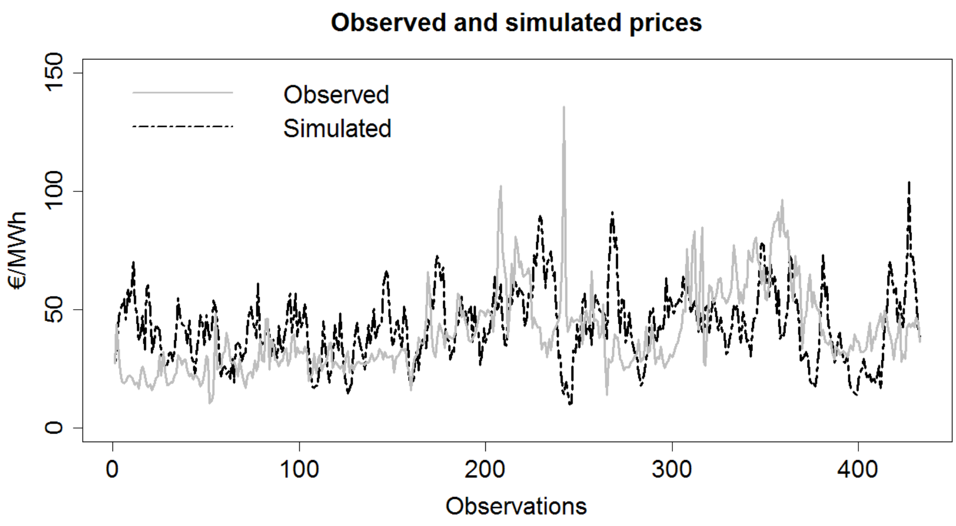



FIGURE 5.6: A simulated path of electricity prices following CIR process

As a second measure of the goodness of fit, I calculate the RMSE: the values of the parameters are given in Table 5.7.

| | mean | se. |
|---|---|---|
| RMSE | 22.98 | 1.88 |

TABLE 5.7: RMSE for in sample forecasts based on CIR process

#### 5.2.3.3 Schwartz one-factor model

The Schwartz one-factor model is the third mean-reverting process that I evaluate. The evolution of price is described by the SDE given in Equation 5.13. The interpretation of parameters is the same as in the case of the OU and CIR processes except that $a$ is used instead of $\mu$.

$$dp_t = k(a - lnp_t)p_t dt + \sigma p_t dz_t \tag{5.13}$$

In order to estimate the parameters of the Schwartz one-factor model, one needs to use the Ito lemma and transform the process into logarithmic scale by defining $x = lnp$. This results in a new SDE given in Equation 5.14, where I define $m = a - \frac{\sigma^2}{2k}$.

$$dx_t = k(m - x_t)dt + \sigma dz_t \tag{5.14}$$

The expected value and variance of the logarithm of the Schwartz process are given in Equations 5.15a and 5.15b.

$$E[x_t] = \mu_x = x_0 e^{-kt} + m(1 - e^{-kt}) \tag{5.15a}$$

$$V[x_t] = \frac{\sigma^2}{2k}\left(1 - e^{-2kt}\right) \tag{5.15b}$$

The logarithm of the Schwartz model is distributed normally, therefore the log-likelihood function used to estimate the parameters is given in Equation 5.16.

$$lnL(\boldsymbol{\theta}|\boldsymbol{x}) = -(n-1) \cdot ln\sqrt{(Var(x))} - \sum_{i=2}^{n}\left[\frac{(x_i - E(x))^2}{2 \cdot Var(x)}\right] \tag{5.16}$$

Upon inserting appropriate expressions for the expected value and variance of the logarithm of price from Equations 5.15a and 5.15b into Equation 5.16, the log-likelihood function becomes:

$$lnL(\boldsymbol{\theta}|\boldsymbol{x}) = -(n-1) \cdot ln\zeta - \sum_{i=2}^{n}\left[\frac{(x_t - (x_{t-1}e^{-kt} + m(1 - e^{-kt})))^2}{2\zeta^2}\right] \tag{5.17}$$

where I use the following short-hand notation in Equation 5.17.

$$\zeta = \sqrt{Var[x_t]} = \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}} \tag{5.18}$$

Estimated parameters are given in Table 5.8.

| | Weekly values |
|---|---|
| Long-run price ($\mu$) | 3.6055 (0.07) |
| Mean reversion ($k$) | 0.1560 (0.03) |
| Standard deviation ($\zeta$) | 0.2090 (0.007) |
| Volatility ($\sigma$) | 0.2256 (0.008) |

TABLE 5.8: Estimated parameters for Schwatz process

To assess the quality of the process, I first run 10,000 simulations of the logarithm of the prices using equation 5.19. In the following, I convert 10,000 simulated paths of the logarithm of the price into level prices.

$$x_t = x_0 e^{-k\Delta t} + (a - \frac{\sigma^2}{2k})(1 - e^{-k\Delta t}) + \sigma\sqrt{\frac{1 - e^{-2k\Delta t}}{2k}} \cdot N(0,1) \tag{5.19}$$

Summary statistics for the simulated prices are given in Table 5.9. The process captures the first four moments rather well, although it overshoots the kurtosis. Also, the process generates some extremely high prices. Histograms of the simulated and observed prices are given in Figure 5.7. For illustration purposes, a sample trajectory of the simulated Schwartz process is given in Figure 5.8.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 10.14 | 35.61 | 39.82 | 135.69 | 17.25 | 1.35 | 5.53 |
| Simulated | 5.25 | 36.59 | 39.73 | 327.57 | 16.73 | 1.36 | 6.49 |

TABLE 5.9: Summary statistics for observed and simulated prices

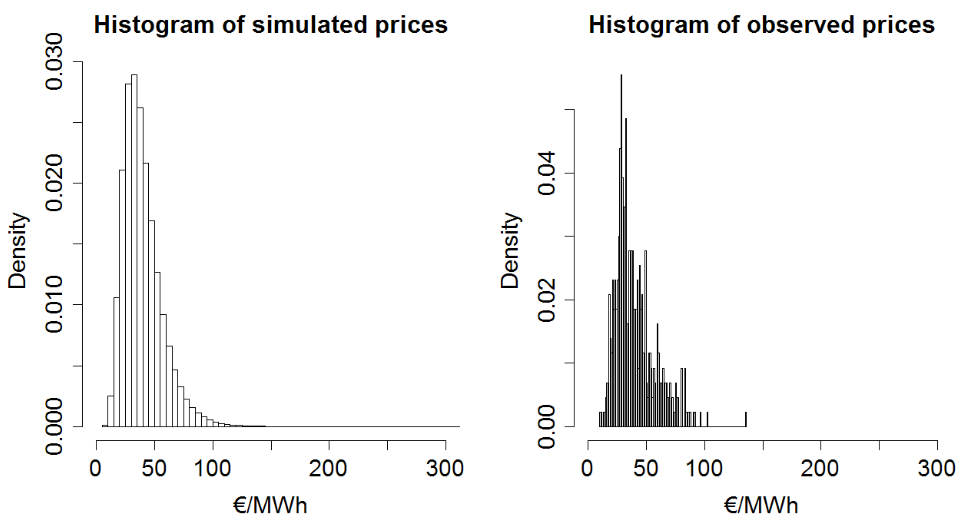



FIGURE 5.7: Histograms of simulated and observed prices for Schwartz process

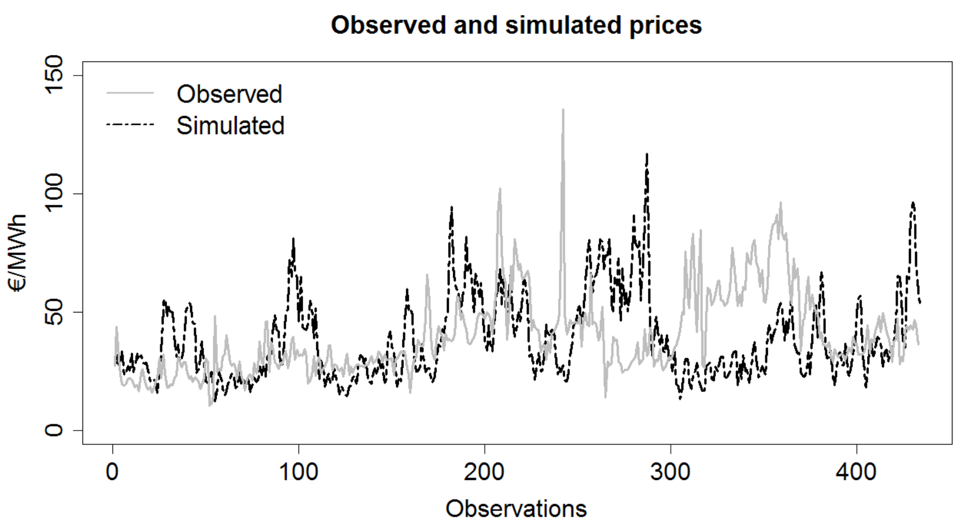



FIGURE 5.8: A simulated path of electricity prices following Schwartz process

Finally, I calculate values of the RMSE and report them in Table 5.10.

| | mean | se. |
|---|---|---|
| RMSE | 23.81 | 2.27 |

TABLE 5.10: RMSE for in sample forecasts based on Schwartz process

### 5.2.4 Selection of stochastic process for modeling electricity prices

#### 5.2.4.1 Comparison based on distributional properties

Table 5.11 provides a summary of the distributional properties of the observed and simulated prices for the OU, CIR and Schwartz processes.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 10.14 | 35.61 | 39.82 | 135.69 | 17.25 | 1.35 | 5.53 |
| OU | -26.89 | 37.90 | 38.12 | 110.43 | 14.99 | 0.11 | 2.97 |
| CIR | 1.13 | 37.76 | 39.77 | 174.59 | 15.50 | 0.79 | 3.93 |
| Schwartz | 5.25 | 36.59 | 39.73 | 327.57 | 16.73 | 1.36 | 6.49 |

TABLE 5.11: Comparison of distributional properties of different stochastic processes

Based on Table 5.11, I compute the absolute difference between observed and simulated prices for each statistic, as shown in Table 5.12.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| OU | 37.03 | 2.29 | 1.7 | 25.26 | 2.26 | 1.24 | 2.56 |
| CIR | 9.01 | 2.15 | 0.05 | 38.90 | 1.75 | 0.56 | 1.60 |
| Schwartz | 4.89 | 0.98 | 0.09 | 191.88 | 0.52 | 0.01 | 0.96 |

TABLE 5.12: Absolute difference between simulated and observed prices

Following, I calculate the deviation in percentage points for each process and for each statistic from the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| OU | 365.2 | 6.4 | 4.3 | 18.6 | 13.1 | 91.9 | 46.3 |
| CIR | 88.9 | 6.0 | 0.1 | 28.7 | 10.1 | 41.5 | 28.9 |
| Schwartz | 48.2 | 2.8 | 0.2 | 141.4 | 3.0 | 0.7 | 17.4 |

TABLE 5.13: Deviation in % from observed data

All processes achieve comparable results for the first two moments: they generate mean and standard errors which are rather close to observed data. In terms of the third and fourth moments, performance is also very similar but still the CIR process performs better than the OU process and the Schwartz process performs better than the CIR process.

Finally, I calculate the ranking of each stochastic process for each statistic. The process that comes closest to the observed data in terms of a particular statistic is assigned the first place. I sum up the positions of each process to arrive at a final ranking for each particular process. The process with the lowest total count in Table 5.14 is considered to be the best performing. This is the case with the Schwartz model which performs best, and it is followed closely by the CIR and then the OU process. Therefore, on the basis of distributional properties and by giving equal weight to each statistic, the Schwartz model would be the first choice in modeling electricity prices followed closely by the CIR and OU processes.

| | min. | median | mean | max. | se. | skew. | kurtosis | total |
|---|---|---|---|---|---|---|---|---|
| OU | 3 | 3 | 3 | 1 | 3 | 3 | 3 | 19 |
| CIR | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 13 |
| Schwartz | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 10 |

TABLE 5.14: Ranking of stochastic processes

#### 5.2.4.2 Comparison based on root mean squared error

Even though the Schwartz one-factor model appears to be the best performing process in terms of distributional properties, the question remains whether this rather informal measure is the appropriate tool for selecting between stochastic processes. Therefore, as a second measure of the goodness of fit, I compute the RMSE for each stochastic process.

Unlike the distributional properties where the Schwartz model performs the best, it is the OU process that performs the best in terms of RMSE. The OU process is followed closely by the CIR and then Schwartz process, although the difference between them is not significant.

| | mean | se. |
|---|---|---|
| OU | 22.93 | 0.40 |
| CIR | 22.98 | 1.88 |
| Schwartz | 23.81 | 2.27 |

TABLE 5.15: Comparison of RMSE for different stochastic processes

In the following, I test whether the difference in RMSE among the three stochastic processes is statistically significant using the Diebold and Mariano (1995) statistic.

**OU versus CIR process** First, I compare the OU and CIR processes by testing the null hypothesis of equal forecast accuracy against the two-sided alternative that these two models produce forecasts of different accuracies. To reject the null hypothesis at the $5\%$ significance level, the DM statistic has to be greater than $1.96$ in absolute value. Using this critical value, I reject the null hypothesis in only $9.1\%$ of all sample paths. This is rather strong evidence that the OU and CIR processes perform equally well. A histogram of DM statistics is given in Figure 5.9.

$$H_0 : \; e^2_{OU,t+1|t} = e^2_{CIR,t+1|t} \tag{5.20a}$$

$$H_1 : \; e^2_{OU,t+1|t} \neq e^2_{CIR,t+1|t} \tag{5.20b}$$

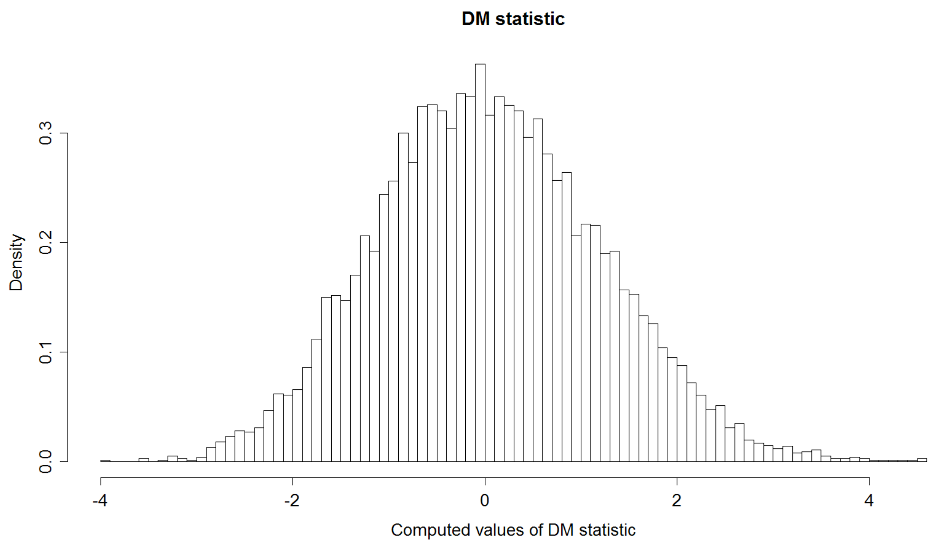



FIGURE 5.9: OU versus CIR - histogram of computed DM statistics

**OU versus Schwartz one-factor process** Next, I compare the forecast accuracy of the OU process against the Schwartz one-factor model. I test the null hypothesis of equal forecast accuracy at the $5\%$ significance level against the two-sided alternative that the forecasts have different accuracies. Here, I reject the null hypothesis in $10.5\%$ of all sample trajectories, which is also rather compelling evidence that neither process dominates. A histogram of computed DM statistics is given in Figure 5.10.

$$H_0: \ e^2_{OU,t+1|t} = e^2_{Sch,t+1|t} \tag{5.21a}$$
$$H_1: \ e^2_{OU,t+1|t} \neq e^2_{Sch,t+1|t} \tag{5.21b}$$

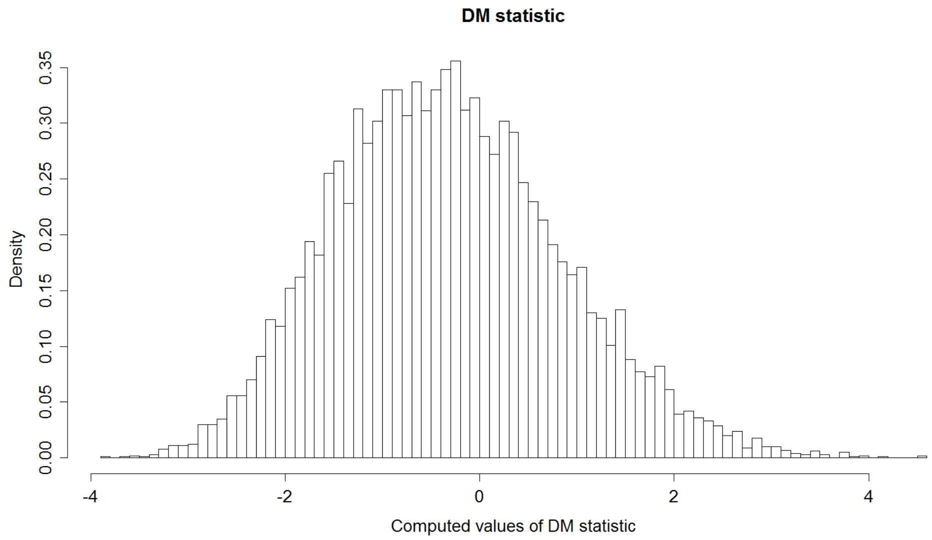



FIGURE 5.10: OU versus Schwartz - histogram of computed DM statistics

**Schwartz one-factor model versus CIR** Finally, I test whether the Schwartz one-factor and CIR models produce forecasts of equal accuracy, against the two-sided alternative that they produce different forecasts. In this case, I reject the null hypothesis in $19.51\%$ of all sample forecasts.

$$H_0: \ e^2_{Sch,t+1|t} = e^2_{CIR,t+1|t} \tag{5.22a}$$
$$H_1: \ e^2_{Sch,t+1|t} \neq e^2_{CIR,t+1|t} \tag{5.22b}$$

Because of this result, I test different specifications of the alternative hypothesis. I test the one-sided alternative that the CIR process generates forecasts of better accuracy. To reject the null hypothesis, the DM statistic has to be greater than $1.65$ which is the critical value at the $95\%$ significance level. In this case, I reject the null hypothesis in $19.05\%$ of all sample paths.

$$H_0: \ e^2_{Sch,t+1|t} = e^2_{CIR,t+1|t} \tag{5.23a}$$
$$H_1: \ e^2_{Sch,t+1|t} > e^2_{CIR,t+1|t} \tag{5.23b}$$

Finally, I test the one-sided alternative hypothesis that the Schwartz process generates better forecast than the CIR process. For this specification, to reject the null hypothesis, the value of the DM statistic has to be smaller than the critical value at the $95\%$ significance level which is $-1.65$. With this specification, I reject the null hypothesis in $9.1\%$ of sample paths. Therefore, I conclude that the CIR process performs equally well as the Schwartz process or a little better. A histogram of calculated DM statistics is given in Figure 5.11.

$$H_0: \; e^2_{Sch,t+1|t} = e^2_{CIR,t+1|t} \tag{5.24a}$$

$$H_1: \; e^2_{Sch,t+1|t} < e^2_{CIR,t+1|t} \tag{5.24b}$$

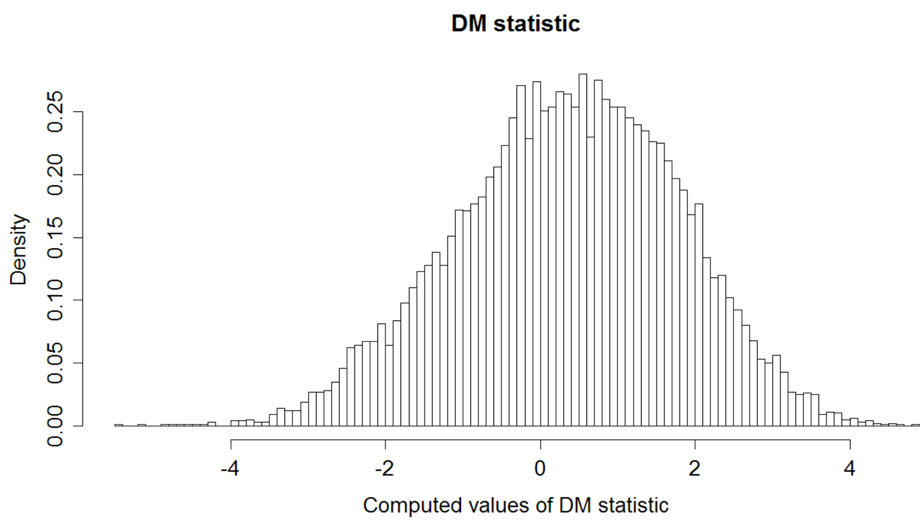



FIGURE 5.11: Schwartz versus CIR - histogram of computed DM statistics

#### 5.2.4.3 Conclusion regarding the selection of the stochastic process

The above analysis has not shown significant statistical difference between the three mean-reverting processes. It appears that all of them produce forecasts of equal accuracy. Nevertheless, I select the OU process over the two other mean-reverting processes. Unlike the Schwartz model, the OU process gives an analytical solution to the project value, something which is not possible with the Schwartz model where the project value has to be calculated numerically. Given that I perform MCLS simulations to determine the optimal investment threshold, this would imply calculating the project value numerically for each sample path and at every stopping time, which represents a great computational challenge. Given that I reject the null hypothesis of equal forecast accuracy between the OU and Schwartz processes in only $10.5\%$ of sample paths, the use of the OU process is justified.

The CIR process on the other hand is difficult to simulate. To simulate a trajectory following the CIR process requires using a chi-squared random variable, which takes longer to simulate than a normal random variable. Again, for computational reasons, I use the OU process. As in the case of the Schwartz process, this is also statistically justifiable, as I reject the null hypothesis of equal forecast accuracy between the CIR and OU processes in only $9.1\%$ of all sample paths.

## 5.3 Coal prices

### 5.3.1 Statistical properties of coal prices

To determine an appropriate stochastic model for coal prices, I use data that span the period from January 4, 2002 until November 13, 2009. Data are obtained from McCloskey Co. and they are compiled on weekly basis. A plot of the level of coal prices together with the corresponding histogram is given in Figure 5.12 while summary statistics for the same period are given in Table 5.16. To model coal prices, I use four processes: Geometric Brownian motion (GBM), OU, CIR and the Schwartz one-factor model.

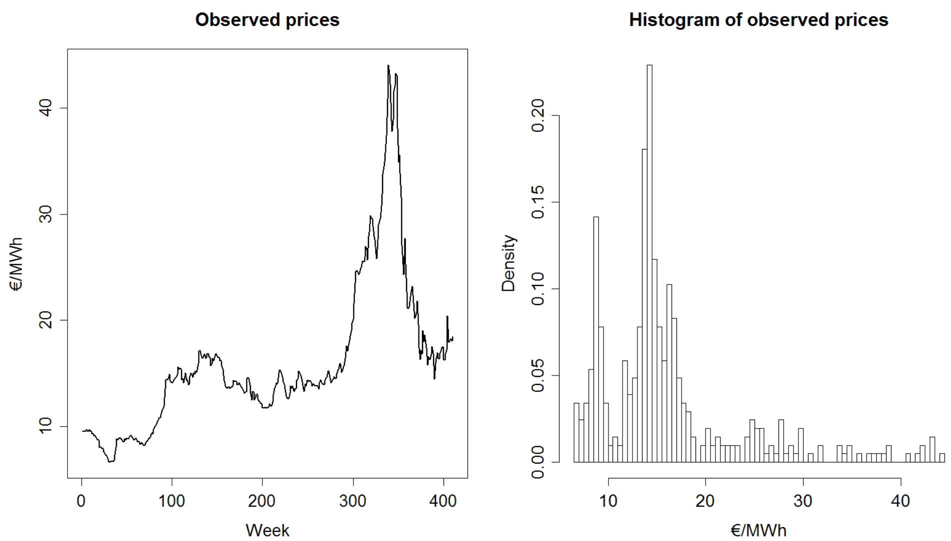



FIGURE 5.12: Weekly coal prices

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Coal prices | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |

TABLE 5.16: Summary statistics for level of coal price

### 5.3.2 Modeling coal prices

#### 5.3.2.1 Geometric Brownian motion

As opposed to the arithmetic version of the same process which I used for modeling dark spread prices, GBM is a strictly non-negative process. Also, the drift and variance of the process are non-constant and depend on price level. The SDE governing the evolution of prices following GBM is given in Equation 5.25.

$$dP = aPdt + \sigma Pdz \tag{5.25}$$

As a first step in the estimation process, I define a new variable $x = lnP$. The dynamics of the new variable $x$ are given in Equation 5.26 (the derivation is shown in Appendix C.1). Given that $x$ appears in the logarithmic form, it follows a Gaussian distribution.

$$dx = (a - \frac{1}{2}\sigma^2)dt + \sigma dz \tag{5.26}$$

Equation 5.26 states that the logarithm of the price follows a random walk with drift. To estimate the parameters, I first determine the expected value and variance of the process in Equation 5.26. They are given by the following two expressions:

$$E[x_t] = x_{t-1} + mt \tag{5.27a}$$
$$Var(x_t) = \sigma^2 t \tag{5.27b}$$

where I define $m = a - \frac{1}{2}\sigma^2$. To estimate the parameters, I follow the ML approach. Given a vector of logarithm of prices $(\boldsymbol{x})$, I estimate the parameters $a, \sigma$ that are part of the parameter vector $\boldsymbol{\theta}$. The log-likelihood function is given by:

$$lnL(\boldsymbol{\theta}|\boldsymbol{x}) = -(n-1) \cdot ln\sqrt{(Var(x))} - \sum_{i=2}^{n} \left[ \frac{(x_i - E(x))^2}{2 \cdot Var(x)} \right] \tag{5.28}$$

Using the expression for the expected value and variance of $x$ from Equations 5.27a and 5.27b, the complete specification of the log-likelihood function becomes:

$$lnL = (\boldsymbol{\theta}|\boldsymbol{x}) = -(n-1) \cdot ln\sigma - \sum_{i=2}^{n} \left[ \frac{(x_i - x_{i-1} - m)^2}{2\sigma^2} \right] \tag{5.29}$$

The values of the estimated parameters are given in Table 5.17.

| | Weekly values |
|---|---|
| Drift ($a$) | 0.0023 (0.002) |
| Volatility ($\sigma$) | 0.0369 (0.001) |

TABLE 5.17: Estimated parameters for GBM process

To test the goodness of fit of the GBM process, I simulate 10,000 trajectories. To simulate a price trajectory following GBM, I use the following exact formula:

$$P_t = P_{t-1} e^{(a-\frac{1}{2}\sigma^2)\Delta t+\sigma\sqrt{\Delta t}\cdot N(0,1)} \tag{5.30}$$

Because I have weekly parameters and as I want to simulate weekly prices, I set $\Delta t = 1$. As a starting point for the simulation, I use the first observed coal price (9.5 €/MWh). Table 5.18 shows the statistical properties of the simulated prices while Figure 5.13 shows a histogram of observed and simulated prices. The simulated trajectories have mean and median values which are close to the observed data. The standard error is also relatively close to the value for the observed data, albeit higher. The skewness of the simulated prices is more than twice the skewness of the observed data, while the kurtosis is roughly seven times greater than the value seen in the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| Simulated | 0.81 | 12.32 | 15.81 | 423.04 | 11.73 | 4.11 | 42.52 |

TABLE 5.18: Summary statistics for observed and simulated coal prices following the GBM process

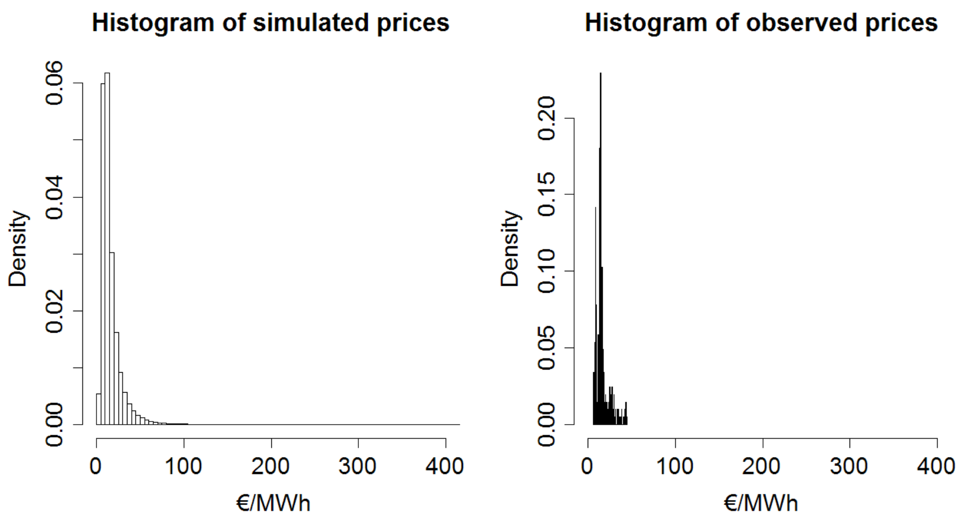



FIGURE 5.13: Histograms of simulated and observed prices for GBM process

As was the case with ABM and dark spread prices, summary statistics of coal prices following GBM are influenced by the starting simulation value. Thus, if as a starting value I use the last observed dark spread price (18.45 €/MWh), summary statistics are given by Table 5.19. As can be seen, changing the starting value has a substantial impact on the simulated prices: now GBM does not come close to any of the distributional properties of coal prices.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| Simulated | 0.81 | 23.93 | 30.71 | 821.78 | 22.79 | 4.11 | 42.52 |

TABLE 5.19: Summary statistics for observed and simulated coal prices following GBM process

For illustration purposes, Figure 5.14 shows a sample trajectory of simulated price using the first observed coal price as a starting value.

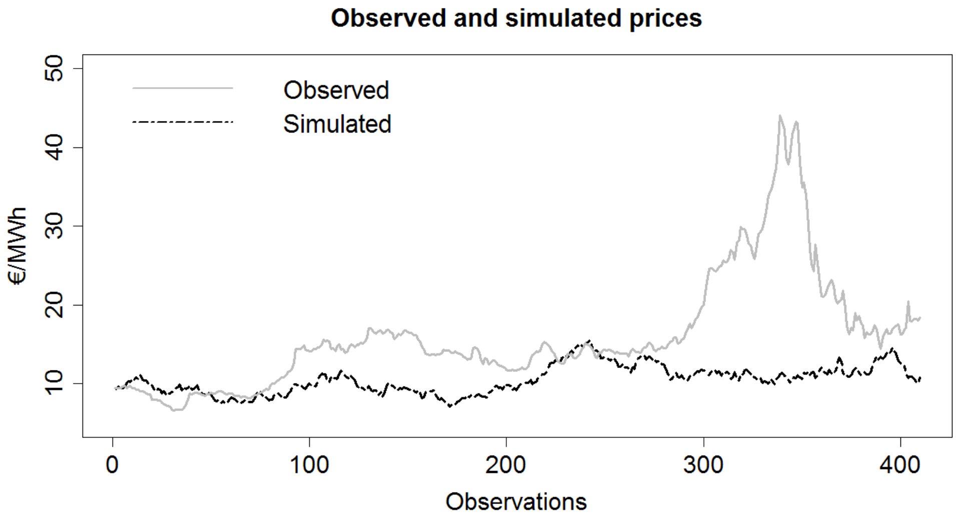



FIGURE 5.14: A simulated path of coal prices following GBM process

To make the process comparable to other stochastic processes, I calculate the RMSE for two different simulations based on two starting values, shown in Table 5.20.

| | mean | se. |
|---|---|---|
| RMSE 1 | 10.42 | 6.32 |
| RMSE 2 | 20.28 | 17.16 |

TABLE 5.20: RMSE for in sample forecasts based on GBM process

#### 5.3.2.2 Ornstein-Uhlenbeck model

There is theoretical and empirical evidence that commodity prices are mean-reverting (Pindyck (1999)). Therefore, as a first mean-reverting model, I fit the OU process to the coal prices. Equation 5.31 gives a SDE governing the evolution of coal prices, where the parameters have the same interpretation as in Equation 5.1. The expected value and variance of the process are given by Equations 5.3a and 5.3b.

$$dp_t = k(\mu - p_t)dt + \sigma dz_t \tag{5.31}$$

Given the vector of coal prices $\boldsymbol{p}$, I estimate a parameter vector $\boldsymbol{\theta}$ consisting of elements $\mu, \sigma, k$ by following the ML approach. The likelihood function is given in Equation 5.5. The values of estimated parameters are given in Table 5.21.

| | Weekly values |
|---|---|
| Long-run price ($\mu$) | 19.1334 (6.14) |
| Mean reversion ($k$) | (0.0071 (0.0056) |
| Standard deviation ($\zeta$) | 0.8113 (0.028) |
| Volatility ($\sigma$) | (0.8142 (0.029) |

TABLE 5.21: Estimated parameters for OU process

To determine the goodness of fit, I perform a MC simulation of 10,000 trajectories. To simulate the trajectories, I use the discretized expression given in Equation 5.7. Summary statistics for the simulated trajectories are given in Table 5.22, while a histogram of simulated prices is given in Figure 5.15.

As opposed to GBM, the OU process generates negative prices. Nevertheless, by looking at the histogram, I see that the frequency of negative values is not large. Overall, the OU process does well in terms of the first two moments. On the other hand, it is unable to come very close to the observed data in terms of skewness and kurtosis.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| Simulated | -15.25 | 15.70 | 16.02 | 46.97 | 6.65 | 0.23 | 2.97 |

TABLE 5.22: Summary statistics for level of observed and simulated coal prices following OU process

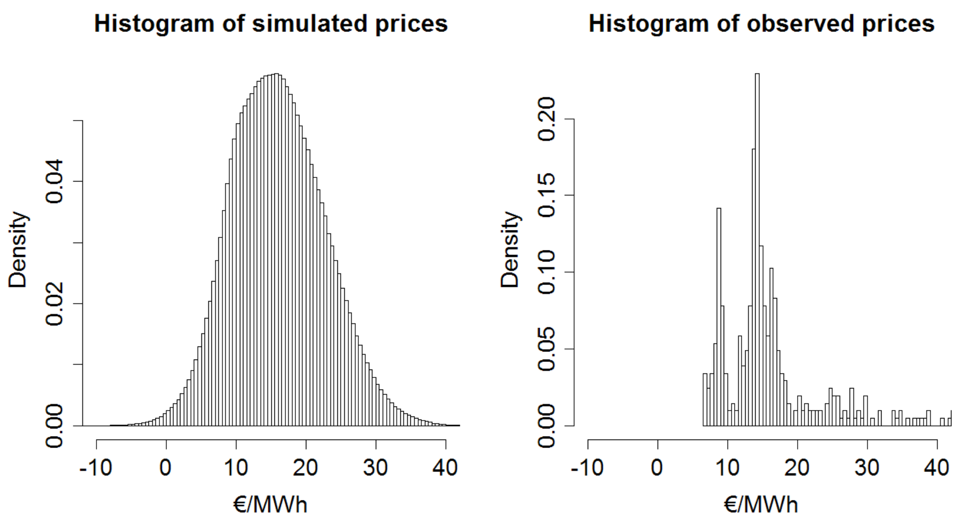



FIGURE 5.15: Histograms of simulated and observed prices for OU process

For illustrative purposes, a plot of the random simulated path of coal prices following the OU process is shown in Figure 5.16.

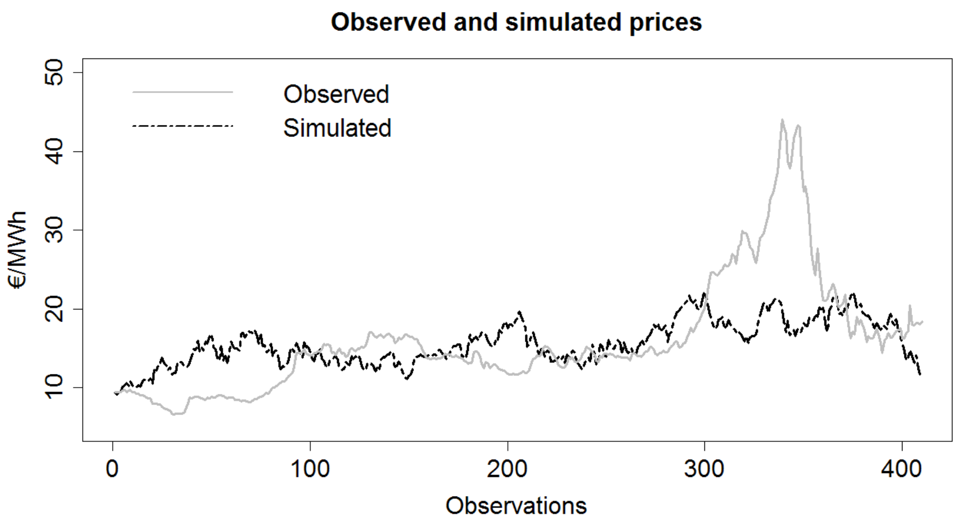



FIGURE 5.16: A simulated path of coal prices following OU process

I also compute the value of the RMSE which is given in Table 5.23.

| | mean | se. |
|---|---|---|
| RMSE | 8.42 | 1.90 |

TABLE 5.23: RMSE for in sample forecasts based on OU process

#### 5.3.2.3 Cox-Ingersoll-Ross model

A third process that I use to model the behavior of coal prices is the CIR model given in Equation 5.32. The properties of the CIR process are given in Section 5.2.3.2.

$$dp_t = k(m - p_t)dt + \sigma\sqrt{p}dz_t \tag{5.32}$$

To estimate the parameters of the CIR process, I use the log-likelihood function given in Equation 5.9. The values of the estimated parameters are displayed in Table 5.24.

| | Weekly values |
|---|---|
| Long-run price ($\mu$) | 19.6565 (7.24) |
| Mean reversion ($k$) | 0.0061 (0.0054) |
| Volatility ($\sigma$) | 0.1659 (0.0058) |

TABLE 5.24: Estimated parameters for CIR process

To test the goodness of fit of the CIR model, I simulate 10,000 trajectories using Equation 5.11. I report summary statistics in Table 5.25 while the corresponding histogram is shown in Figure 5.17. Comparing it to the OU process, CIR performs equally well in terms of mean and standard error. In terms of skewness and kurtosis, it slightly improves over the OU process, but still does not come very close to the observed data.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| Simulated | 2.01 | 15.07 | 15.93 | 54.63 | 5.87 | 0.82 | 3.76 |

TABLE 5.25: Summary statistics for observed and simulated coal prices following CIR process

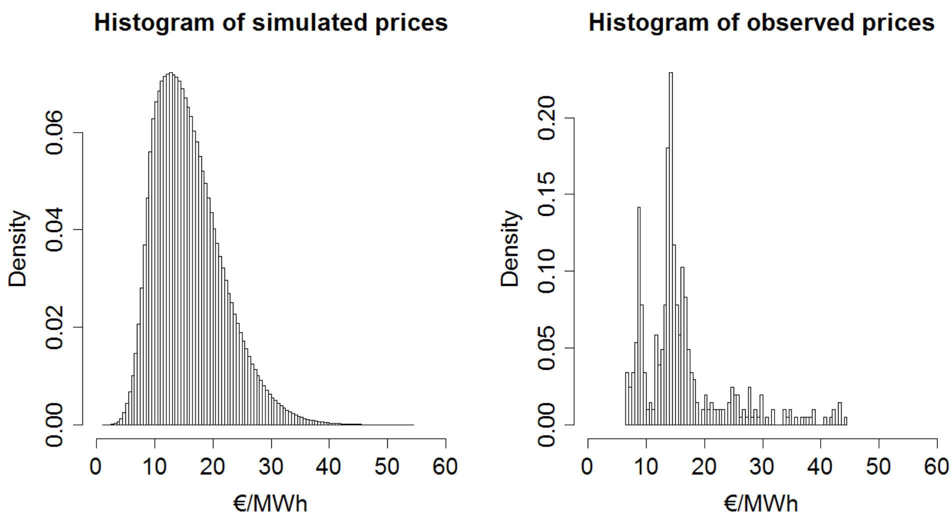



FIGURE 5.17: Histograms of simulated and observed coal prices for CIR process

For illustration purposes, a graph of simulated and observed trajectory is given in Figure 5.18.

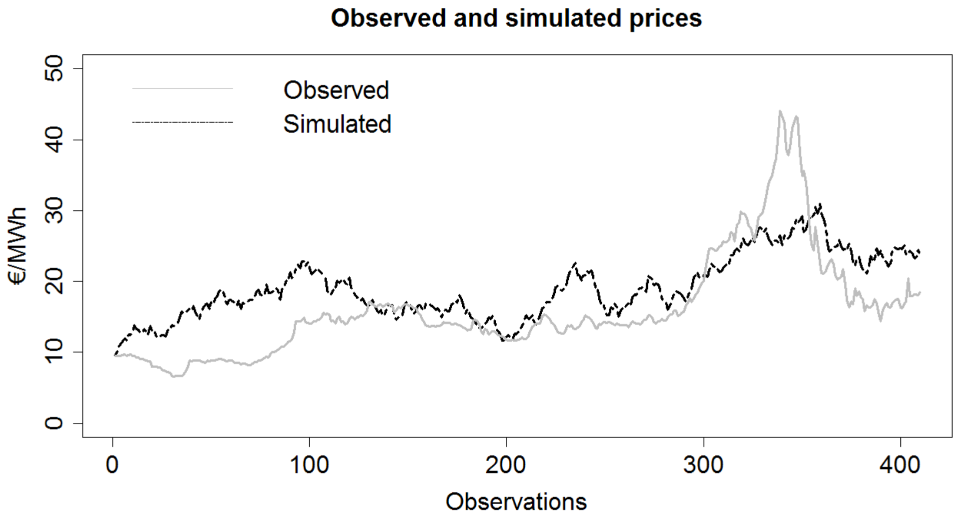



FIGURE 5.18: A simulated path of coal prices following CIR process

Finally, I compute the RMSE as shown in Table 5.26.

| | mean | se. |
|---|---|---|
| RMSE | 7.76 | 1.61 |

TABLE 5.26: RMSE for in sample forecasts based on CIR process

#### 5.3.2.4 Schwartz one-factor model

The last process that I use for modeling the evolution of coal prices is the Schwartz one-factor model. The properties of the model are given in Section 5.2.3.3. To estimate the parameters, I use the expression from Equation 5.17. The estimated parameters are reported in Table 5.27.

| | Weekly values |
|---|---|
| Long-run price ($\mu$) | 3.046 (0.477) |
| Mean reversion ($k$) | 0.0046 (0.0041) |
| Standard deviation ($\zeta$) | 0.0328 (0.0011) |
| Volatility ($\sigma$) | 0.0328 (0.0011) |

TABLE 5.27: Estimated parameters for Schwartz process

In the following, to test the goodness of fit of the process, I simulate 10,000 trajectories using the expression in Equation 5.19. A summary of the distributional properties is given in Table 5.28. The process comes close to the observed data in terms of all four moments. Moreover, it significantly improves over the other processes in terms of the third and fourth moments. Histograms of observed and simulated prices are given in Figure 5.19. Also, for illustration purposes, I show a random simulated trajectory in Figure 5.20.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| Simulated | 3.79 | 14.46 | 15.67 | 93.01 | 5.92 | 1.33 | 5.95 |

TABLE 5.28: Summary statistics for observed and simulated coal prices following Schwartz process

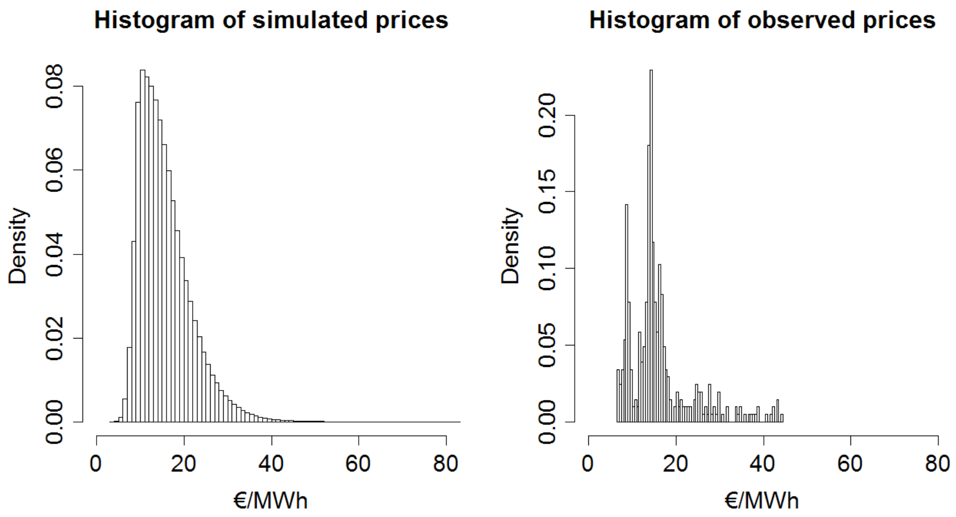



FIGURE 5.19: Histograms of simulated and observed prices for Schwartz process

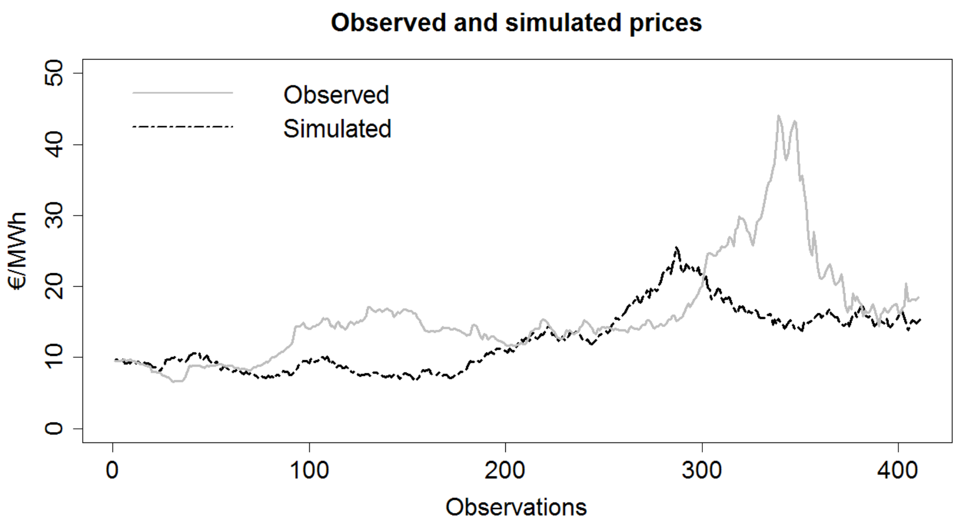



FIGURE 5.20: A simulated path of coal prices following Schwartz process

Table 5.29 shows the RMSE for the Schwartz process.

| | mean | se. |
|---|---|---|
| RMSE | 7.47 | 1.70 |

TABLE 5.29: RMSE for in sample forecasts based on Schwartz process

### 5.3.3 Selection of stochastic process for modeling coal prices

#### 5.3.3.1 Comparison based on distributional properties

Summary statistics for 10,000 simulated trajectories for four different stochastic processes are given in Table 5.30.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| Observed | 6.60 | 14.40 | 16.05 | 44.10 | 7.27 | 1.72 | 6.23 |
| GBM | 0.81 | 12.32 | 15.81 | 423.04 | 11.73 | 4.11 | 42.52 |
| OU | -15.25 | 15.70 | 16.02 | 46.97 | 6.65 | 0.23 | 2.97 |
| CIR | 2.01 | 15.07 | 15.93 | 54.63 | 5.87 | 0.82 | 3.76 |
| Schwartz | 3.79 | 14.46 | 15.67 | 93.01 | 5.92 | 1.33 | 5.95 |

TABLE 5.30: Comparison of distributional properties of different stochastic processes

Based on Table 5.30 I calculate absolute difference between observed and simulated prices which is shown in Table 5.31.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| GBM | 5.79 | 2.08 | 0.24 | 378.94 | 4.46 | 2.39 | 36.29 |
| OU | 21.85 | 1.3 | 0.03 | 2.87 | 0.62 | 1.49 | 3.26 |
| CIR | 4.59 | 0.67 | 0.12 | 10.53 | 1.40 | 0.90 | 2.47 |
| Schwartz | 2.81 | 0.06 | 0.38 | 48.91 | 1.35 | 0.39 | 0.28 |

TABLE 5.31: Absolute difference between simulated and observed prices

Finally, I calculate the deviation in percentage points for each process and for each statistic from the observed data, shown in Table 5.32.

| | min. | median | mean | max. | se. | skew. | kurtosis |
|---|---|---|---|---|---|---|---|
| GBM | 87.7 | 14.4 | 1.5 | 859.3 | 61.3 | 139.0 | 582.5 |
| OU | 331.1 | 9.0 | 0.2 | 6.5 | 8.5 | 86.6 | 52.3 |
| CIR | 69.5 | 4.7 | 0.7 | 23.9 | 19.3 | 52.3 | 39.6 |
| Schwartz | 42.6 | 0.4 | 2.4 | 110.9 | 18.6 | 22.7 | 4.5 |

TABLE 5.32: Deviation in % from observed data

As in the case of electricity prices, the GBM process performs the worst, while the mean-reverting processes perform similarly. Assuming equal weight for each statistic, it turns out that the Schwartz model performs the best, followed closely by the CIR and OU processes, while GBM performs significantly worse.

| | min. | median | mean | max. | se. | skew. | kurtosis | total |
|---|---|---|---|---|---|---|---|---|
| GBM | 3 | 4 | 3 | 4 | 4 | 4 | 4 | 26 |
| OU | 4 | 3 | 1 | 1 | 1 | 3 | 3 | 16 |
| CIR | 2 | 2 | 2 | 2 | 3 | 2 | 2 | 15 |
| Schwartz | 1 | 1 | 4 | 3 | 2 | 1 | 1 | 13 |

TABLE 5.33: Ranking of stochastic processes

#### 5.3.3.2 Comparison based on root mean squared error

As a second criterion used to choose between stochastic models I employ RMSE. Value of RMSE for each model is given in Table 5.34. As in the case of distributional properties, Schwartz model performs the best while GBM performs the worst. Nevertheless, it is questionable whether the difference in RMSE is statistically significant. Therefore, I also compute Diebold and Mariano (1995) (DM) test statistic to determine whether the difference in RMSE between stochastic processes is statistically significant.

| | mean | se. |
|---|---|---|
| ABM | 10.42 | 6.32 |
| OU | 8.42 | 1.90 |
| CIR | 7.76 | 1.61 |
| Schwartz | 7.47 | 1.70 |

TABLE 5.34: Comparison of RMSE for different stochastic processes

General form of DM test is given in Equations 5.33a and 5.33b where $A$ and $B$ are two competing processes. The null hypothesis is that both processes, $A$ and $B$, perform equally well. Alternative, two sided hypothesis is that two competing processes generate forecasts of different accuracy. Testing is performed at $5\%$ significance level which implies that the null is rejected if the absolute value of DM statistic is greater than $1.96$.

$$H_0 : \; e^2_{A,t+1|t} = e^2_{B,t+1|t} \tag{5.33a}$$

$$H_1 : \; e^2_{A,t+1|t} \neq e^2_{B,t+1|t} \tag{5.33b}$$

Table 5.35 shows results of hypothesis testing. Overall, in more than half of all the sample paths I reject the null of equal forecast accuracy between any two models.

| Test | rejected (%) |
|---|---|
| GBM vs OU | 64.8 |
| GBM vs CIR | 66.5 |
| GBM vs Schwartz | 69 |
| OU vs CIR | 60.2 |
| OU vs Schwartz | 59.9 |
| Schwartz vs CIR | 57.4 |

TABLE 5.35: Percentage of sample paths of equal forecast accuracy rejected in favor of two sided alternative of different forecast accuracy

Given that models generate statistically different forecasts, I test whether one model performs better than the others: results are shown in Table 5.36. Interpretation of the table is as follows. I test in what percentage of the sample paths do the models listed in the left column perform better than the models listed in the top row. According to the table, GBM rarely performs better than the three mean reverting process: it performs better than OU, CIR and Schwartz process in $26.6\%$, $19.4\%$ and $16\%$ of all sample paths respectively. Among the mean reverting processes, the Schwartz one-factor model performs relatively better than other mean reverting processes.

| | GBM | OU | CIR | Sch. |
|---|---|---|---|---|
| GBM | 0 | 26.6 | 19.4 | 16 |
| OU | 44.3 | 0 | 23.1 | 17.9 |
| CIR | 52.9 | 43.3 | 0 | 26.2 |
| Sch. | 57.9 | 48.2 | 38 | 0 |

TABLE 5.36: Model comparison (%)

#### 5.3.3.3 Conclusion regarding the selection of the stochastic process

The above analysis showed that the Schwartz one-factor process is the best performing among the four stochastic processes. It performs better than the other two mean reverting processes, albeit not by too much. It also performs better than the GBM process. Nevertheless, I choose the OU process for modeling the evolution of the coal prices. Reasons are similar to the ones used in justifying the use of the OU process in modeling the evolution of electricity prices.

The OU is outperformed by other two mean reverting processes: the CIR and the Schwartz one-factor model. Nevertheless, the advantage of the two mentioned processes is not enough to outweigh the benefits of using the OU process. First, the OU process gives the analytical solution to the investment problem, which is not the case with the best performing Schwartz process, as explained in Section 5.2.4.3. Besides the Schwartz one-factor model, the CIR process also performs marginally better than the OU process. Nevertheless, the use of CIR process requires generating chi-squared random variables which is computationally very demanding. In summary, I choose the OU process because it can reasonably represent the evolution of coal prices and at the same time it is computationally efficient.

## 5.4 Investment analysis

Unlike Chapter 4, where the investment problem was solved using the finite difference method, in this chapter I will solve it using the Monte Carlo least squares (MCLS) algorithm proposed by Longstaff and Schwartz (2001). I will determine the optimal investment decision on a coal-fired power plant whose value depends on the price of coal and electricity.

### 5.4.1 NPV Calculation

As a benchmark, I first determine the optimal investment threshold using NPV. The input parameters used in the calculations are shown in Table 5.37 where as in Chapter 3 I assume power plant can be built instantaneously. According to NPV analysis, the project will remain profitable as long as the price difference between electricity and coal prices remains at 20.6 €/MWh. For instance, if one assumes an electricity price of 42 €/MWh and coal price of 21.4 €/MWh, it makes sense to build the power plant. Therefore, under the NPV framework, the introduction of separate stochastic processes does not alter the investment decision in any way.

| Parameter | Value |
|---|---:|
| Project life (T) | 40 years |
| Capacity | 500 MW |
| Capacity factor % | 85% |
| Annual production (q) | 3,723,000 MWh |
| Investment cost per kW | 1,500 € |
| Discount rate (r) | 10% |

TABLE 5.37: Investment parameters

### 5.4.2 Monte Carlo Least Squares

Previously, it was thought that Monte Carlo (MC) simulations could not handle early exercise problems and thus they were used only to price European options. For instance, in the case of a simple European option on a stock, the value of the option at time $t$ which expires at time $T$ is given by:

$$F(t, p(t)) = E_t\{[e^{-r(T-t)}\ \Pi(p(T), T)\} \tag{5.34}$$

where $\Pi$ denotes a known function of the option payoff at expiration ($T$), $p$ stock price, $r$ the risk-free interest rate, and $E$ denotes the expectation based on information set at time $t$. The basic idea is that the option value is equal to the discounted payoff from exercising the option at maturity if the option is in the money. To estimate the value of the option in Equation 5.34, one needs to simulate a large number of price trajectories ($N$). Thus, the estimated option value ($\hat{F}$) can be calculated as shown in Equation 5.35. Equation 5.35 says that, in order to calculate the value of a European option, one needs to average the discounted payoffs for each simulated path $\omega$.

$$\hat{F}(t, p(t)) = \frac{1}{N}\sum_{\omega=1}^{N} e^{-(T-t)r}\Pi\left(p(\omega, T), T\right) \tag{5.35}$$

Unlike European options which can be exercised only at expiration, American options can be exercised at any time prior to expiration. Therefore, the value of an American option is defined as (Rodrigues and Armada (2006)):

$$F(t, p(t)) = \max_{\tau \leq T}\{E_t[e^{-r(\tau-t)}\Pi(p(\tau), \tau)]\} \tag{5.36}$$

In other words, the value of the option at time $t$ is found by determining the optimal stopping time $\tau \in [t, T]$ when one should exercise the option. The value of the option in Equation 5.36 can be estimated using the Longstaff and Schwartz (2001) algorithm (MCLS).

The mechanism of the MCLS algorithm is as follows. The option life is divided into $N$ discrete points in time when the option can be exercised. As a first step, a large number of trajectories are simulated for each state variable (in this case trajectories of coal and electricity prices). At each of these $N$ time steps, the option holder maximizes the option value by choosing the greater between an immediate exercise or continuation. The problem can be stated in terms of the Bellman equation:

$$F(t_n, p(\omega, t_n)) = \max\{\Pi(t_n, p(\omega, t_n)), e^{-r(t_{n+1}-t_n)}E_{t_n}[F(t_{n+1}, p(\omega, t_{n+1}))]\} \tag{5.37}$$

Equation 5.37 states that, at time $t_n$ and along path $\omega$, one chooses between an immediate exercise which gives payoff $\Pi(t_n, p(\omega, t_n))$ and the decision to wait until the next period $t_{n+1}$ when the value of the option is given by $F(t_{n+1}, p(\omega, t_{n+1}))$. To compare the option value from immediate exercise with the option value from waiting, I discount the expected option value from waiting given by $E_{t_n}[F(t_{n+1}, p(\omega, t_{n+1}))]$ at time $t_{n+1}$ to time $t_n$. I denote the discounted expected value from waiting by $\Lambda$:

$$\Lambda(t_n, p(\omega, t_n)) = e^{-r(t_{n+1}-t_n)} E_{t_n}[F(t_{n+1}, p(\omega, t_{n+1}))] \tag{5.38}$$

The MCLS algorithm starts at expiration when I know that the continuation value is zero, i.e. $\Lambda(T, p(\omega, T)) = 0$. Therefore, at expiration I exercise the option if the exercise value is greater than zero, i.e. if the option is in the money:

$$\text{If} \quad \Pi(T, p(\omega, T)) > 0 \quad \text{exercise the option} \tag{5.39}$$

In the following, I take one step back and repeat the analysis. If, for a given simulated path $\omega$, the continuation value is greater than the exercise value, I should not exercise the option. On the other hand, if the exercise value is greater than the continuation value, I should exercise the option for that path, namely:

$$\text{if} \quad \Lambda(t_n, p_(\omega, t_n)) \leq \Pi(t_n, p(\omega, t_n))] \quad \text{then} \quad \tau(\omega) = t_n \tag{5.40}$$

The value of an American option is then calculated as the average of all optimally exercised discounted cash flows:

$$F(0, p) = \frac{1}{K} \sum_{\omega=1}^{K} e^{-r\tau(\omega)} \Pi(\tau(\omega), p(\omega, \tau)) \tag{5.41}$$

A contribution by Longstaff and Schwartz (2001) is an estimate of the continuation value, which is equal to the expected future cash flow value from the optimal exercise given the information available at the current time. They suggest that the continuation value can be expressed by a countable number ($J$) of basis functions ($B$):

$$\Lambda(t_n, p(\omega, t_n)) = \sum_{j=0}^{J} \lambda_j(t) B_j(t_n, p(\omega, t_n)) \tag{5.42}$$

where $\lambda(t)$ is estimated by the OLS regression proposed by Longstaff and

Schwartz (2001). The estimated continuation value is given by the following expression:

$$\hat{\Lambda}(t_n, p(\omega, t)) = \sum_{j=0}^{J} \hat{\lambda}_j(t) B_j(t_n, p_n(\omega, t)) \tag{5.43}$$

Once the continuation value is estimated, I can choose at each step whether to exercise the option or to wait.

### 5.4.3 Project valuation

The coal power plant does not have any operational options, thus its value is equal to the difference between revenues, which depends on the price of $1$ MWh of electricity ($p$) and the cost of coal required to produce $1$ MWh of electricity ($c$), as given by Equation 5.44.

$$V(p, c) = q \int_0^T E(p_t) e^{-rt} dt - q \int_0^T E(c_t) e^{-rt} dt \tag{5.44}$$

In Equation 5.44, $q$ denotes the annual production[2], $r$ the annual discount rate and $T$ the life of the project. In the previous sections, I determined that the optimal process for modeling electricity and coal prices is the OU process. Therefore, to determine the optimal investment decision, I will assume that both electricity and coal prices follow the OU process. The expected values of electricity and coal prices following the OU process are given in Equations 5.45a and 5.45b, respectively.

$$E[p_t] = p_0 e^{-k_p t} + \mu(1 - e^{-k_p t}) \tag{5.45a}$$

$$E[c_t] = c_0 e^{-k_c t} + \gamma(1 - e^{-k_c t}) \tag{5.45b}$$

Inserting the expressions for the expected prices of electricity ($E(p_t)$) and coal ($E(c_t)$) from Equations 5.45a and 5.45b into Equation 5.44 and integrating, I obtain the value of the project given by Equation 5.46:

$$V(p, c) = q \left( \frac{\left(1 - e^{-(k_p + r)T}\right)(p - \mu)}{k_p + r} - \frac{\left(1 - e^{-(k_c + r)T}\right)(c - \gamma)}{k_c + r} + \frac{\left(1 - e^{-rT}\right)(\mu - \gamma)}{r} \right) \tag{5.46}$$

[2]Actually q represents annual production of electricity as well as amount of coal required to generate this amount of electricity

For completeness, in Table 5.38 I give the parameters describing the evolution of electricity and coal prices.

| | |
|---|---|
| Electricity prices | |
| Long-run price ($\mu$) | 39.9663 |
| Mean reversion ($k_p$) | 0.1793 |
| Volatility ($\sigma_p$) | 10.3230 |
| Coal prices | |
| Long-run price ($\gamma$) | 19.1334 |
| Mean reversion ($k_c$) | 0.0071 |
| Volatility ($\sigma_c$) | 0.8142 |

TABLE 5.38: Parameters for electricity and coal prices following OU process

### 5.4.4 Implementation of MCLS

As a first step in the MCLS model, I simulate 10,000 price trajectories for electricity ($p$) and coal prices ($c$) (5,000 plus 5,000 antithetic). For this purpose, I use the exact expressions given by Equations 5.47a and 5.47b.

$$p_{t+1} = p_t e^{-k_p \Delta t} + \mu(1 - e^{-k_p \Delta t}) + \sigma_p \sqrt{\frac{1 - e^{-2k_p \Delta t}}{2k_p}} \cdot \epsilon_p \tag{5.47a}$$

$$c_{t+1} = c_t e^{-k_c \Delta t} + \gamma(1 - e^{-k_c \Delta t}) + \sigma_c \sqrt{\frac{1 - e^{-2k_c \Delta t}}{2k_c}} \cdot \epsilon_c \tag{5.47b}$$

where $\epsilon \sim N(0,1)$ and $Cov(\epsilon_t, \epsilon_s) = 0$. Once the trajectories have been simulated, I calculate the value of the project ($V(p,c)$) given by Equation 5.46.

The MCLS algorithm starts at the final step, expiration date $T$. At expiration, I determine the value of the option given by $F(p,c) = max[V(p,c) - I, 0]$, where $I$ stands for the investment cost. In other words, the option value is equal to or greater than the intrinsic value (project value minus investment cost) or zero. This well known relation from finance implies that the value of the option cannot be negative. One can always choose not to exercise the option if the payoff from exercising is negative, thus putting a limit on the downside value of the option. Thus If the option value is positive, it is said to be *in the money*, otherwise it is *out of money*. An option holder exercises an option that is in the money and receives the option payoff: otherwise, the option is not exercised. Next, I go back in time to $(T-1)$ where I find the project values that are in the money at this time. Along these paths, I decide whether to exercise the option or to wait for the next period to make the same decision. In other words, I need to compare the value between the immediate exercise and expected future value, i.e. *continuation value*. But the continuation value is not known, thus I need to estimate it. To estimate the continuation value, I regress the discounted option values ($F$) from time $T$ for the paths that are in the money at time ($T-1$) on basis functions of state variables ($p, c$) at time ($T-1$) for the same paths. Formally, denoting by $\omega$ the paths that are in the money at time ($T-1$), $r$ the risk-free interest rate, and by $B_i$ the basis functions where $i = 0...N$, I minimize the following expression to estimate the coefficient vector $\boldsymbol{\lambda}$:

$$\min_{\lambda}\left[F(\omega_T)e^{-r\Delta t} - \sum_{i=1}^{N}\lambda_i B_i(p(\omega_{T-\Delta t}), (c(\omega_{T-\Delta t})))\right]^2 \tag{5.48}$$

Once the optimal coefficients $\hat{\lambda}$ are estimated from Equation 5.48, I calculate the expected continuation value, $\hat{\Lambda}(p, c)$:

$$\hat{\Lambda}(p(\omega_{T-\Delta t}), c(\omega_{T-\Delta t})) = \sum_{i=1}^{N}\hat{\lambda}_i B_i(p(\omega_{T-\Delta t}), (c\omega_{T-\Delta t})) \tag{5.49}$$

For the basis functions, I use simple monomials i.e.:

$$B_i = 1 + \lambda_1 p_i + \lambda_2 p_i^2 + \lambda_3 c_i + \lambda_4 c_i^2 + \lambda_5 p_i c_i \tag{5.50}$$

Now, at time ($T-1$), I exercise the option and receive the payoff $\Pi(p(\omega_{T-\Delta t}), c(\omega_{T-\Delta t}))$ if:

$$\Pi(p(\omega_{T-\Delta t}), c(\omega_{T-\Delta t})) > \hat{\Lambda}(p(\omega_{T-\Delta t}), c(\omega_{T-\Delta t}))$$

I continue in this fashion backward until the beginning of the life of the option, $t = 0$. To obtain the value of the option, I average the payoffs and discount them to the present value. Next, I compare the value of the option with the project value to determine whether to exercise the option. Therefore:

$$\text{if} \quad F(p, c) - V(p, c) + I \leq \epsilon \quad \text{then exercise the option} \tag{5.51}$$

where $\epsilon$ denotes tolerance factor.

### 5.4.5 Results

The investment threshold in the MCLS framework is given in terms of a combination of electricity and coal prices. Needless to say, there are numerous combinations and it is computationally infeasible to come up with a plane linking all possible electricity and coal prices for which one should exercise the option and invest in the project. Instead, for a given level of electricity prices, my approach is to evaluate what is the maximum level of coal prices which can be sustained to make the project profitable. To calculate the optimal investment threshold, I use 10,000 simulations where I assume that the option matures in 3 years. Table 5.39 gives combinations of electricity and coal prices at which one would invest.

| Electricity price €/MWh | Maximum allowed coal price €/MWh | Difference (dark spread) |
|---|---|---|
| 30 | 2.3 | 27.7 |
| 40 | 7.75 | 32.5 |
| 50 | 11.3 | 38.7 |
| 60 | 13.6 | 46.4 |
| 70 | 15.15 | 54.9 |

TABLE 5.39: Combinations of electricity and coal prices at which to invest

## 5.5 Conclusion

In this chapter, I evaluated the investment in a base load coal-fired power plant assuming that plant profitability depends on the prices of electricity and coal. To this end, I modeled separately the evolutions of electricity and coal prices. I determined that the most appropriate process to model both electricity and coal prices is the OU process. Finally, I determined some of the potential electricity and coal price combinations at which it is optimal to invest in a coal-fired power plant.

Comparing the results obtained in this chapter to the result obtained in the previous chapter, I notice that dark spread at which one should invest is not constant. For lower levels of coal prices, dark spread is lower while for higher levels of coal prices it is higher. An explanation for this behavior can be found in the values of half life. The half life of electricity prices is roughly four weeks while, for coal prices, it is close to two years. Therefore, once coal prices wander away from their long-run average, they take a very long time to revert back. Consequently, if the decision to build a coal power plant is made during times of high coal prices, the value of dark spread will also need to be higher.

| | Value (days) |
|---|---|
| Electricity prices | 27 |
| Coal prices | 680 |

TABLE 5.40: Half life for electricity and coal prices in days

# Chapter 6

# Concluding remarks

## 6.1 Contributions and summary of main findings

Investment analysis in energy sector is mostly performed using discounted cash flow (DCF) approach: tool devised several decades ago to value certain cash-flow streams. Researchers and practitioners have been dissatisfied with this approach for a long time because DCF cannot properly deal with the issues of risk and flexibility present in investment projects.

At the end of 1980s researchers have came up with a new method to evaluate investment projects based on options approach, called real options (RO) analysis. RO have so far been applied to analysis of investments in electricity generation, but never to investments in coal fired power plants. Therefore, in this thesis I contribute to the growing knowledge of RO by applying it to investment in coal fired power plant.

The goal of the thesis was to develop a RO model that could be used by energy practitioners in evaluating investments in coal fired power plants in liberalized markets. In the thesis I develop two RO models. First, in Chapter 3 I estimate a one factor dark spread model. I put forward four different stochastic processes among which I select the most appropriate one based on its statistical properties. In Chapter 4 I develop one factor RO model where I calculate investment thresholds for all four stochastic processes estimated in Chapter 3. Finally, in Chapter 5 I develop a two factor RO model. First, I choose the most appropriate process for electricity and coal prices among three and four different stochastic processes respectively. Following, I use MCLS to perform an investment analysis determining the optimal investment threshold.

The research showed that impact of investment irreversibility and stochastic nature of electricity and coal prices results in investment thresholds which are higher than those obtained under traditional capital budgeting techniques. It showed that mean reverting processes should be preferred over Brownian motion. It also showed that two factor model is superior to one factor model. In particular, the following conclusions and contributions arise from the research:

- In Chapter 3 I model evolution of dark spread prices, which according to my knowledge, has not been done before. In this chapter I show that Brownian motion[1] is not an appropriate process for modeling dark spread prices. Among the three mean reverting processes that I also use, CIR model performs the best, while performance of other two mean reverting processes (OU and Schwartz one-factor) is slightly less satisfactory.
- RO have been so far applied to evaluation of investments in gas fired, nuclear and hydro power plants (see literature review section 4.2 in Chapter 4). In Chapter 4 a RO model for evaluation of investments in coal fired power plants is developed for the first time. In developing RO models authors generally assume a certain stochastic process without giving much explanation regarding the choice of the process. It is common to use two processes: Brownian motion (e.g. Takizawa et al. (2001), Venetsanos et al. (2002), Gollier et al. (2005), Kjaerland (2007), Bockman et al. (2008)) or some kind of mean reverting process (e.g. Abadie and Chamorro (2009), Nasakkala and Fleten (2005)). In Chapter 4, for the first time impact of different stochastic processes on investment threshold is assessed. This chapter shows that use of Brownian motion is not acceptable for evaluating investments in coal fired power plants despite the fact that it is commonly used in the literature. It also shows that various mean reverting processes result in very similar investment thresholds.
- Following the development of one factor RO model, in Chapter 5 I develop a two factor model. According to my knowledge this is the first time that two factor RO model of investment in coal fired power plant is developed. In this chapter I show that use of separate electricity and coal prices significantly affects the investment threshold. Due

[1]Besides ABM, GBM was also estimated but the results were more disappointing than in the case of ABM.

to different half lives of electricity and coal prices, higher coal prices require higher dark spread prices to make the investment profitable.

## 6.2 Limitations and future research

The research presented in the thesis can be further extended. In the following I show some of the limitations of the research and indicate the potential avenues for future research:

- RO analysis results in a single threshold value at which one should invest, which might be interpreted as that there is a perfect certainty regarding this value. At the same time, there is a certain degree of uncertainty surrounding the parameters used in calculation of the threshold value evident in the standard error of the estimates. An interesting line of future research might be in determining confidence intervals for the investment threshold values[2].
- In Chapter 5 I modeled separately evolution of electricity and coal prices, assuming these processes are not correlated. An interesting extension of the model would be to introduce correlation between the processes where correlation factor would be estimated from observed data.
- In Chapter 3 the econometric analysis showed that optimal process for modeling of dark spread prices is CIR process. On the other hand, in Chapter 5 I conclude that for modeling of electricity prices it does not make much difference what mean reverting process is chosen. For coal prices the optimal process seemed to be Schwartz one-factor model. This result was not expected: rather it was expected that CIR would be preferred process at least in modeling one of the state variables in Chapter 5. Therefore, future research should include analysis of the aggregation issues in Chapter 3 to analyze how come CIR process is the optimal process for dark spread prices but not for either electricity or coal prices separately.
- Two factor model in Chapter 5 could be further extended to include separate modeling of long term level in electricity and coal prices. This would allow long term price levels to vary over time, which is more realistic assumption than to take them to be fixed values.

[2]This idea was initially suggested by Jean-Yves Pitarakis and again brought to my attention by the examiners Christian Schluter and Richard Green.

- One of the major concerns facing investors in coal fired power plants relates to the treatment of emission risk. Currently, there are no economically viable technologies that can deal with this issue. Only carbon capture and storage technology might be considered a potential candidate, but it is still far from commercial use. Therefore, future research could include an extra state variable in the RO model: price of emission permits.

- Empirical evidence on the validity of RO is sparse, especially when compared to the amount of work in applying RO to various problems. Thus, an interesting experiment would be to asses the profitability of a new plant using the investment threshold and price dynamics estimated in the thesis[3].

[3]This idea was suggested by the external examiner Richard Green.

# Appendices

# Appendix A

# Appendix to Chapter 3

## A.1 Maximum likelihood estimation

Assume a random variable $\boldsymbol{y}$ with $N$ independently distributed observations $(y_1, y_2...y_N)$. Define a probability density function of the random variable conditional on a vector of population parameters $\boldsymbol{\theta}$ (Greene (2002)):

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = f(y_1...y_N|\boldsymbol{\theta})$$

The likelihood function ($L$) is given as:

$$L(\boldsymbol{\theta}|\boldsymbol{y}) = f(y_1...y_N|\boldsymbol{\theta}) = \prod_{i=1}^{N} f(y_i|\boldsymbol{\theta}) \tag{A.1}$$

Unlike the probability density function which is written as a function of data conditional on population parameters, the likelihood function is written as a function of parameters conditional on observed data. This is so as to emphasize that we try to find the parameter vector $\boldsymbol{\theta}$ which is most likely to have generated the observed data. Or, in other words, we try to find $\boldsymbol{\theta}$ which maximizes Equation A.1.

Equation A.1 says that the likelihood function is given as a product of density functions of individual observations. This is possible as we assume that the observed variables are independent. Therefore, the likelihood of observing a sample is equal to the joint probability of observing each individual observation. Because it is easier to work with sums than with products, we take the logarithm of the likelihood function to obtain the log-likelihood function ($lnL$) given by the following equation:

$$lnL(\boldsymbol{\theta}|\boldsymbol{y}) = \sum_{i=1}^{N} lnf(y_i|\boldsymbol{\theta}) \tag{A.2}$$

To illustrate the workings of the maximum likelihood approach, assume a two variable model $y_i = \beta_1 + \beta_2 x_i + \epsilon_i$, where $\boldsymbol{y}$ is independently and normally distributed with mean $\mu = \beta_1 + \beta_2 x_i$ and variance $\sigma^2$.

Based on the observed data, we try to find the value of sample parameter $\sigma$, $\beta_1$ and $\beta_2$, which are considered to be elements of the parameter vector $\theta$. The log-likelihood function is given by the following expression:

$$lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x}) = \sum_{i=1}^{n} lnf(y_i|\boldsymbol{\theta}, x_i) \tag{A.3a}$$

$$lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x}) = \sum_{i=1}^{n} ln\left[\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y_i-\mu)^2}{\sigma^2}\right)\right] \tag{A.3b}$$

$$lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x}) = -\frac{n}{2}ln(2\pi) - nln(\sigma) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{(y_i-(\beta_1+\beta_2 x_i))^2}{\sigma^2}\right) \tag{A.3c}$$

The goal of a maximum likelihood estimation is to find the value of $\boldsymbol{\theta} = \beta_1, \beta_2, \sigma$ that maximizes equation A.3c: this estimate is called the maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$. The value that maximizes the log-likelihood function also maximizes the likelihood function because the logarithm is a monotonic function.

The value that maximizes the log-likelihood function is found by differentiating the log-likelihood function with respect to parameter vector $\boldsymbol{\theta}$ and is known as a necessary condition:

$$\frac{\partial lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x})}{\boldsymbol{\theta}} = 0 \tag{A.4}$$

In the above example, the MLE estimate of $\boldsymbol{\theta}$ is calculated in the following way:

$$\frac{\partial lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x})}{\beta_1} = 0 \qquad \frac{\partial lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x})}{\beta_2} = 0 \qquad \frac{\partial lnL(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{x})}{\sigma} = 0 \tag{A.5}$$

## A.2 Root Mean Squared Error

Denote with $y_{t+1}$ the actual value of the time series and with $\hat{y}_{t+1|t}$ the predicted value of the time series. Error associated with each prediction is denoted by $\epsilon_{t+1|t} = \hat{y}_{t+1|t} - y_{t+1}$. The root mean squared error (RMSE) is calculated using the expression from Equation A.6, where $T$ denotes the number of observations.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} \epsilon_{t+1|t}}{T}} \tag{A.6}$$

## A.3 Diebold-Mariano test statistic

The steps in calculation of Diebold-Mariano statistic are as follows:

1. Calculate the forecast error associated with each model ($A$ and $B$):

$$e_{A,t+1|t} = \hat{y}_{A,t+1|t} - y_{t+1} \tag{A.7a}$$
$$e_{B,t+1|t} = \hat{y}_{B,t+1|t} - y_{t+1} \tag{A.7b}$$

2. Define a variable $d_t$ such that: $d_t = e^2_{A,t+1|t} - e^2_{B,t+1|t}$. The value of Diebold-Mariano statistic (DM) is given by the following expression:

$$DM = \frac{\bar{d}}{\sqrt{Var(d)}} \tag{A.8}$$

   To determine the variance of $d$ ($Var(d)$) I regress $d$ on a constant and compute the Newey West robust standard error. To determine the number of lags ($L$) I use the following rule (Greene (2002)): $L = T^{0.25}$, where $T$ denotes the number of observations. The number of lags is rounded to the nearest higher integer. For instance, if $L = 4.3$, value of $L = 5$ is used. For the numerator I use the average of $d$, namely $\bar{d} = \sum_i^T d_i/n$.

3. For each simulated sample path, I test the following hypothesis:

$$H_0 : \ d_t = 0 \tag{A.9a}$$
$$H_1 : \ d_t \neq 0 \quad \text{or } H_1 : \ d_t > 0 \quad \text{or } H_1 : \ d_t < 0 \tag{A.9b}$$

4. Following, I determine the paths for which the DM statistic is above or below the critical value.

5. Finally, I report the percentage of sample paths for which I can reject the null in favor of the alternative hypothesis.

# Appendix B

# Appendix to Chapter 4

## B.1 Finite differences

In order to determine the optimal investment threshold within a RO framework, one generally has to solve a differential equation. The complexity of the differential equation depends on the choice of stochastic process and on the number of state variables. In the case of a single stochastic variable, and only for a few stochastic processes, analytical solutions are possible. Unfortunately, instances where one can obtain analytical solution are rare and in most situations one has to rely on numerical methods.

To solve a real options problem numerically, one can choose between several methods. In the case of one state variable, the choice is usually between lattice and finite difference methods (FD): the choice depends mostly on preference. In the case of multi-dimensional problems, lattice schemes are generally not very practical because of the curse of dimensionality, and one generally chooses between FD and Monte Carlo least-squares (MCLS) methods. Finally, when the number of dimensions exceeds three (time plus two state variables), FD methods become cumbersome and the most common approach is to use MCLS.

In valuing real options, one is primarily concerned with linear, second order partial differential equations (PDEs) which take the following general form:

$$a(x,y)f_{xx} + 2b(x,y)f_{xy} + c(x,y)f_{yy} + d(x,y)f_x + e(x,y)f_y + g(x,y)f + q = 0 \quad \text{(B.1)}$$

where I have used the following shorthand notation:

$$
\begin{aligned}
f_{xx} &= \frac{\partial^2 f(x,y)}{\partial x^2} & f_x &= \frac{\partial f(x,y)}{\partial x} \\
f_{yy} &= \frac{\partial^2 f(x,y)}{\partial y^2} & f_y &= \frac{\partial f(x,y)}{\partial y} \\
f_{xy} &= \frac{\partial^2 f(x,y)}{\partial x \partial y}
\end{aligned}
\tag{B.2}
$$

Equation B.1 is second order as the highest partial derivative is of second order. It is also linear because coefficients $a$ to $g$ depend only on independent variables $x$ and $y$ and not on the unknown function $f = f(x, y)$ itself.

PDEs are classified into three major categories (Duffy (2006)):

$$\text{elliptic:} \quad b^2 - ac < 0 \tag{B.3a}$$

$$\text{parabolic:} \quad b^2 - ac = 0 \tag{B.3b}$$

$$\text{hyperbolic:} \quad b^2 - ac > 0 \tag{B.3c}$$

For instance, the traditional Black-Scholes PDE has the following form:

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + rS\frac{\partial f}{\partial S} + \frac{\partial f}{\partial t} - rf = 0 \tag{B.4}$$

This equation is of parabolic type because $b = c = 0$ i.e.

$$b = f_{St} = \frac{\partial^2 f(S,t)}{\partial S\,\partial t} = 0 \quad c = f_{tt} = \frac{\partial^2 f(S,t)}{\partial t^2} = 0 \tag{B.5}$$

### B.1.1 Derivation of finite difference scheme

The FD scheme is based on the idea of replacing partial derivatives with approximations based on a Taylor series expansion (Wilmott (1994)).
The Taylor series states that any sufficiently smooth function *f(x)* can be approximated at a point *a* by the following (where the superscript on *f* denotes derivative):

$$f(x) = \frac{f(a)}{0!} + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{B.6}$$

Using the Taylor series, one can approximate the function $f(x \pm h)$ at a point

$x$ by (Brandimarte (2006)):

$$f(x+h) = \frac{f(x)}{0!} + \frac{f^{(1)}(x)}{1!}h + \frac{f^{(2)}(x)}{2!}h^2 + \frac{f^{(3)}(x)}{3!}h^3... \tag{B.7}$$

If we neglect terms of order $h^2$ and higher, then it follows from Equation B.7:

$$f^1(x) = \frac{f(x+h) - f(x)}{h} + O(h) \tag{B.8}$$

where $O(h)$ denotes the truncation error. Equation B.8 actually represents an approximation of the first derivative around point $x$ and is also known as the *forward difference*. Using the same approach as in Equation B.7, one can also write:

$$f(x-h) = \frac{f(x)}{0!} - \frac{f^{(1)}(x)}{1!}h + \frac{f^{(2)}(x)}{2!}h^2 - \frac{f^{(3)}(x)}{3!}h^3... \tag{B.9}$$

Again, I neglect terms of order $h^2$ and higher to get:

$$f^1(x) = \frac{f(x) - f(x-h)}{h} + O(h) \tag{B.10}$$

Equation B.10 is known as the *backward difference*. In the case of both forward and backward differences, the truncation error is of order *O(h)*. The smaller the value of $h$, the smaller the truncation error. This also means that the truncation error goes to zero linearly with *h*.

I can also approximate the first derivative by subtracting Equation B.9 from Equation B.7:

$$f(x+h) - f(x-h) = \frac{2f^1(x)}{1!}h + \frac{2f^3}{3!}h^3... \tag{B.11}$$

Re-arranging and ignoring terms of order $h^3$ and higher, I get:

$$f^1(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) \tag{B.12}$$

Equation B.12 is known as the *central or symmetric difference* and has a truncation error of order *O*($h^2$). Because the error is of order $O(h^2)$, it is a better approximation to the first derivative compared to forward and backward differences. In this case, central differences exhibit quadratic convergence: the truncation error goes to zero like $h^2$ as $h$ goes to zero. A graphical interpretation of forward, backward and central differences is given in Figure B.10.

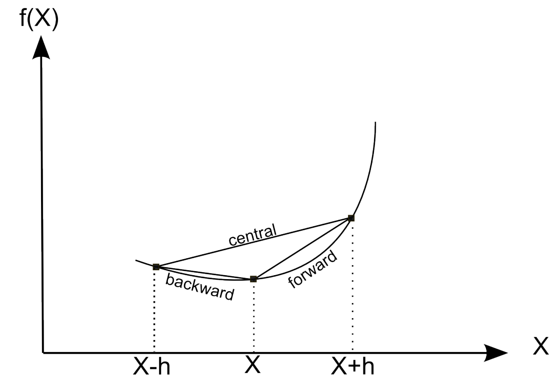



FIGURE B.1: Forward, central and backward difference approximations to point $(x, f(x))$

To obtain an approximation for the second derivative, I add together Equations B.9 and B.7 to get:

$$f(x+h) + f(x-h) = 2f(x) + f^2(x)h^2 + \frac{2f^4}{4!}h^4... \tag{B.13}$$

Re-arranging Equation B.13 I get:

$$f^2(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + O(h^2) \tag{B.14}$$

There are other ways of approximating the second derivative, but this particular scheme is the most stable (Wilmott (1994)).

To use the above formulations of approximations to derivatives to solve a PDE, a discrete grid is formed. The $x$-axis is divided into $m$ equally spaced nodes $dx$ units apart, and the $t$-axis is divided into $n$ equally spaced nodes $dt$ units apart. This in turn divides the plane into a mesh of points $(idx, jdt)$. The value of the function $f(idx, jdt)$ is found at discrete points on the mesh. A graphical representation of the mesh is given in Figure B.2.

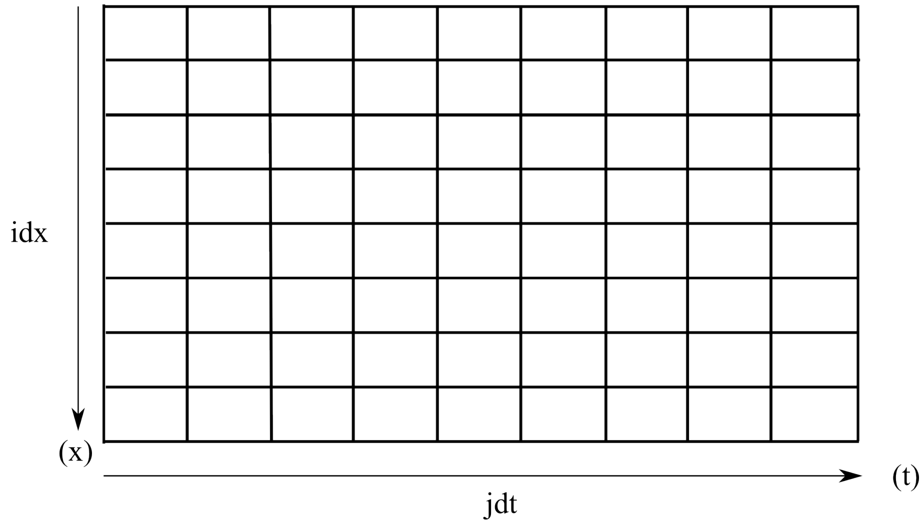



FIGURE B.2: Mesh used in finite difference approximation

Values of a PDE obtained using the mesh are just approximations to the true values, but we hope that this approximation is better the smaller the increments $dx$ and $dt$. There are several ways of solving a two-factor PDE (time plus one stochastic variable), but here I will describe the three most common methods: explicit, implicit and Crank-Nicolson scheme.

Regardless of the approach adopted, the discretization is done in the same manner. Here, I will describe the case of a two-factor PDE which depends on time $(t)$ and price $(p)$, i.e. $f = f(p,t)$. The price is divided into $M$ equally spaced intervals of length $dp$ and the value of the price ranges from $p = 0...pmax$, where $pmax = Mdp$ and $p = idp$. Time is divided into $N$ equally spaced intervals of length $dt$ with the time variable going from $t = 0...T$, where $t = jdt$ and $T = Ndt$. Thus, there are $M+1$ price steps and $N+1$ time steps in total. The value of the option at points $(i,j)$ is $f_{i,j} = f(idp, jdt)$.

To illustrate the use of FDs and the three different schemes, I will use an example that has an analytical solution and compare it to the solution obtained using numerical procedures. For this, I will value a European call option on a non-dividend paying stock. The value of this option is given by the Black-Scholes equation already given in Equation B.4. I will assume the parameters given in Table B.1. Using the analytical solution for the Black-Scholes equation, I get a value of €$12.8545$ for this option.

| | Parameter values |
|---|---|
| Stock price (p) | €60 |
| Strike price (E) | €50 |
| Maturity (T) | 6 months |
| Risk free rate (r) | 10% |
| Standard error p.a. ($\sigma$) | 25% |

TABLE B.1: Parameter values for valuation of European call option

In order to use finite differences and solve a differential equation uniquely, it is necessary to prescribe boundary and terminal or initial conditions. Terminal or initial conditions refer to the behavior of the option at the beginning of its life (initial) or at its expiration (terminal), depending on what convention we use for the passage of time. Furthermore, one needs to describe how the option behaves at some extreme points at all times, which are known as boundary conditions. In solving financial problems like the Black-Scholes equation, boundary conditions describe how the option behaves at some minimum and maximum prices ($p$). Formally, we need to describe the value of the option as $p \to \infty$ and $p \to 0$ if we use a level price or $p \to \infty$ and $p \to -\infty$ if we use the log of price.

If we assume that the option matures at $t = T$, then we know the value of the option at this point in time (expiration), which represents the option payoff (final condition). The value of the call option at expiration is given by:

$$f(idp, T) = max[idp - E, 0] \tag{B.15}$$

In the following, we need to determine the option value at the boundary condition. Because the Black-Scholes equation is based on geometric Brownian motion, we know that if the stock price ever reaches a value of $0$, it will stay there forever. This implies that the option is worthless, giving a lower boundary condition:

$$f(0, t) = 0 \tag{B.16}$$

To determine the value of the option at maximum values of stock price, the upper boundary, I will assume that the option value at maximum price remains constant over time and is equal to the option value at maximum price at expiration (also known as the *Dirichlet* boundary condition), i.e.:

$$f(pmax, t) = f(pmax, T) \tag{B.17}$$

### B.1.2 Explicit finite difference scheme

The explicit finite difference scheme is the most simple and easiest to implement of all three finite difference methods. Figure B.3 gives a graphical representation of the explicit scheme. Starting with the last column, which gives the option payoff at maturity, one calculates the value of the option in the previous time step $(T-1)$ from the option values at time step $T$ (using the option values at points $A, B, C$ at time $T$ to calculate the option value at point $D$ at time $T-1$). This algorithm is carried out until reaching $t=0$.

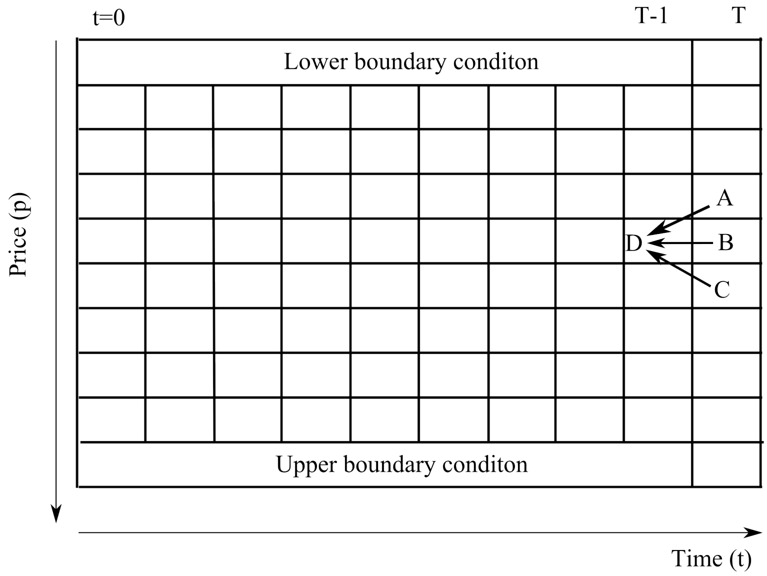



FIGURE B.3: Explicit finite difference scheme

The first step is to approximate the derivatives from equation B.4. For derivatives with respect to price, I will use symmetric differences and backward differences with respect to time:

$$\frac{\partial f(p,t)}{\partial t} = \frac{f_{i,j} - f_{i,j-1}}{dt} \qquad \frac{\partial^2 f(p,t)}{\partial p^2} = \frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{dp^2}$$
$$\frac{\partial f(p,t)}{\partial p} = \frac{f_{i+1,j} - f_{i-1,j}}{2dp} \tag{B.18}$$

Inserting approximations to the derivative from equation B.18 into the Black-Scholes equation B.4 and simplifying, I get:

$$f_{i,j-1} = A_i f_{i-1,j} + B_i f_{i,j} + C_i f_{i+1,j} \tag{B.19}$$

where I define:

$$A_i = \frac{dt}{2}(\sigma^2 i^2 - ri) \qquad B_i = 1 - dt(\sigma^2 i^2 + r) \qquad C_i = \frac{dt}{2}(\sigma^2 i^2 + ri) \tag{B.20}$$

The error of the explicit FD scheme is of the order of $O(dt, dp^2)$. Brennan and Schwartz (1978) point out that for the scheme to be stable, it is necessary that the coefficients $A, B$ and $C$ be positive, which then can be interpreted as probabilities. The fact that the coefficients should be positive puts constraints on the sizes of the $dp$ and $dt$ increments. In practice, the constraint results in taking extremely small increments of $dt$. Formally, in order to make the explicit FD scheme stable, the following restriction must hold (Wilmott (2006)):

$$dt \leq \frac{1}{\sigma^2}\left(\frac{dp}{p_{max}}\right)^2 \tag{B.21}$$

For example, by taking the price step $dp = 1.5€$ and by taking the time step $dt = 1/365$ I obtain the option value of €$12.8452$. Making the grid finer and taking $dp = 0.5€$ and $dt = 1/8760$ gives the value of the option of €$12.8515$, plot of which is shown in Figure B.4.

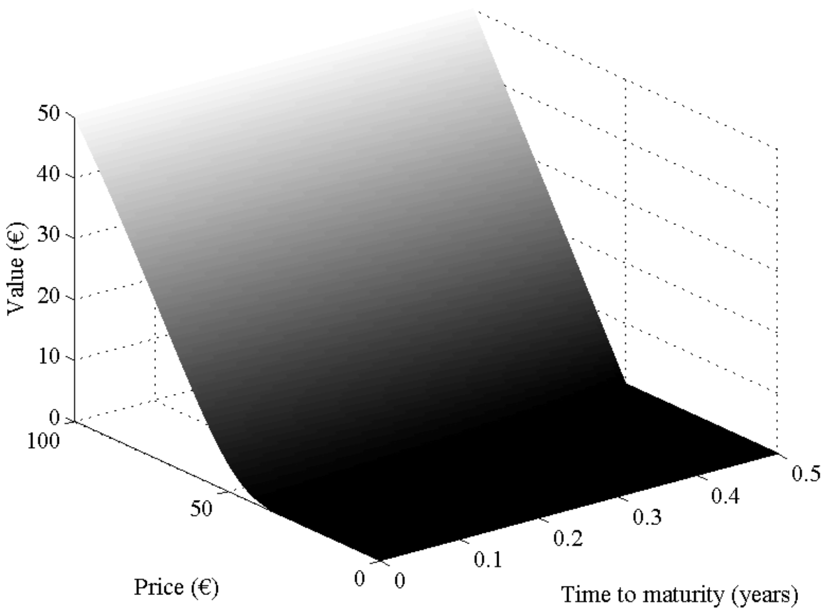



FIGURE B.4: Option value for explicit finite difference scheme

As indicated in Equation B.21, for the explicit scheme to be stable one has to choose the size of the time and price steps carefully. For instance, by taking $dp = 0.5€$, one has to take $dt \leq 3.5/8760$. By using incorrect time step the scheme will give wrong results. For example, by setting $dt = 4/8760$ the meaningless result is obtained: the plot of the option value is given in Figure B.5. Therefore, even though the explicit scheme is simple to use, it's main disadvantage is the limitation on the size of time and price increments.

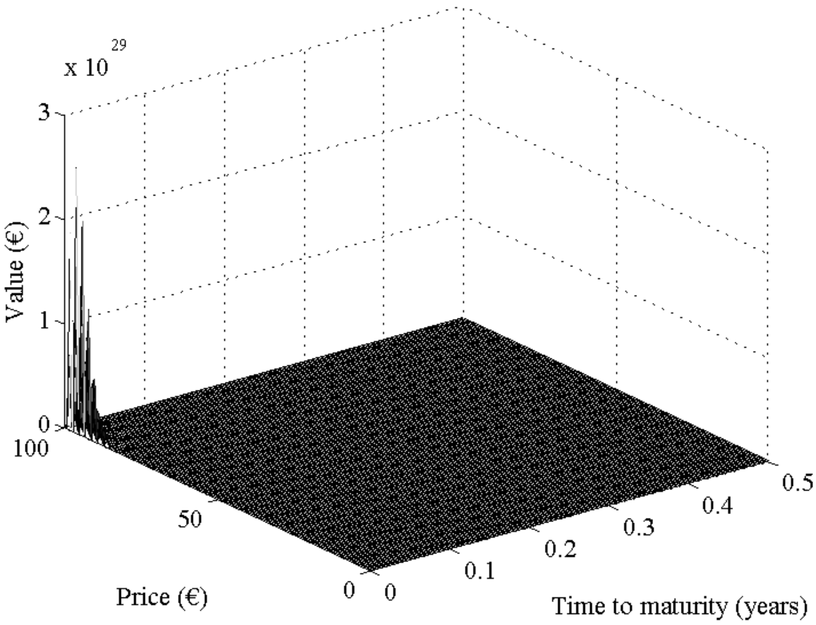



FIGURE B.5: Instability of explicit scheme

### B.1.3 Implicit finite difference scheme

Despite the fact that the explicit method is intuitive and easy to program, its main weakness is instability. The implicit method is used to overcome this limitation: the implicit scheme is unconditionally stable. As was the case with the explicit FD scheme, implicit scheme has error of the order of $O(dt, dp^2)$. Unlike explicit finite differences where the option values can be calculated explicitly, in the case of the implicit method, the option values are calculated implicitly: they are obtained as solutions to a set of simultaneous equations. A graphical representation of the implicit method is given in Figure B.6. As the figure shows, the values of the option at points $B, C, D$ at time $(T-1)$ are obtained simultaneously from the option value at point $A$ at time $(T)$.

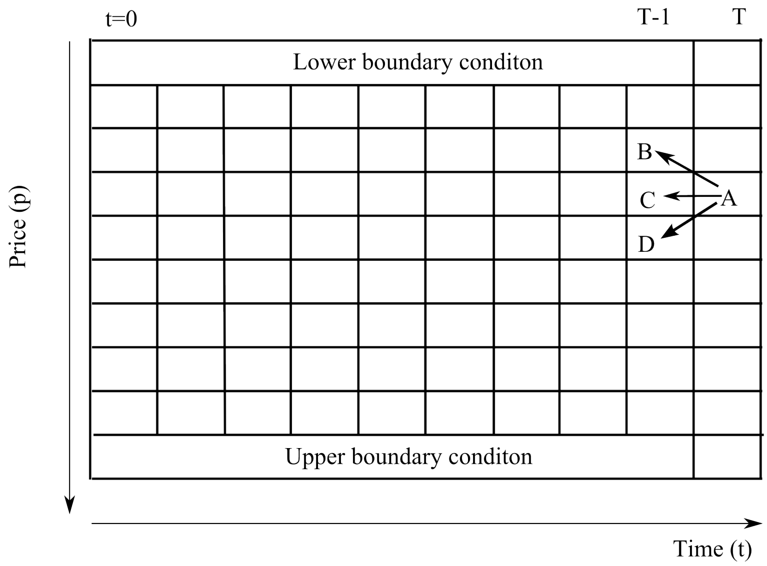



FIGURE B.6: Implicit finite difference scheme

To approximate the derivatives, I use central differences with respect to the price variable and forward differences with respect to time. This gives the following approximation to the Black-Scholes equation:

$$f_{i,j+1} = A_i f_{i-1,j} + B_i f_{i,j} + C_i f_{i+1,j} \tag{B.22}$$

where the coefficients have the following interpretation:

$$A_i = \frac{dt}{2}(ri - \sigma^2 i^2) \qquad B_i = 1 + dt(\sigma^2 i^2 + r) \qquad C_i = \frac{dt}{2}(-ri - \sigma^2 i^2) \tag{B.23}$$

In equation B.22, there are three unknown option values at time $j$ while the option value at time $j+1$ is known from the terminal condition. To obtain the option value at time $j$, I solve the following set of equations simultaneously, i.e (where price ranges from $j = 0...M$):

$$\begin{aligned} f_{1,j+1} &= A_1 f_{0,j} + B_1 f_{1,j} + C_1 f_{2,j} \\ f_{2,j+1} &= A_2 f_{1,j} + B_2 f_{2,j} + C_2 f_{3,j} \\ &\dots \\ f_{M-1,j+1} &= A_{M-1} f_{M-2,j} + B_{M-1} f_{M-1,j} + C_{M-1} f_{M,j} \end{aligned} \tag{B.24}$$

Equation B.24 can be written in matrix form as:

$$\begin{pmatrix} f_{1,j+1} \\ f_{2,j+1} \\ \dots \\ f_{M-1,j+1} \end{pmatrix} = \begin{pmatrix} B_1 & C_1 & \dots & \dots \\ A_2 & B_2 & C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & A_{M-1} & B_{M-1} \end{pmatrix} \cdot \begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \dots \\ f_{M-1,j} \end{pmatrix} + \begin{pmatrix} A_1 \cdot f_{0,j} \\ 0 \\ \dots \\ C_{M-1} \cdot f_{M,j} \end{pmatrix} \tag{B.25}$$

I need to obtain the option value at time $j$, thus the solution is:

$$\begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \dots\dots \\ f_{M-1,j} \end{pmatrix} = \begin{pmatrix} B_1 & C_1 & \dots & \dots \\ A_2 & B_2 & C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & A_{M-1} & B_{M-1} \end{pmatrix}^{-1} \cdot \left[ \begin{pmatrix} f_{1,j+1} \\ f_{2,j+1} \\ \dots \\ f_{M-1,j+1} \end{pmatrix} - \begin{pmatrix} A_1 \cdot f_{0,j} \\ 0 \\ \dots \\ C_{M-1} \cdot f_{M,j} \end{pmatrix} \right] \tag{B.26}$$

In practice, one rarely inverts coefficient matrix from equation B.26 because coefficient matrix is a tridiagonal matrix. This means that except for the elements on the diagonal, above and below the diagonal, matrix is empty, i.e. has zero entries. Inverting such a matrix will create a non tridiagonal matrix which in turn will require more memory [1]. Instead of inverting the matrix one can use more efficient procedures such as LU decomposition (Wilmott (1994)).

Using $dt = 1/365$ and $dp = 1.5$ I get an option value equal to €$12.8343$. By making the grid finer and assuming $dt = 1/8760$ and $dp = 0.5$, the value of the option is €$12.8502$. Refining the grid even more to $dp = 0.1$ gives an option value of €$12.8509$.

### B.1.4 Crank-Nicolson

The third method for valuing options via FDs is the Crank-Nicolson (CN) scheme. Unlike the explicit and implicit method, CN scheme is accurate at the order of $O(dt^2, dp^2)$: this implies that larger time steps can be used and still get accurate results (Wilmott (2006)). CN can be seen as an average of the explicit and implicit schemes:

$$\begin{aligned} \frac{1}{2}\sigma^2 p^2 \frac{f_{i+1,j+1} - 2f_{i,j+1} + f_{i-1,j+1}}{dp^2} + rp\frac{f_{i+1,j+1} - f_{i-1,j+1}}{2dp} + \frac{f_{i,j+1} - f_{i,j}}{dt} - rf_{i,j+1} \\ + \frac{1}{2}\sigma^2 p^2 \frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{dp^2} + rp\frac{f_{i+1,j} - f_{i-1,j}}{2dp} + \frac{f_{i,j+1} - f_{i,j}}{dt} - rf_{i,j} = 0 \end{aligned} \tag{B.27}$$

[1] Memory indeed becomes limiting factor in calculations as problems become more complex. For instance, one number requires 8 bytes of memory. Assume a square matrix of size N. If we invert the matrix, the memory requirement is $N^2$. If we do not invert the matrix, the memory requirement is $3N - 2$. With large matrices, e.g. 10,000 rows and columns this results in a requirement of 800 MB ($10,000^2 \times 8$) if one inverts the matrix and a bit less than 240 KB ($29,998 \times 8$) if one uses only tridiagonal values.

After some manipulation, the CN scheme can be written as:

$$A_i f_{i-1,j+1} + B_i f_{i,j+1} + C_i f_{i+1,j+1} = -A_i f_{i-1,j} + BB_i f_{i,j} - C_i f_{i+1,j} \tag{B.28}$$

where the coefficients $A, B, BB$ and $C$ have the following interpretation:

$$\begin{aligned} A_i &= \frac{dt}{4}(\sigma^2 i^2 - ri) & B_i &= 1 - \frac{dt}{2}(\sigma^2 i^2 + r) \\ BB_i &= 1 + \frac{dt}{2}(\sigma^2 i^2 + r) & C_i &= \frac{dt}{4}(ri + \sigma^2 i^2) \end{aligned} \tag{B.29}$$

Starting at the expiration of the option at time $t = j + 1$, I can express equation B.28 in matrix form as:

$$\begin{aligned} &\begin{pmatrix} B_1 & C_1 & \dots & \dots \\ A_2 & B_2 & C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & A_{M-1} & B_{M-1} \end{pmatrix} \cdot \begin{pmatrix} f_{1,j+1} \\ f_{2,j+1} \\ \dots \\ f_{M-1,j+1} \end{pmatrix} + \begin{pmatrix} A_1 f_{0,j+1} \\ 0 \\ \dots \\ C_{M-1} f_{M,j+1} \end{pmatrix} \\ &= \begin{pmatrix} BB_1 & -C_1 & \dots & \dots \\ -A_2 & BB_2 & -C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & -A_{M-1} & BB_{M-1} \end{pmatrix} \cdot \begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \dots \\ f_{M-1,j} \end{pmatrix} - \begin{pmatrix} A_1 f_{0,j} \\ 0 \\ \dots \\ C_{M-1} f_{M,j} \end{pmatrix} \end{aligned} \tag{B.30}$$

In the case of the Black-Scholes equations, the values of the boundary conditions remain constant i.e. $f(0, j) = f(0, j + 1)$ and $f(M, j) = f(M, j + 1)$, therefore equation B.30 can be simplified to:

$$\begin{aligned} &\begin{pmatrix} B_1 & C_1 & \dots & \dots \\ A_2 & B_2 & C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & A_{M-1} & B_{M-1} \end{pmatrix} \cdot \begin{pmatrix} f_{1,j+1} \\ f_{2,j+1} \\ \dots \\ f_{M-1,j+1} \end{pmatrix} \\ &= \begin{pmatrix} BB_1 & -C_1 & \dots & \dots \\ -A_2 & BB_2 & -C_2 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & -A_{M-1} & BB_{M-1} \end{pmatrix} \cdot \begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \dots \\ f_{M-1,j} \end{pmatrix} - \begin{pmatrix} 2 \cdot A_1 f_{0,j} \\ 0 \\ \dots \\ 2 \cdot C_{M-1} f_{M,j} \end{pmatrix} \end{aligned} \tag{B.31}$$

Equation B.31 can be written more compactly using the following notation:

$$\mathbf{Pmat} \cdot \mathbf{pvec^{j+1}} = \mathbf{Nmat} \cdot \mathbf{pvec^{j}} - \mathbf{2} \cdot \mathbf{bvec} \tag{B.32}$$

Finally, the value of the option at time $t = j$ can be found using:

$$\mathbf{pvec^j} = \mathbf{Nmat}^{-1} \cdot (\mathbf{Pmat} \cdot \mathbf{pvec^{j+1}} + \mathbf{2} \cdot \mathbf{bvec}) \tag{B.33}$$

By taking $dt = 1/365$ and $dp = 1.5$, the value of the option is €$12.8441$. Increasing the number of time steps and assuming $dt = 1/8760$ and $dp = 0.5$, the option value becomes €$12.8515$. Finally, by making the grid even finer and assuming $dp = 0.1$ and $dt = 1/8760$ the value of the option is €$12.8520$.

Table B.2 gives a summary of the option values for the three different methods. For the time step $dt = 1/365$ price increment is taken to be $dp = 1.5$ and for the time step $dp = 1/8760$ price increments are taken to be $dp = 0.5$ and $dp = 0.1$.

| | dt | | |
|---|---|---|---|
| | 1/365 | 1/8760 | 1/8760 & dp=0.1 |
| Explicit | 12.8452 | 12.8515 | NA |
| Implicit | 12.8343 | 12.8502 | 12.8509 |
| Crank-Nicolson | 12.8441 | 12.8515 | 12.8520 |

TABLE B.2: Value of European call option using three different FD schemes (analytical solution €$12.8545$)

## B.2 Half life

The half life shows how long it takes for the state variable to return half way back to its long-run level, when there are no shocks to the system, i.e. no stochastic disturbances (Clewlow and Strickland (2000)). To calculate the half life, I take the deterministic component of equation 3.6:

$$dp = k(m - p)dt \tag{B.34}$$

Equation B.34 can be written as:

$$\frac{dp}{dt} + kpdt = kmdt \tag{B.35}$$

By using integrating factor $e^{\int kdt} = e^{kt}$ (Edwards and Penney (1992)), I get:

$$\frac{dp}{dt}e^{kt} + kpe^{kt}dt = kme^{kt}dt \tag{B.36a}$$

$$\frac{d}{dt}(pe^{kt}) = kme^{kt}dt \tag{B.36b}$$

Integrating equation B.36b, I get:

$$(p_t - m) = e^{-kT}(p_0 - m) \tag{B.37}$$

Denoting the half life as T/2, it must hold:

$$0.5(p_0 - m) = e^{-kT/2}(p_0 - m) \tag{B.38}$$

Taking the natural logarithm and re-organizing the equation B.38, I get a value of half life given by:

$$T/2 = \frac{ln2}{k} \tag{B.39}$$

Appendix C

# Appendix to Chapter 5

## C.1 Derivation of ABM

The stochastic differential equation governing the evolution of a variable following GBM is given by Equation C.1, where $a$ denotes the drift parameter and $\sigma$ the volatility.

$$dP = aPdt + \sigma Pdz \tag{C.1}$$

To estimate the parameters, I first define a new variable $x$ which is defined as the logarithm of price, i.e. $x = lnP$. Using the Ito lemma, I find that the change in $x$ is given by the following expression:

$$dx = \frac{1}{P}dP - \frac{1}{2P^2}dP^2 \tag{C.2}$$

Inserting the value for $dP$ from Equation C.1 into Equation C.2, I obtain the stochastic differential equation describing the evolution of the log of price given in Equation C.3.

$$dx = (a - \frac{1}{2}\sigma^2)dt + \sigma dz \tag{C.3}$$

## C.2 Monte Carlo Least Squares

Valuing real options with one stochastic variable is relatively simple and, as I have shown, it can be done using a finite difference (FD) scheme. Moving to two stochastic variables makes FDs very tedious. Also, in the case of correlated stochastic processes, a FD scheme faces problems with approximating cross-derivatives. With three or more stochastic variables, a FD scheme becomes infeasible. An elegant and efficient way of solving such multidimensional problems consists of using the Monte Carlo Least Squares (MCLS) scheme proposed by Longstaff and Schwartz (2001).

It has been known for some time that simulations can be used to value European options (Boyle (1977)). In the recent years a few authors have proposed simulation-based techniques to value American options (Barraquand and Martineau (1995), Broadie and Glasserman (1997)). But the approach adopted by Longstaff and Schwartz (2001) seems to be as precise as any other approach, with the added benefit of simplicity in terms of calculating the conditional expectation payoff.

In an MCLS setting, I value an American option in a state space $\Omega$, where $\omega \in \Omega$ denotes a sample path on a finite time horizon $[0, T]$ . I assume the option can be exercised at $K$ discrete points in time $0 < t_1 < t_2 < ....t_k = T$. To value a put option ($P$), for instance, I start at the end of the simulated paths at point $t_k = T$, i.e. at maturity where the value of the put option along a single stock price trajectory $\omega$ is given by

$$P(\omega; t_k) = max[E - S(\omega : t_k), 0]$$

Here, $E$ denotes the exercise price and $S$ is the stock price along the $\omega$ price path. For each trajectory, I calculate the option value at maturity. Then, I move backwards to time $t_{k-1}$ and find the price trajectories that are in the money. If the option is in the money, I compare the payoff from an immediate exercise with expected continuation value. The value of an immediate exercise is a simple payoff, $max[E - S(\omega, t_{k-1}), 0]$. The expected continuation value along path $\omega$ at time $t_{k-1}$ given the information set at time $t_{k-1}$ ($I_{k-1}$) and assuming that the option is not exercised until after $t_{k-1}$ is given by:

$$F(\omega, t_{k-1}) = E\left[\sum_{k-1}^{k} exp\left(-\int_{t_{k-1}}^{t_k} r(\omega, s)ds\right) C(\omega, t_k; t_{k-1}, T)|I_{t_{k-1}}\right]$$

where $r$ denotes the discount rate and $C(\omega, t_k; t_{k-1}, T)$ is the option cash flow conditional on the option not being exercised at or prior to $t_{k-1}$ and the option holder following the optimal stopping strategy for $t_k$. Now it becomes optimal to exercise the option as soon as the value of an immediate exercise is equal to or greater than this expected continuation value.

To actually calculate the continuation value $F(\omega, t_{k-1})$, I assume that the functional form of the continuation function can be approximated by a countable linear combination of $B$ basis functions $L_j(S)$ such as simple polynomials or Laguerre, Hermite, Legendre or Chebyshev polynomials and where $S$ denotes an asset following a Markovian process (such as a stock price) underlying the option:

$$F_B(\omega, t_{k-1}) = \sum_{j=0}^{B} a_j L_j(S)$$

The coefficients $a_j$ are determined by regressing the discounted cash flows from time $k$ for price trajectories that are in money at time $k-1$ onto the basis functions $L_j(S)$ using least squares: this gives us the estimated conditional expectation function $\hat{F_B}(\omega, t_{k-1}) = \sum_{j=0}^{B} \hat{a}_j L_j(S)$. Once the conditional expectation function is computed, I can calculate the expected continuation value by using the estimated coefficients of the conditional expectation function and the value of the stock prices that are in the money at time $t_{k-1}$. If the value of the immediate exercise is greater than or equal to the expected continuation value, I immediately exercise the option and receive a payoff $max[E - S(\omega, t_{k-1}), 0]$ and set all subsequent cash flows along the same path to zero. In the case that the continuation value is greater than the exercise value, the cash flows are left unchanged and I move one step backwards. This algorithm is carried out until time $t_1$. Once I have determined the exercise decisions for all price paths, the estimated cash flows are discounted to the present value and averaged across all sample paths $\omega$ to give the value of the option.

In terms of convergence of the algorithm, the authors offer two propositions. The first is that the number of basis functions should be increased up to the point at which the value of the option under the MCLS algorithm no longer changes.

**Proposition 1** *For any finite number of basis functions $B$ and discrete stopping times $K$, and denoting the true value of an American option by $V(S)$ and the discounted cash flows obtained from the MCLS algorithm by LSM, the following relation holds almost surely:*

$$V(S) \geq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} LSM(\omega_i; B, K)$$

***Proof***: *see Longstaff and Schwartz (2001)* ■

The second proposition states that the option value under MCLS can be made arbitrarily close to the true option value by selecting a finite number of basis functions ($B$) and by letting the number of simulated paths ($N$)

approach infinity.

**Proposition 2** *For any $\epsilon > 0$, there exists $B < \infty$ such that*

$$\lim_{N\to\infty} Pr[|V(S) - \frac{1}{N}\sum_{i=1}^{N} LSM(\omega_i; B, K)| > \epsilon] = 0$$

***Proof**: see Longstaff and Schwartz (2001)* ■

## C.3 Monte Carlo Least Squares Example

The example is adopted from Longstaff and Schwartz (2001). To illustrate the workings of MCLS algorithm consider an American put option on a non-dividend paying stock. This option gives the owner the right, but not the obligation to sell the stock at the predetermined price at any time before or at maturity. The option can be exercised at a strike price (E) of $1.1$ at times $t = 1$ to $t = 3$ where time $t = 3$ is the final time representing expiration of the option (after this time the option becomes worthless) and each unit of time represents one year. Assume also that risk-free rate is $r = 6\%$. Simulated sample paths for the stock price are given in Table C.1.

| Path | t = 0 | t = 1 | t = 2 | t = 3 |
|---|---|---|---|---|
| 1 | 1.00 | 1.09 | 1.08 | 1.34 |
| 2 | 1.00 | 1.16 | 1.26 | 1.54 |
| 3 | 1.00 | 1.22 | 1.07 | 1.03 |
| 4 | 1.00 | .93 | .97 | .92 |
| 5 | 1.00 | 1.11 | 1.56 | 1.52 |
| 6 | 1.00 | .76 | .77 | .90 |
| 7 | 1.00 | .92 | .84 | 1.01 |
| 8 | 1.00 | .88 | 1.22 | 1.34 |

TABLE C.1: Stock price paths

The idea of using MCLS is to find stopping times that will maximize the value of the option. The algorithm starts on the last date, time $t = 3$. Conditional that option has not been exercised before $t = 3$, the value of the option along each path realized by the option holder following optimal strategy is given in Table C.2. Basically, if the option is in the money, the option holder exercises the option, otherwise the option is not exercised.

| Path | t = 1 | t = 2 | t = 3 |
|---|---|---|---|
| 1 | - | - | .00 |
| 2 | - | - | .00 |
| 3 | - | - | .07 |
| 4 | - | - | .18 |
| 5 | - | - | .00 |
| 6 | - | - | .20 |
| 7 | - | - | .09 |
| 8 | - | - | .00 |

TABLE C.2: Cash flow matrix at $t = 3$

Now we move one step back to time $t = 2$. At this point in time, the option holder needs to decide whether to exercise the option at time $t = 2$ or to wait until time $t = 3$. From Table C.1 we can see that there are $5$ paths at time $t = 2$ for which the option is in the money. Let $X$ denote these $5$ paths that are in the money at time $t = 2$ and let $Y$ denote discounted value of option payoff at time $t = 3$ for the same paths if the option is not exercised at time $t = 2$. As a result I obtain Table C.3.

| Path | Y | X |
|---|---|---|
| 1 | .00 x .94176 | 1.08 |
| 2 | - | - |
| 3 | .07 x .94176 | 1.07 |
| 4 | .18 x .94176 | .97 |
| 5 | - | - |
| 6 | .20 x .94176 | .77 |
| 7 | .09 x .94176 | .84 |
| 8 | - | - |

TABLE C.3: Regression at $t = 2$

Where $0.94176$ represents discount factor ($e^{-0.06}$). To estimate the expected cash flow from continuing the option life conditional on stock price at time $t = 2$, I regress Y on a constant, $X$ and $X^2$ ($X$ and $X^2$ represent basis functions). The resulting conditional expectation function is $E[Y|X] = -1.070 + 2.983X - 1.813X^2$. Using the expression for conditional expectation I compare the value of the option from immediate exercise (second column in Table C.4) and continuation, i.e. waiting for time $t = 3$ (third column in Table C.4). Value of immediate exercise is calculated as $1.1 - X$ (intrinsic value), whereas the continuation value is obtained by inserting $X$ into conditional expectation function from above ($E[Y|X]$).

| Path | Exercise | Continuation |
|---|---|---|
| 1 | .02 | .0369 |
| 2 | - | - |
| 3 | .03 | .0461 |
| 4 | .13 | .1176 |
| 5 | - | - |
| 6 | .33 | .1520 |
| 7 | .26 | .1565 |
| 8 | - | - |

TABLE C.4: Optimal early exercise decision at time $t = 2$

Comparing the values of immediate exercise and continuation from Table C.4, it appears it is optimal to exercise the option at time $t = 2$ for paths $4, 6$ and $7$. Using this result I obtain the following matrix of cash flows:

| Path | t = 1 | t = 2 | t = 3 |
|---|---|---|---|
| 1 | - | .00 | .00 |
| 2 | - | .00 | .00 |
| 3 | - | .00 | .07 |
| 4 | - | .13 | .00 |
| 5 | - | .00 | .00 |
| 6 | - | .33 | .00 |
| 7 | - | .26 | .00 |
| 8 | - | .00 | .00 |

TABLE C.5: Cash flow matrix at time $2$

It is important to note that once the option is exercised at time $t = 2$, cash flow at time $t = 3$ for the same path becomes zero. This is just the consequence of the fact that option can be exercised only once.

Next, I check whether it is optimal to exercise the option at time $t = 1$. There are five paths for which the option is in the money at time $t = 1$. Again, denote by $X$ option values that are in the money at time $t = 1$ and by $Y$ discounted cash flows for the same paths at time $t = 2$. What is important to notice is that for $Y$ I use actual realized cash flows along each path: I do not use conditional value of $Y$ estimated at time $t = 2$. As noted, there are five paths in the money at time $t = 1$: paths $1, 4, 6, 7$ and $8$. Matrix of current $(X)$ and future discounted cash flows $(Y)$ is given in the following table.

| Path | Y | X |
|---|---|---|
| 1 | .00 x .94176 | 1.09 |
| 2 | - | - |
| 3 | - | - |
| 4 | .13 x .94176 | .93 |
| 5 | - | - |
| 6 | .33 x .94176 | .76 |
| 7 | .26 x .94176 | .92 |
| 8 | .00 x .94176 | .88 |

TABLE C.6: Regression at $t = 1$

The conditional expectation function at time $t = 1$ is estimated by regressing Y on a constant and $X$ and $X^2$, obtaining: $E[Y|X] = 2.038 - 3.335X + 1.356X^2$. By substituting values of X into the equation, I obtain estimated conditional expectation function. These estimated continuation values are then compared to immediate exercise values and if immediate exercise values are greater, the option is exercised, otherwise it is optimal to wait for the next period. Option values from immediate exercise and continuation are given in the following table.

| Path | Exercise | Continuation |
|---|---|---|
| 1 | .01 | .0139 |
| 2 | - | - |
| 3 | - | - |
| 4 | .17 | .1092 |
| 5 | - | - |
| 6 | .34 | .2866 |
| 7 | .18 | .1175 |
| 8 | .22 | .1533 |

TABLE C.7: Optimal early exercise decision at time $t = 1$

Now I have determined optimal exercise strategy for all times $t = 1$ to $t = 3$. In turn I can create an optimal stopping matrix where $1$ is used to denote exercise date.

| Path | t = 1 | t = 2 | t = 3 |
|---|---|---|---|
| 1 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 1 | 0 | 0 |
| 5 | 0 | 0 | 0 |
| 6 | 1 | 0 | 0 |
| 7 | 1 | 0 | 0 |
| 8 | 1 | 0 | 0 |

TABLE C.8: Stopping rule

Having determined the optimal stopping rule, I can now easily calculate cash flow realized from following the rule. This is done by exercising the option at dates where there is one in the matrix in Table C.8, giving rise to the following table:

| Path | t = 1 | t = 2 | t = 3 |
|---|---|---|---|
| 1 | .00 | .00 | .00 |
| 2 | .00 | .00 | .00 |
| 3 | .00 | .00 | .07 |
| 4 | .17 | .00 | .00 |
| 5 | .00 | .00 | .00 |
| 6 | .34 | .00 | .00 |
| 7 | .18 | .00 | .00 |
| 8 | .22 | .00 | .00 |

TABLE C.9: Optimal cash flow matrix

Finally I can calculate the value of the option. This is achieved by discounting back to time zero each cash flow in the above table and averaging over all paths. This in turn gives the value of th option equal to $0.1144$.

## C.4 Antithetic variance

Antithetic variance is one of the procedures used to reduce the variance in a MC simulation. If I have to simulate $N$ sample paths for which I need to generate $N$ random variables, I might generate only $N/2$ random variables and make the other half perfectly negatively correlated. More precisely, denote by $X_1$ the first $N/2$ random variables and by $X_2$ the other $N/2$ random variables and by $Y$ the total number of random variables. If $X_1$ and $X_2$ represent just the first and second half sets of random variables from $Y$, then the variances of these two is equal to the variance of $Y$:

$$Var(Y) = Var(X_1 + X_2) \tag{C.4}$$

But I know that $Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)$. For equation C.4 to hold, the covariance between $X_1$ and $X_2$ has to be zero. But if $Cov(X_1, X_2)$ is negative, then:

$$Var(Y) \geq Var(X_1 + X_2) \tag{C.5}$$

For Equation C.5 to hold it should be the case that $Cov(X_1, X_2) < 0$. Actually, by sampling only $N/2$ random variables (where $X_1 = N/2$) and by making $X_2 = -X_1$, I minimize the variance.

# Bibliography


**Abadie, Luis M. and Jos M. Chamorro**, “Valuing flexibility: The case of an Integrated Gasification Combined Cycle power plant,” *Energy Economics*, 2008, *30*, 1850–1881.

_ **and** _ , “Income risk of EU coal-fired power plants after Kyoto,” *Energy Policy*, 2009, *37* (12), 5304 – 5316.

**Abramowitz, M. and I. A. (Eds.). Stegun**, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, 1972.

**Administration, U.S. Energy Information**, “Annual Energy Outlook 2010 with Projections to 2035,” Technical Report, U.S. Energy Information Administration 2010.

**Aitchison, J. and J.A.C Brown**, *The lognormal distribution*, Cambridge University Press, 1957.

**Amram, Martha and Nalin Kulatilaka**, *Real Options: Managing Strategic Investment in an Uncertain Worldg*, Oxford University Press, 1998.

**Barraquand, Jrme and Didier Martineau**, “Numerical Valuation of High Dimensional Multivariate American Securities,” *The Journal of Financial and Quantitative Analysis*, 1995, *30* (3), 383–405.

**Black, Fischer and Myron Scholes**, “The Pricing of Options and Corporate Liabilities,” *The Journal of Political Economy*, 1973, *81* (3), 637–654.

**Bockman, Thor, Stein-Erik Fleten, Erik Juliussen, Havard J. Langhammer, and Ingemar Revdal**, “Investment timing and optimal capacity choice for small hydropower projects,” *European Journal of Operational Research*, 2008, *190* (1), 255 – 267.

**Boyle, Phelim P.**, “Options: A Monte Carlo approach,” *Journal of Financial Economics*, 1977, *4* (3), 323 – 338.

**Brandimarte, Paolo**, *Numerical Methods in Finance and Economics: A MATLAB-Based Introduction*, Wiley-Interscience, 2006.

**Brennan, Michael J. and Eduardo S. Schwartz**, "Finite difference methods and jump processes arising in the pricing of contingent claims: a synthesis," *Journal of Financial and Quantitative Analysis*, 1978, *13*, 461–474.

_ **and** _ , "Evaluating Natural Resource Investments," *The Journal of Business*, 1985, *58* (2), 135–157.

**Broadie, Mark and Paul Glasserman**, "Pricing American-style securities using simulation," *Journal of Economic Dynamics and Control*, 1997, *21* (8-9), 1323 – 1352.

**Clewlow, Les and Chris Strickland**, *Energy Derivatives: Pricing and Risk Management*, Lacima Publications, 2000.

**Cox, John C., Jonathan E. Ingersoll Jr, and Stephen A. Ross**, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 1985, *53 (2)*, 385–407.

**Diebold, Francis X. and Roberto S. Mariano**, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, July 1995, *13* (3), 253–63.

**Dixit, Avinash**, *Art of Smooth Pasting* Fundamentals of Pure and Applied Economics, Routledge, 1993.

**Dixit, Avinash K. and Robert S. Pindyck**, *Investment under Uncertainty*, Princeton University Press, 1994.

**Duffy, Daniel J.**, *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*, Wiley, 2006.

**Edwards, C. H. and David E. Penney**, *Elementary Differential Equations With Boundary Value Problems*, Prentice Hall, 1992.

**Escribano, Alvaro, Juan Ignacio Pena, and Pablo Villaplana**, "Modeling Electricity Prices: International Evidence," *Working paper*, June 2002.

**Ewald, Christian-Oliver and Wen-Kai Wang**, "Irreversible investment with Cox-Ingersoll-Ross type mean reversion," *Mathematical Social Sciences*, 2010, *In Press, Accepted Manuscript.*

**for the Advancement of Chemical Technology, Texas Institute**, "Texas Gulf coast nuclear feasibility study - final report," Technical Report, Department of Energy 2005.

**Geman, Helyette and Andrea Roncoroni**, "Understanding the Fine Structure of Electricity Prices," *Journal of Business*, 2006, *79 (3)*, 1225–1262.

**Glasserman, Paul**, *Monte Carlo Methods in Financial Engineering*, Vol. 53 of *Stochastic Modelling and Applied Probability*, Springer, 2003.

**Gollier, Christian, David Proult, Franoise Thais, and Gilles Walgenwitz**, "Choice of nuclear power investments under price uncertainty: Valuing modularity," *Energy Economics*, 2005, *27* (4), 667 – 685.

**Graham, John R. and Campbell R. Harvey**, "The theory and practice of corporate finance: evidence from the field," *Journal of Financial Economics*, 2001, *60* (2-3), 187–243.

**Greene, William H.**, *Econometric Analysis*, Prentice Hall, 2002.

**Harvey, David, Stephen Leybourne, and Paul Newbold**, "Testing the equality of prediction mean squared errors," *International Journal of Forecasting*, 1997, *13* (2), 281 – 291.

**Hull, John C.**, *Options, Futures and Other Derivatives*, Prentice Hall, 2005.

**Iacus, Stefano M.**, *Simulation and Inference for Stochastic Differential Equations With R Examples*, Springer, 2008.

**IEA**, "Electricity reform: power generation costs and investment," Technical Report 1999.

_ , "Power Generation: Investment in Electricity Markets," Technical Report 2003.

_ , "Projected Costs of Generating Electricity: 2005 Update," Technical Report, OECD/IEA 2005.

_ , "Tackling Investment Challenges in Power Generation," Technical Report 2007.

**JSOnline**, "Power plant cost to top 1 billion USD," June 2008.

**Kjaerland, Frode**, "A real option analysis of investments in hydropower–The case of Norway," *Energy Policy*, 2007, *35* (11), 5901 – 5908.

**Knittel, Christopher R. and Michael R. Roberts**, "An empirical examination of restructured electricity prices," *Energy Economics*, 2005, *27* (5), 791 – 817.

**Longstaff, Francis A. and Eduardo S. Schwartz**, "Valuing American Options by Simulation: A Simple Least-Squares Approach," *The Review of Financial Studies*, 2001, *14* (1), 113–147.

**Lucia, Julio and Eduardo Schwartz**, "Electricity prices and power derivatives: Evidence from the Nordic Power Exchange," *Review of Derivatives Research*, 2002, *5*, 5–50.

**MIT**, "The Future of Coal  Options for a Carbon Constrained World," Technical Report, MIT 2007.

**Moyen, Nathalie, Margare Slade, and Raman Uppal**, "Valuing risk and flexibility: a comparison of methods," *Resources Policy*, 1996, *22*, 63–74.

**Nasakkala, Erkka and Stein-Erik Fleten**, "Flexibility and technology choice in gas fired power plant investments," *Review of Financial Economics*, 2005, *14* (3-4), 371 – 393. Real Options SI.

**Overbeck, Ludger and Tobias Ryden**, "Estimation in the Cox-Ingersoll-Ross Model," *Econometric Theory*, 1997, *13 (3)*, 430–461.

**Pilipovic, Dragana**, *Energy Risk: Valuing and Managing Energy Derivatives*, McGraw-Hill, 1997.

**Pindyck, Robert S.**, "The Long-Run Evolution of Energy Prices," *The Energy Journal*, 1999, *20 (2)*, 1–27.

**Rodrigues, Artur and Manuel J. Rocha Armada**, "The Valuation of Real Options with the Least Squares Monte Carlo Simulation Method," *University of Minho Working paper*, 2006.

**Ross, Stephen A.**, "Uses, abuses, and alternatives to the net-present-value rule," *Financial Management*, 1995, *24 (3)*, 96102.

**Schwartz, Eduardo and James E. Smith**, "Short-Term Variations and Long-Term Dynamics in Commodity Prices," *Management Science*, 2000, *46 (7)*, 893911.

**Schwartz, Eduardo S.**, "The stochastic behavior of commodity prices: Implications for valuation and hedging," *The Journal of Finance*, 1997, *52(3)*, 923–973.

**Seifert, Jan and Marliese Uhrig-Homburg**, “Modelling jumps in electricity prices: theory and empirical evidence,” *Review of Derivatives Research*, 2007, *10 (1)*, 59–85.

**Takizawa, Shinichiro, Ryota Omori, Atsuyuki Suzuki, and Kiyoshi Ono**, “Analysis of critical electricity price for the investment for constructing a nuclear power plant using real options approach,” *Journal of Nuclear Science and Technology*, 2001, *38, No. 10*, 907–909.

**Venetsanos, Konstantinos, Penelope Angelopoulou, and Theocharis Tsoutsos**, “Renewable energy sources project appraisal under uncertainty: the case of wind energy exploitation within a changing energy market environment,” *Energy Policy*, 2002, *30*, 293–307.

**Weisstein, Eric W.**, “Confluent Hypergeometric Function of the First Kind,” MathWorld-.

**Wilmott, Paul**, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, 1994.

_ , *Paul Wilmott on Quantitative Finance*, John Wiley & Sons, 2006.