Liquidity Risk in the Sovereign Credit Default Swap Market

Author: Rob Sperna Weiland (2520153)

Supervisors: Dr. Peter Spreij (UvA) Dr. Eduard Belitser (VU) Leonie van den Berge, MSc (Rabobank) Dr. Erik Winands (Rabobank)

August 28, 2014

PUBLIC VERSION\textsuperscript{1}

\textsuperscript{1}In this version, some sections and results are altered or deleted due to their confidential nature.
Vrije Universiteit Amsterdam
Faculty of Sciences
Master Stochastics and Financial Mathematics
De Boelelaan 1081a
1081 HV Amsterdam

Rabobank Nederland
Group Risk Management
Risk Model Validation and Methodologies
P.O. Box 17100
3500 HG Utrecht
Preface

This thesis is written in partial fulfillment of the requirements for the degree of Master of Science in Stochastics and Financial Mathematics at VU University Amsterdam. Students can either write their thesis on research conducted at the university or at a company. In the latter case, the company provides a practically relevant problem on which the student can work during a research internship.

Since research in the field of financial mathematics is strongly motivated by practical observations and problems, I decided to write my thesis at a company. Not only the possibility to see what research questions (from a practical point of view) are relevant today, but also the personal experience of working at a company made this an easy choice.

This report is the result of a 6 months research internship at Rabobank, which is one of the largest financial institutions in the Netherlands. Within Rabobank, I was positioned at the department of Risk Model Validation and Methodologies (RMVM), which is responsible for validating all risk models that are used within Rabobank. The project I worked on was on extracting sovereign default probabilities from credit default swap data.

I would like to thank here everyone who helped me during my research. First of all, I would like to thank my colleagues at RMVM, not only for the stimulating and professional work environment, but also for all the fun during the non work related activities. The many drinks, participation in the Dutch FinTech Hackaton and the department’s soccer team are just a few things that come into mind. Although almost everyone helped me one time or another, I would like to thank Leonie van den Berge and Erik Winands in particular for their supervising roles during my internship. Despite their busy schedules, they always had time for my questions and they were always very fast in giving feedback on preliminary results. Furthermore, all practical matters were always perfectly accounted for by them. Of course, I would also like to thank Peter Spreeij from the University of Amsterdam (UvA) for his supervising role. Even though I will not graduate from the UvA, he was still willing to supervise me, and for this I am grateful. I would also like to thank Eduard Belitser from the VU University Amsterdam, who, without hesitation, was willing to act as the second reader of my thesis. Furthermore, I want to thank Koen Visscher for helping me multiple times with programming related issues. Without him, I
would probably still be waiting behind my computer for the first results. Lastly, I would like to thank my family, especially my parents, for their support throughout my study, both morally and financially.

Rob Sperna Weiland
August 28, 2014, Utrecht
Abstract

Due to its large exposures to several sovereigns, Rabobank wants to assess the creditworthiness of these entities. In particular, Rabobank is interested in the sovereign default probabilities. Estimating sovereign default probabilities is however difficult, since, by the lack of historical default events, statistical methods are unreliable.

In order to validate Rabobank’s currently used estimates of sovereign default probabilities, we look at sovereign credit default swaps. A credit default swap (CDS) can be thought of as an insurance contract on default events and therefore the CDS premium should reflect information on the corresponding default probability. Large observed bid-ask spreads indicate, however, that there are liquidity problems in the sovereign CDS market and that therefore the CDS premia are not pure measures of credit risk, but also of liquidity risk.

By introducing an intensity-based model that incorporates liquidity effects, we derive closed-form formulas for the bid and ask premia. By calibrating these formulas to observed bid and ask premia, we are able to decompose the CDS premium into a pure credit part and a liquidity part. Furthermore, we are able to compute the implied default probabilities, since the proposed calibration procedure also generates a time series of the underlying default intensity process.

We tested our model on Turkish and Brazilian credit default swaps using data from the period 01-06-2009 until 28-02-2014. In the case of Brazil, we found that on average 61.5% of the 2 year credit default swap premium could be attributed to credit risk. Furthermore, we computed that the one-year implied default probability was on average 0.27%. In the case of Turkey, we found that only approximately 54% of the 2 year credit default swap premium could be attributed to credit risk. The one-year default probability was, on average, found to be 0.5%. Since the credit default swap premia are higher due to liquidity effects, the sellers receive higher premium payments and therefore we can also conclude that the compensation for liquidity risk in the sovereign CDS market can be attributed to the sellers.

Rabobank’s estimates of the default probabilities of Turkey and Brazil compare well with ours and therefore this strengthens the plausibility of both ours and Rabobank’s estimates. We conclude that the use of credit default swaps in order to extract default probabilities can be a
Abstract

valuable tool for Rabobank to benchmark its internal estimates.
Contents

Preface v

Abstract vii

1 Introduction 1

2 Credit Default Swap Market 5
   2.1 Credit Default Swaps 5
   2.2 Sovereign CDS Market 11
   2.3 Liquidity Issues in the (Sovereign) CDS Market 15
   2.4 Chapter Summary 17

3 Intensity-Based Models 19
   3.1 Definitions and Filtrations 21
   3.2 Construction of the Default Process 25
   3.3 Pricing of Defaultable Contingent Claims 29
   3.4 Change of Probability Measure 35
   3.5 Chapter Summary 45

4 Affine Processes 47
   4.1 Set-Up and Definitions 48
   4.2 Affine Processes and Bond Pricing 51
## CONTENTS

4.3 Existence of Affine Processes .............................................. 54  
4.4 Application to Credit Risk Modeling ..................................... 55  
4.5 Chapter Summary .......................................................... 56  

5 The Credit-Liquidity Model .................................................... 59  
5.1 CDS Liquidity Literature ................................................... 60  
5.2 Model Set-Up ..................................................................... 61  
5.3 How to Use the Model? ....................................................... 71  
5.4 Chapter Summary .......................................................... 78  

6 Data Description and Calibration Procedure .............................. 79  
6.1 Data Description ............................................................ 79  
6.2 Calibration Procedure ..................................................... 84  
6.3 Chapter Summary .......................................................... 89  

7 Results .............................................................................. 91  
7.1 Calibration Results .......................................................... 91  
7.2 Decomposition of the CDS Mid Premium ............................... 103  
7.3 Chapter Summary .......................................................... 104  

8 Maximum Likelihood Estimation of the Process Parameters ......... 107  
8.1 Maximum Likelihood Estimators for Arithmetic Brownian Motion ... 110  
8.2 Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process .... 112  
8.3 Maximum Likelihood Estimators for CIR Process .................... 114  
8.4 Simulation Study ........................................................... 118  
8.5 Maximum Likelihood Results on Intensity Time Series ............. 132  
8.6 Chapter Summary .......................................................... 137  

9 Conclusions .................................................................... 139
**CONTENTS**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A Hazard Process of a Random Time</strong></td>
<td></td>
</tr>
<tr>
<td>A.1 The F-Hazard Process</td>
<td>144</td>
</tr>
<tr>
<td>A.2 Conditional Expectations</td>
<td>145</td>
</tr>
<tr>
<td>A.3 Martingale Results</td>
<td>150</td>
</tr>
<tr>
<td><strong>B Martingale Hazard Process of a Random Time</strong></td>
<td>157</td>
</tr>
<tr>
<td>B.1 The (F, G)-Martingale Hazard Process</td>
<td>157</td>
</tr>
<tr>
<td>B.2 Relationships Between Γ and Λ</td>
<td>161</td>
</tr>
<tr>
<td><strong>C Derivations of the Discount Factors</strong></td>
<td>165</td>
</tr>
<tr>
<td>C.1 Model with Gaussian Processes for Liquidity Intensities</td>
<td>166</td>
</tr>
<tr>
<td>C.2 Model with Ornstein-Uhlenbeck Processes for Liquidity Intensities</td>
<td>168</td>
</tr>
<tr>
<td><strong>D Arithmetic Brownian Motion</strong></td>
<td>171</td>
</tr>
<tr>
<td>D.1 Bond Price under Arithmetic Brownian Motion</td>
<td>172</td>
</tr>
<tr>
<td>D.2 Explicit Solution of the Arithmetic Brownian Motion SDE</td>
<td>173</td>
</tr>
<tr>
<td>D.3 Maximum Likelihood Estimators for Arithmetic Brownian Motion</td>
<td>174</td>
</tr>
<tr>
<td><strong>E Ornstein-Uhlenbeck Process</strong></td>
<td>179</td>
</tr>
<tr>
<td>E.1 Bond Price under Ornstein-Uhlenbeck Process</td>
<td>180</td>
</tr>
<tr>
<td>E.2 Explicit Solution of the Ornstein-Uhlenbeck SDE</td>
<td>183</td>
</tr>
<tr>
<td>E.3 Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process</td>
<td>186</td>
</tr>
<tr>
<td><strong>F CIR Process</strong></td>
<td>193</td>
</tr>
<tr>
<td>F.1 Solution for Riccati Equation With Constant Coefficients</td>
<td>194</td>
</tr>
<tr>
<td>F.2 Bond Price under CIR process</td>
<td>197</td>
</tr>
<tr>
<td>F.3 Maximum Likelihood Estimators for CIR Process</td>
<td>201</td>
</tr>
</tbody>
</table>

References                                                                 | 205  |
1

Introduction

“It is a first priority, then, for future empirical credit research to quantify the effect of nondefault factors [...] and to extract a default intensity function that is “uncontaminated” by these factors.”

– Jarrow, Lando, Yu (2005) [38]

Before the outbreak of the global financial crisis in 2007, large financial institutions and (developed) sovereigns were considered to be practically free of default risk. The phrase “too big to fail” was commonly used to denote the perception that large financial institutions would not be allowed to go into default, since this would cause too much damage to the economy. With the bankruptcy of Lehman Brothers on September 15, 2008, however, this perception changed and the demand for better risk assessments of these ‘low-default’ entities increased.

Rabobank, one of the largest financial institutions in the Netherlands, is also interested in assessing the credit riskiness of these low-default entities. Rabobank is, amongst others, interested in the credit riskiness of sovereigns, which are also typical examples of low-default portfolios. One of the most important measures of sovereign credit risk is the probability of default. This measure is not only used to decide whether or not to invest in a certain sovereign, but it is also used to set the bank’s provisions and, furthermore, it is an important parameter in the computations of economic and regulatory capital requirements.

Of course, Rabobank has estimates of sovereign default probabilities, but, by the lack of historical data on default events, which makes statistical methods and tests are unreliable, it is not clear whether these estimates are correct. In order to validate the currently used estimates, the Risk Model Validation and Methodologies\(^1\) (RMVM) department, which is responsible

\(^1\)Due to an internal reorganization, the department will be renamed to Model Validation (MV) from July 1, 2014.
for validating all risk models within Rabobank, initiated an investigation to the possibility of extracting default probabilities from credit default swap (CDS) data.

A credit default swap is an over-the-counter traded product, which can be viewed as an insurance against default risk on a reference entity. The buyer of such a contract receives a payment from the seller if the reference entity defaults during the lifetime of the contract. In return for this insurance, the protection buyer makes fixed periodic premium payments to the protection seller. These payments are typically made every three months and stop if either a default occurs within the lifespan of the contract or if the contract reaches its maturity. The fixed premium payments are a fractional amount of the face value of the contract and the annualized fraction is called the CDS premium or CDS spread. Similar to interest rate swaps, the CDS premium is such that the present value of the initial contract is the same for the buyer and the seller.

In principle, the level of the CDS premium should reflect the market’s perception of the default risk of the reference entity, but it is highly likely that other ‘distorting’ factors also influence the CDS premium. Some of these possible distorting factors are counterparty credit risk (i.e. the risk that either the protection buyer or seller defaults during the lifetime of the contract), speculation and contagion effects. RMVM conjectures, however, that the largest distorting factor is liquidity risk.

Liquidity risk can be defined as the risk stemming from the lack of marketability of the CDS, meaning that it cannot be bought or sold fast enough to prevent or minimize a loss. A possible cause for liquidity risk in the credit default swap market is the over-the-counter market structure. Because of this market structure, there is not much regulation and transparency. Furthermore, since there are no central clearing houses, market participants incur search costs for finding reliable counterparties. Large observed bid-ask spreads (the differences between the premia that sellers demand and the premia that buyers are willing to give) are clear indicators of liquidity problems in the sovereign CDS market.

In order to get more reliable default probability estimates one needs to filter out the distorting components and figure out what part of the spread can be attributed to default risk. The first goal of our research is, therefore, to investigate the impact of liquidity risk on sovereign CDS premia. We want to be able to decompose the CDS premium into a pure credit part and a liquidity part, since this improves the information on the credit riskiness of the reference entity we can obtain from the CDS premium. The second goal of our research is to actually compute the sovereign default probabilities using this improved information.

Based on earlier research within Rabobank, we propose to use a so-called intensity-based model to extract the default probabilities from CDS premia. Intensity-based models are a class of models that do not try to explain default events by considering economic fundamentals, but instead they assume that every reference entity follows some exogenously specified jump process in which the first jump time of this process is considered to be the default time. The probability that a jump occurs is influenced by the default intensity process, hence the name intensity-based
In order to incorporate liquidity effects, we will introduce extra liquidity discount factors. We will derive closed-form expressions for the CDS bid and the ask premia, and by calibrating these to the observed bid and ask premia, we are able to give a natural decomposition of the CDS mid premium into a credit part and a liquidity part. In order to be able to compute the implied default probabilities, some additional steps have to be taken, since calibrating the model premia to observed premia implies that we work under the risk-neutral measure. By definition of a risk-neutral measure, we know that, in general, the probabilities of events are different under this measure than under the real-world probability measure. Therefore, we need to extract the real-world behavior of the default intensity in order to be able to compute the real-world default probabilities.

Unfortunately, issues related to the change of probability measure are more complicated in intensity-based models than in, for example, standard short-rate interest rate models. In intensity-based models, not only the drift terms of the default intensity process are different under the different measures, also the values of the process itself differ. Rabobank already has a model to compute the differences between the values of the default process under both measures, but it still needs a model to compute the drift term change. Therefore, in order to be able to compute the real-world default probabilities, we need to find the drift term of the default intensity process under the real-world measure and plug this into the model of Rabobank.

The calibration procedure we will propose, generates a time series of the default intensity process from which we are able to obtain the drift term that is required for the use of Rabobank’s model. We will test our model on CDS data on the Turkish and Brazilian governments. We indeed find that large parts of their CDS premia can be attributed to liquidity effects. Furthermore, our estimates of the one-year default probabilities compare well with Rabobank’s estimates, which strengthens the plausibility of both our and their results.

The structure of this report is as follows. In chapter 2, we will give some general background information on credit default swap contracts and the major market developments. We will specifically focus on the sovereign credit default swap market and we will discuss some indicators of liquidity issues in this market. In chapter 3, we will introduce the class of intensity-based models and we will explain how one can use these models to price defaultable claims. Furthermore, we will discuss some issues related to the change of probability measures within these models. In chapter 4, we will introduce the class of affine processes, which is particularly useful for modeling intensity processes within the intensity-based set-up, since processes of this class give rise to analytically tractable expressions.

Having built up all the necessary ingredients, we are able to specify our ‘credit-liquidity’ model in chapter 5. In chapter 6 we will give a description of the data and we will discuss the calibration procedure. The results of the calibration will be stated in chapter 7. In order to be
able to compute the default probabilities of the Turkish and Brazilian governments, we have to perform some additional steps on the calibration output and this will be the topic of chapter 8. In this chapter, we will discuss some maximum likelihood estimators for the drift parameters of the default intensity process, given a discrete time series of the default intensity process. We will first analyze their performance by means of a Monte Carlo simulation study and, after that, we will apply these estimators to the output data of our calibration procedure. By combining these results to the model of Rabobank, we are able to derive the default probabilities of the Turkish and Brazilian governments. In chapter 9 we will state our conclusions and recommendations.
Credit Default Swap Market

Even though the CDS market is moving towards more standardization, a credit default swap is still a very flexible contract and, therefore, we will discuss its main specifications in more detail in this chapter. Furthermore, we will give some background information on the sovereign CDS market and we will discuss some liquidity issues within this market.

2.1 Credit Default Swaps

Since its beginning in the early 1990s, the credit derivatives market has experienced a dramatic development. Figure 2.1 shows, for example, that the gross outstanding notional amount was approximately 200 billion USD in 1997, whereas the gross outstanding notional amount was over 60 trillion USD in 2007. Since the start of the financial crisis, the credit derivatives market has declined to roughly 30 trillion USD. The decline of the credit derivative market likely relates to the lack of adequate derivative counterparty credit risk management and accounting rules, and the fact that exactly this counterparty credit risk largely contributed to the crisis (IMF, 2013 [33]).
Chapter 2. Credit Default Swap Market

Figure 2.1: Gross notional amount outstanding on credit derivatives in trillions of dollars (Kiff et al., 2009 [41])

Following Schönbucher (2003) [52], a credit derivative can be defined as “a derivative security that has a payoff which is conditioned on the occurrence of a credit event. The credit event is defined with respect to a reference credit (or several reference credits), and the reference credit asset(s) issued by the reference credit.” Figure 2.2 gives an overview of the different credit derivative products in the credit derivative market.

![Diagram of credit derivatives market]

Figure 2.2: Overview of the credit derivatives market(Weistroffer, 2009 [59])

Typical credit derivative contracts require the specification of the reference entity, the reference assets, the credit event and the payment in case of a credit event. The International Swap and Derivatives Association (ISDA) provides detailed definitions on credit events and procedures involving credit derivatives. Although credit derivatives are traded over-the-counter (OTC),
2.1. Credit Default Swaps

most contracts follow the rules and guidelines set by the ISDA.

The most commonly traded credit derivative is the credit default swap (CDS). The ISDA defines a credit default swap as “a bilateral agreement to shift credit risk between two parties”. The reference assets in a CDS contract are typically loans or bonds issued by a corporate or sovereign entity. The most common CDS contracts can be divided into two categories: the single-name CDSs and multi-name CDSs. In a single-name CDS there is one reference entity, whereas in a multi-name CDS, there are multiple reference entities. These two groups together add up to about 88% of the whole credit derivatives market in 2007 (Weistroffer, 2009 [59]). Within these groups, the single-name CDSs are the most important group in terms of market share. Figure 2.3 shows the relative market share of single-name and multi-name CDSs in percentages of the gross notional outstanding amounts.

![Figure 2.3: Market share of single-name and multi-name CDSs in 2011 (Vogel, Bannier & Heidorn, 2013 [57]).](image)

Although credit default swaps dominate the credit derivatives market, their market share in the overall over-the-counter derivatives market is not so large. As can be seen from Figure 2.4, the majority of the OTC market is composed of interest-rate contracts with a market share of 70%. The market share of credit default swaps was at a peak of 10% in 2007, but fell back to 4.5% in 2012.
Chapter 2. Credit Default Swap Market

For our research, we are specifically interested in sovereign CDSs (that is CDSs with a sovereign as reference entity). Since sovereign CDSs are mostly single-name CDSs (see figure 2.3), we will discuss the specifications and pay-off structure of a single-name credit default swap in more detail in the next section. After that, we will give a quick and broad overview of the sovereign CDS market and its development. Furthermore, we will discuss some issues related specifically to the sovereign CDS market.

2.1.1 Contract Specifications

A single-name credit default swap can be thought of as an insurance contract against default events on some reference entity specified by the contract\(^1\). The protection seller agrees to pay an amount to the protection buyer if the reference entity experiences a default event within the contract time of the CDS. The amount that is paid is usually not fixed upfront, but depends on the so-called recovery rate. In return for this insurance, the protection buyer pays a periodic protection fee to the protection seller. These fee payments stop if a default event occurs before the maturity of the contract. If a default occurs between two fee payment dates, the protection buyer still has to pay the accrued fee payment. Figure 2.5 shows the pay-off structure.

---

\(^1\)Strictly speaking, a CDS contract is not equal to an insurance contract, since buyers and sellers of a CDS do not necessarily have to hold an underlying asset. Therefore, holders of a CDS contract need not suffer a loss, whereas insurance policyholders do.
2.1. Credit Default Swaps

Credit default swaps are flexible contracts and consist of the following specifications:

1. The reference entity and its reference assets.
2. The default event definition.
3. The specification of the payment at default.
4. The notional value and credit default swap spread/premium.
5. The contract length and maturity date.
6. The frequency of the protection fee payments and the corresponding day count convention.

The specification of the reference entity is an obvious identifier since the default risk of this entity is the key point of the CDS contract. Reference assets (usually a set of bonds issued by the reference entity) need to be specified in order to be able to identify default events. Furthermore, they serve as a basis for determining the recovery rate in case of default (for cash settlement), and, in case of physical settlement, they include the deliverable assets. The notions of cash- and physical settlement will be explained below.

The definition of a default typically exists of (a subset of) the following components\(^2\):

- bankruptcy (the reference entity files for relief under bankruptcy law),
- failure to pay (the reference entity fails to make interest or principal payments on one of the reference assets),
- obligation default or obligation acceleration (similar to failure to pay),
- repudiation/moratorium (the reference entity disaffirms, disclaims or otherwise challenges the validity of a reference asset),
- restructuring (the reference entity changes the terms of one of the reference assets in such a way that the asset becomes less favourable to the holders (postponement or reduction of payments)).

\(^2\)These are the default events as specified by the ISDA. Precise definitions can be found in the ISDA 2003 Credit Derivatives Definitions (ISDA, 2003 [34]).
Chapter 2. Credit Default Swap Market

The bankruptcy event refers to the reference entity itself, whereas the other default events refer to the reference assets specified by the CDS contract. Of course, other default events may be agreed upon by the buyer and seller of the contract.

If a default occurs, the protection seller has to make a default payment to the protection buyer. A CDS contract typically specifies one of the following two ways to carry out the settlement between the protection buyer and the protection seller: physical settlement and cash settlement.

In the case of physical settlement, the protection buyer gives deliverable assets, with a total notional value equal to the notional value specified in the CDS contract, to the protection seller and he receives this total notional value in return. The net payment from the protection seller to the protection buyer is thus the loss given default (LGD) on the deliverable assets. Often, there is a set of possible deliverable assets and the protection buyer will, in that case, want to deliver the asset with the lowest market value. The possibility to choose among a set of deliverable assets is often referred to as the delivery option, and may enhance value for the protection buyer.

In the case of cash settlement, the protection buyer does not have to deliver any asset to the protection seller. Instead, the protection seller pays an amount equal to the difference between the reference assets market value at the time of settlement and its notional value (scaled to the notional value of the CDS contract). The determination of the market value of the reference asset is not straightforward due to liquidity problems in the distressed, post-default, market. Therefore, a robust procedure is needed to determine the market value.

In 2009, the ISDA introduced the so-called “big-bang protocol” and the “small-bang-protocol” for the CDS market, which are supplements to the 2003 Credit Derivatives Definitions. One of the main contributions of these supplements is the introduction of an auction settlement procedure that determines the final market value that establishes the physical and cash settlement. This auction settlement procedure has been the standard in the global sovereign CDS market since (Vogel, Bannier & Heidorn, 2013 [57]).

If one refers to the price of a CDS contract, one typically means the CDS premium (also referred to as the CDS spread). The CDS premium is quoted as an annual percentage (in basis points) of the notional value of the CDS contract. Although the quote is annual, the protection fee payments (coupons) are made quarterly, and are equal to the (quarterly) CDS spread times the notional amount (adjusted for day count conventions). Similar to interest rate swaps, the level of the CDS spread is the level that makes the initial value of the contract equal to zero. Higher spreads are usually associated with higher default probabilities.

Since the 2009 ISDA protocols, CDS premia have become standardized. The market, however, still quotes CDS premia as described above. If the market quote is higher than the standard premium, the buyer of a CDS makes an upfront payment, equal to the present value of the difference between the quoted spread and the standard coupon, to the seller. If the quoted CDS
2.2. Sovereign CDS Market

Premium is lower than the standard premium, the payment is made the other way around. Table 2.1 shows the standardized annual CDS premia (in basis points) for different reference entities as fixed by the ISDA. One can see, for example, that CDSs on Latin American sovereigns are either traded with (annualized) premia of 100 or 500 basis points and if, for example, the market would quote a CDS premium of 50 basis points, then the seller of the CDS would have to make a payment to the protection buyer equal to the present value of the difference between the quoted premium and the contract premium (which is either 100 or 500 basis points).

<table>
<thead>
<tr>
<th>Reference Entity Type</th>
<th>Implementation Date</th>
<th>25</th>
<th>100</th>
<th>500</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian Corporate &amp; Sovereign CDS</td>
<td>December, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian Corporate &amp; Sovereign CDS</td>
<td>September, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern &amp; Central European Corporate &amp; Sovereign CDS</td>
<td>September, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>European Corporate CDS</td>
<td>June, 2009</td>
<td>☑</td>
<td>☑</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td>Japanese Corporate &amp; Sovereign CDS</td>
<td>December, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latin American Corporate &amp; Sovereign CDS</td>
<td>September, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Eastern Corporate &amp; Sovereign CDS</td>
<td>September, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand Corporate &amp; Sovereign CDS</td>
<td>September, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North American Corporate &amp; Sovereign CDS</td>
<td>April, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western European Sovereign CDS</td>
<td>June, 2009</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Standardized coupons per reference entity type (Markit, 2009 [47]).

Next to the premium size, the premium payment dates also have become fixed. The standard payment dates are March 20, June 20, September 20 and December 20, and these standard dates also function as maturity dates. If, for example, a 5-year maturity CDS contract is agreed on April 1, 2010, the protection starts at April 2, 2010, and the first (accrued) fee payment is made on June 20, 2010. The contract matures at June 20, 2015. The contract length of a CDS contract typically ranges between 1 year and 10 years.

2.2 Sovereign CDS Market

Although the sovereign CDS market is only a small part of the total CDS market, it is rapidly increasing in (market) size, whereas the total CDS market is decreasing in size. Figure 2.6 shows the development of the sovereign CDS market compared to the single-name and multi-name CDS market as a whole.
The gross notional amount of sovereign CDS outstanding increased from roughly 500 billion USD in 2006, to 3 trillion USD in 2012. In the same period, the market share of single-name sovereign CDS in the single-name CDS market increased from 5% to 15%. Even after the beginning of the global financial crisis the sovereign CDS market kept on growing, whereas the total CDS market declined.

The drop of the overall CDS market size can again be explained by the need to hedge and limit counterparty credit risk exposure. An explanation for the growth of the sovereign CDS market is that, in light of the crisis, investors changed their perception of the riskiness of (developed) sovereigns and, therefore, they wanted to hedge against default risk of sovereigns, which were before assumed to be free of default risk. Figure 2.7 shows, for example, that the gross notional amount outstanding on developed sovereigns such as Germany, Belgium, France, Finland, Austria and the Netherlands (GBFFAN) increased from 104 billion USD in 2008 to 409 billion USD in 2012. Also (but maybe less surprisingly) the gross notional amount outstanding on the more distressed countries Portugal, Italy, Ireland and Greece (PIIGS) increased during that period.
2.2. Sovereign CDS Market

Figure 2.7: Gross notional amounts outstanding for European sovereigns (Vogel, Bannier & Heidorn, 2013 [57]).

Since sovereign CDS are over-the-counter contracts, it is hard to quantify the trading amounts of individual counterparties. On a less detailed level, however, the main market participants can be identified. Dealer banks, banks that are authorized to buy and sell government debt securities, dominate the buy and sell sides of the sovereign CDS market, because they want to manage their risk on their exposure on sovereigns. In a survey, Fitch Ratings (2010) [28] finds that the top 10 counterparties in the CDS market are involved in 80% of the trades. This result is supported by Figure 2.8, which is based on data from the Depositary Trust & Clearing Corporation (DTCC) and is taken from Vogel, Bannier and Heidorn (2013) [57] (note that a 0% means an insignificant market share).

Figure 2.8: CDS counterparty type (Vogel, Bannier & Heidorn, 2013 [57]).
The second largest market participant types are nondealer banks and security firms. Furthermore, hedge funds play an active role in the CDS market. Nondealer banks tend to be net sellers of sovereign CDS, whereas hedge funds are, at least from 2010 onward, net buyers (IMF, 2013 [33]).

Credit default swaps on sovereign reference entities have many things in common with credit default swaps on corporate reference entities. There are, however, also some differences between these two types. First of all, standard sovereign CDSs specify slightly different default events than corporate CDSs. For corporate CDSs, the default events are typically bankruptcy, failure to pay and restructuring (ISDA, 2012 [35]), whereas for sovereign CDSs the default events are failure to pay, repudiation or moratorium on debt and restructuring (ISDA, 2013 [36]). The absence of bankruptcy in the default definition of sovereign CDSs is because there is no bankruptcy court that applies to sovereign issuers.

Another major difference between the corporate and sovereign CDS market is the currency denomination of the CDS contract. Sovereign CDSs are typically denoted in another currency than the domestic currency. For example, European sovereign CDS are most often denoted in USD. The reason for this is that a sovereign default may induce a currency devaluation or redenomination of the domestic currency. If CDSs were also denominated in the domestic currency, the credit event would further destabilize the currency. In order to diversify currency exposure, sovereign CDSs are thus typically denoted in another currency (Vogel, Bannier & Heidorn, 2013 [57]).

The last difference we want to note is that the market for sovereign CDSs is (relatively) more active on a broad spectrum of maturities than the corporate CDS market. Almost all trading on corporate CDSs is concentrated on 5-year maturity contracts, whereas for sovereign CDSs also other maturities are traded. Figure 2.9 shows the relative market volumes of different maturity contracts for different types of CDS.
2.3 Liquidity Issues in the (Sovereign) CDS Market

In principle, the CDS spread should reflect the credit riskiness of the reference entity, and, under some assumptions, it should be possible to extract default probabilities from these CDS spreads. In reality, however, it is quite likely that the CDS spread is not only composed of components reflecting credit risk, and that other ‘distorting’ factors play a role as well. Some factors that might influence the CDS spread, for example, are liquidity risk, counterparty credit risk and contagion risk. Of all these distorting components, it is conjectured (by Rabobank) that liquidity risk has the most influence on the CDS spread\(^3\).

An often used indicator of liquidity problems is the bid-ask spread (which is defined as the ask spread minus the bid spread). Higher bid-ask spreads are related to more liquidity problems. In (sovereign) CDS markets, the bid-ask spread tends to be very high, and is, therefore, a clear sign of liquidity problems. Figure 2.10 shows, for example, the bid-ask spread of the five-year CDS on the Argentine government from the first of January 2008 till the beginning of 2014. It can be seen that the bid-ask spread is well over 500 basis points in some periods.

---

\(^3\)The academic literature also recognizes that liquidity effects should be taken into account. Some authors explicitly deal with this topic (see e.g., Bühler & Trapp (2008) [11]), whereas other authors indicate that they left out liquidity issues from their analysis, but that this topic should be investigated further (see e.g., Pan & Singleton (2008) [49], and Jarrow, Lando, Yu (2005) [38]).
Also the market participants themselves recognize the significant role of liquidity in the CDS market. In a survey among the most significant market participants, Fitch Ratings (2010) [28], asked what they thought the top challenges were for the CDS market in the future. The top three answers were (central) clearing, regulation and market liquidity. All these three answers directly or indirectly deal with the liquidity of the CDS market. The participants mainly named regulation, because they were afraid that overregulation could damage market liquidity\textsuperscript{4}.

Some measures to improve market liquidity are the introduction of central clearing and standardization of CDS contracts. The introduction of central clearing reduces search costs and counterparty credit risk, which are typical for over-the-counter market, thereby improving market liquidity. Central clearing is not only advocated by regulators, also market participants themselves see the benefits of such a market structure. In 2009, for example, the major European market participants for index- and single-name CDS agreed on trading through a central clearing party. The standardization of contracts further enhances the use of central clearing and market liquidity in general. Both market participants and regulators are, therefore, working towards more standardization of the market (Vogel, Bannier & Heidorn, 2013 [57]).

In light of the, still observed, large bid-ask spreads, we can, however, conclude that liquidity problems still remain an important issue in the (sovereign) CDS market. And, although, CDS spreads certainly reveal information about the credit riskiness of the reference entity, liquidity factors do distort this information.

\textsuperscript{4}A recent example of regulation that reduced market liquidity was the ban on ‘naked’ CDSs by the European Union in 2012 (IMF, 2013 [33]).
2.4 Chapter Summary

In this chapter, we discussed some background information on credit default swaps. Credit default swaps, especially single-name CDSs, are the most commonly traded credit derivatives. A CDS can be thought of as an insurance contract on default risk of a reference entity. The buyer pays premium fees to the seller until either a default event occurs or the contract reaches its maturity. In the case of a default event within the lifespan of the contract, the seller pays a protection amount to the buyer. This amount depends on the recovery rate, which is determined after the default event.

The most important specifications of a CDS contract are the reference entity, the reference assets, the default definition, the settlement procedure, the frequency and size of the premium payments, and the contract length. In principle, all these specifications are flexible, but the market is moving towards more standardization. CDS premia and payments dates, for example, are typically fixed by rules set by the International Swaps and Derivatives Association (ISDA).

If the reference entity of a CDS is a sovereign, one speaks of a sovereign credit default swap. Although sovereign CDSs still make up only a small part of the overall CDS market (15% market-share in the single-name CDSs in 2012), they are increasingly important. Corporate and sovereign CDSs are very similar, but there are some differences. Typically, the default event definition in sovereign CDSs differs slightly from that in corporate CDSs. Furthermore, sovereign CDSs are typically denoted in a different currency than the underlying reference assets, making them vulnerable to currency risks. A further difference is that, whereas for corporate CDSs almost all trades focus on 5 year maturity CDSs, the sovereign CDS market is more evenly spread among different maturities.

In principle the CDS premium should reflect the credit riskiness of the reference entity, but ‘distorting’ factors may also play a role. Because of the over-the-counter market structure and the large observed bid-ask spreads it is conjectured that especially liquidity risk is also priced into the CDS premia.
Quantitative models of credit risk can, broadly speaking, be divided into two classes: *structural models* and *reduced-form models*. The class of reduced-form models can again be subdivided in *intensity-based models* and *credit migration models*, where models from the latter class can be viewed as extensions of those from the former. Both structural and reduced-form models mainly deal with the modeling of the (random) default time (or the time at which the credit rating changes) and the recovery payments in case of default. Their approaches, however, differ.

Structural models, initiated by Merton (1974) [48] and Black and Cox (1976) [7], deal with modeling and pricing of credit risk at the individual firm (or sovereign) level. The focus of these models is on modeling the firm’s liabilities, which are seen as contingent claims issued against the total value of the firm’s asset. Default events are triggered when the firm’s value falls below some barrier that is specified either exogenously or endogenously with respect to the total value of the firm. The structural approach attempts to explain default events by economic fundamentals and is therefore attractive from a theoretical point of view. From a practical point of view, however, these models are less attractive, since they (unrealistically) assume that all assets of the firm are directly observable and tradable. Furthermore, complex priority structures of the firm’s liabilities need to be specified in the valuation procedure, and often these models tend to be analytically complex and computationally expensive (Wang, 2009 [58]).

Reduced-form models, on the other hand, do not try to explain default events by economic factors. Within the sub-class of intensity-based models, default events are modeled in terms of an exogenously specified jump process. The default time is then specified as the (first) jump time of this jump process. Often, the starting point of these models is the modeling of the so-called *default intensity process*, also referred to as the *hazard rate process*. These processes directly affect the (conditional) probabilities of a jump of the underlying jump process, and,
hence, the (conditional) default probabilities. Next to the modeling of the default intensity process itself, an important role is played by the conditioning information used to model the default intensity process. One of the main attractions of intensity-based models, as we will show below and in chapter 4, is that there is a great analogy with the well-known class of short-rate interest rate models. Because of this analogy, these models provide analytical tractability and they are relatively easy to implement. Furthermore, intensity-based models can be useful for extracting default probabilities from market prices (Brigo & Mercurio, 2006 [9]).

The credit migration models are an extension of intensity-based models. Within these models, the quality of (corporate/sovereign) debt is categorized into a finite number of disjoint credit rating classes. The quality of debt can change over time and these changes can induce changes in the rating of the debt. This migration of credit quality is often modeled as a Markov chain, where the default state is an absorbing state. The main modeling issue is to specify the transition intensities matrix for the migration process. Traditional intensity-based models can be viewed as credit migration models with two states: the ‘normal’ state and the default state (Bielecki & Rutkowski, 2002 [6]).

Our research focuses on extracting default probabilities from credit default swap data and, by definition, this means that we are interested in the opinion of the market on the credit riskiness of the reference entity and not in the economic fundamentals driving the default probability. This, in combination with the disadvantages mentioned before, makes the class of structural models inappropriate for our purposes. The class of intensity-based models, on the other hand, is suitable, since one can relatively easy construct pricing formulas that can be calibrated to observed data. The calibrated pricing formulas then reveal information on the default intensity process, which can be used in order to compute default probabilities.

In this thesis, we will introduce our ‘credit-liquidity’ model, which we will use in order to extract default probabilities from CDS data. This model is an intensity-based model that also incorporates liquidity effects by means of extra liquidity discount factors. The introduction of this model will, however, be postponed until chapter 5, since we first want to give a thorough introduction to the class of intensity-based models and issues related to the modeling of the default intensity process in this chapter and chapter 4, respectively.

In this chapter, we will explain how intensity-based models are constructed and how one can price defaultable claims within this framework. Furthermore, we will add a discussion on some issues related to the change of probability measure within this framework, since this topic is more complicated than in the case of, for example, short-rate interest rate models. We need this discussion, since calibrating CDS premia formulas to data is done under the risk-neutral measure, but the actual default probabilities have to be computed under the real-world probability measure. The Rabobank already has a model that partially deals with this topic and, therefore, we will briefly mention this and discuss the implications for our research. More details on the mathematical foundation of intensity-based models can be found in Appendices A.
and B, which are primarily based on Bielecki and Rutkowski (2002) [6] and Filipović (2009) [27].

3.1 Definitions and Filtrations

In the introduction of this chapter, we mentioned that the starting point of modeling within intensity-based models is the modeling of the hazard (rate) process. Of course, the concept of a hazard process only makes sense if there is an actual underlying default process. In section 3.2, we will, therefore, explain how we can construct a default process for a given hazard process. In this section, we will reverse the order and we will assume that we are given a default process and we will define what the associated hazard process is in order to get familiar with this concept.

Let us thus consider the default time $\tau$, which is modeled as a random time on a probability space $(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}\{\tau = 0\} = 0$ and $\mathbb{P}\{\tau > t\} > 0$ for any $t \in \mathbb{R}_+$ and $\mathcal{G} = \mathcal{H} \lor \mathcal{F}$ (that is $\mathcal{G}_t = \mathcal{H}_t \lor \mathcal{F}_t$ for all $t \in \mathbb{R}_+$). Here $\mathcal{H}$ is the natural filtration of the jump process $H_t = 1_{\{\tau \leq t\}}$ and $\mathcal{F}$ is some reference filtration, which in applications follows naturally as the filtration generated by a certain stochastic process. For example, Schönbucher (2003) [52], introduces a $d$-dimensional background driving process, $(\mathcal{X}_t)_{t \geq 0}$, and states that $\mathcal{F}$ is the filtration generated by $\mathcal{X}_t$. All default-free processes, such as the default-free interest rates, are then assumed to be adapted to this filtration $(\mathcal{F}_t)_{t \geq 0}$. It is not essential that the background filtration is generated by such a stochastic process $\mathcal{X}_t$, but it is convenient to think of it that way\(^1\). We assume that the filtrations $\mathcal{H}$ and $\mathcal{F}$ satisfy the ‘usual’ conditions of right-continuity and completeness and that $\mathcal{F} \subset \mathcal{G}$. For convenience, we summarize these assumptions as follows:

**Condition 1.** The filtration $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$ on $(\Omega, \mathcal{G}, \mathbb{P})$ represents the total information available at time $t$ and is such that we can write $\mathcal{G} = \mathcal{H} \lor \mathcal{F}$, so $\mathcal{G}_t = \mathcal{H}_t \lor \mathcal{F}_t$ for every $t \in \mathbb{R}_+$. Here $\mathcal{F}$ is a given auxiliary filtration and $\mathcal{H}$ is the natural filtration of the jump process $H_t = 1_{\{\tau \leq t\}}$. We also assume, for simplicity, that $\mathcal{F}_0$ is trivial and, therefore, $\mathcal{G}_0$ is trivial as well. Furthermore, all filtrations are assumed to satisfy the usual conditions of right-continuity and completeness.

Let $F_t = \mathbb{P}\{\tau \leq t|\mathcal{F}_t\}$ for every $t \in \mathbb{R}_+$. We denote by $G$ the $\mathcal{F}$-survival process of $\tau$ with respect to the filtration $\mathcal{F}$, where

$$G_t := 1 - F_t = \mathbb{P}\{\tau > t|\mathcal{F}_t\}, \quad \forall t \in \mathbb{R}_+.$$

**Definition 3.1.** Assume that $F_t < 1$ for all $t \in \mathbb{R}_+$. We denote by $\Gamma$ the $\mathcal{F}$-hazard process of $\tau$ under $\mathbb{P}$ and we have $\Gamma_t = -\ln G_t = -\ln(1 - F_t)$. If we furthermore assume that $\Gamma$ is absolutely continuous with respect to the Lebesgue measure, we can write

\(^1\)In most applications, we have $\mathcal{F} = \mathcal{F}^W$ for some (multi-dimensional) Brownian motion $W$. 


where \( \Gamma_t = \int_0^t \gamma_u du \),

where \( \gamma \) is some \( \mathbb{F} \)-progressively measurable process. We refer to \( \gamma \) as the \( \mathbb{F} \)-intensity of \( \tau \) or \( \mathbb{F} \)-hazard rate process of \( \tau \).

Throughout this chapter, we will, next to Condition 1, assume that the following condition holds:\(^2\)

**Condition 2.** For all \( t \in \mathbb{R}_+ \) and every \( u \leq t \), we have

\[
P\{ \tau > u | \mathcal{F}_\infty \} = P\{ \tau > t | \mathcal{F}_t \}.
\]

It can be shown that Condition 2 is equivalent to the *martingale invariance property of \( \mathbb{F} \) with respect to \( \mathbb{G} \),* which states that every \( \mathbb{F} \)-martingale also follows a \( \mathbb{G} \)-martingale\(^3\).

Closely related to the \( \mathbb{F} \)-hazard process of a random time \( \tau \) is the so-called \((\mathbb{F}, \mathbb{G})\)-martingale hazard process of a random time \( \tau \), which is defined as follows:

**Definition 3.2.** An \( \mathbb{F} \)-predictable, right-continuous, increasing process \( \Lambda \), with \( \Lambda_0 = 0 \), is called a \((\mathbb{F}, \mathbb{G})\)-martingale hazard process if and only if the process \( \tilde{M}_t := H_t - \Lambda_{t\wedge \tau} \) follows a \( \mathbb{G} \)-martingale. If \( \Lambda_t = \int_0^t \lambda_u du \), then the \( \mathbb{F} \)-progressively measurable non-negative process \( \lambda \) is referred to as the \((\mathbb{F}, \mathbb{G})\)-martingale intensity process.

Since a random time \( \tau \) and a given reference filtration \( \mathbb{F} \) uniquely determine \( \mathbb{G} = \mathbb{H} \lor \mathbb{F} \), we can also refer to the \((\mathbb{F}, \mathbb{G})\)-martingale hazard process as the \( \mathbb{F} \)-martingale hazard process of \( \tau \).

In the present set-up, we have that the \((\mathbb{F}, \mathbb{G})\)-martingale hazard process of \( \tau \) always exists and is unique (Bielecki & Rutkowski, 2002 [6]). Furthermore, the following relationship between the \( \mathbb{F} \)-hazard process and \((\mathbb{F}, \mathbb{G})\)-martingale hazard process exists:

**Proposition 3.3.** Let all the conditions stated so far hold.

(i) If the increasing process \( \mathbb{F} \) is \( \mathbb{F} \)-predictable, but \( \mathbb{F} \) is not continuous, then the \( \mathbb{F} \)-martingale hazard process \( \Lambda \) is also a discontinuous process and we have

\[
e^{-\Gamma_t} = e^{-\Lambda^c_t} \prod_{0<u\leq t} (1 - \Delta \Lambda_u),
\]

where \( \Lambda^c \) is the continuous component of \( \Lambda \).

\(^2\)This condition follows naturally from the construction of the default time for a given hazard rate process, which will be explained in section 3.2.

\(^3\)See Lemma B.2 in Appendix B.
3.1. Definitions and Filtrations

(ii) If the increasing process $F$ is continuous, then the $\mathcal{F}$-martingale hazard process $\Lambda$ is also continuous an

$$\Gamma_t = \Lambda_t = -\ln(1 - F_t), \quad \forall \, t \in \mathbb{R}^+.$$  

If, in addition, the process $\Lambda = \Gamma$ is absolutely continuous, then for an integrable $\mathcal{F}_s$-measurable random variable $Y$ we get

$$\mathbb{E}^\mathbb{P}\left[1_{\{\tau > s\}}Y \bigg| \mathcal{G}_t\right] = 1_{\{\tau > t\}} \mathbb{E}^\mathbb{P}\left[Y e^{-\int_t^\tau \lambda_u \, du} \bigg| \mathcal{F}_t\right].$$

Proof. See Proposition B.5 in Appendix B.

Usually, the starting point of intensity-based models is the specification of the default intensity process. Modeling the default intensity process automatically implies that the hazard process is absolutely continuous with respect to the Lebesgue measure (otherwise the intensity process would not be defined). Furthermore, the default time $\tau$ is usually either the first jump time of a Poisson process, or it is constructed using the canonical approach. In both cases, all conditions and assumptions stated above hold, and in combination with the absolute continuity of the hazard process, we can conclude that the hazard process is equal to the martingale hazard process and that the hazard rate process is equal to the martingale hazard rate process.

The following result allows us to ‘switch’ from conditioning filtration in conditional expectations. Brigo and Mercurio (2006) [9] therefore call this result the “Filtration Switching Formula”.

**Lemma 3.4.** Let $Y$ be a $\mathcal{G}$-measurable random variable, then

$$\mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}}Y \bigg| \mathcal{G}_t\right] = 1_{\{\tau > t\}} \frac{\mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}}Y \bigg| \mathcal{F}_t\right]}{\mathbb{P}\{\tau > t|\mathcal{F}_t\}}.$$  

(3.1)

In particular, for any $t \leq s$,

$$\mathbb{P}\{t < \tau \leq s|\mathcal{G}_t\} = 1_{\{\tau > t\}} \frac{\mathbb{P}\{t < \tau \leq s|\mathcal{F}_t\}}{\mathbb{P}\{\tau > t|\mathcal{F}_t\}}.$$  

(3.2)

This result is important, since the computation of certain expressions conditioned on $\mathcal{F}$ is easier than the computation of these expressions conditioned on $\mathcal{G}$. This is a consequence of the fact that one usually directly models the $\mathcal{F}$-(martingale) hazard process and by Lemma 3.4, we

---

*See section 3.2.*
obtain an expression with this process in it. For example, we have that $P\{\tau > t|F_t\} = e^{-\Gamma_t}$, and, therefore, we can rewrite (3.1) as follows:

$$E^P[1_{\{\tau > t\}}Y|G_t] = 1_{\{\tau > t\}}E^P[1_{\{\tau > t\}}e^{\Gamma_t}Y|F_t].$$

In the case that the $F$-(martingale) hazard process is absolutely continuous, we can write $\Gamma_t = \int_0^t \gamma_u du$ for some $F$-progressively measurable process $\gamma$, and the following holds:

$$P\{\tau > s|G_t\} = 1_{\{\tau > t\}}E^P\left[1_{\{\tau > s\}}e^{\int_s^t \gamma_u du}|F_t\right].$$

The last formula explains why we refer to the hazard rate process $\gamma$ as the default intensity process. Furthermore, we see that the expression for the default probability is similar to a zero-coupon bond price formula if we would interpret $\gamma$ as a short-rate process. This analogy with short-rate models makes intensity-based models so powerful and popular.

The following two propositions turn out to be very helpful in the pricing of defaultable claims (see, e.g., Proposition 3.13):

**Proposition 3.5.** Let $Y$ be $G$-measurable random variable and let $t \leq s$.

(i) If for every $t \in \mathbb{R}_+$ we have $F_t \subseteq G_t \subseteq H_t$ $\cap F_t$, then

$$E^P[1_{\{t < \tau \leq s\}}Y|G_t] = 1_{\{\tau > t\}}E^P[1_{\{t < \tau \leq s\}}e^{\Gamma_s}Y|F_t].$$

(ii) If the assumptions stated in Condition 1 hold, then

$$E^P[1_{\{t < \tau \leq s\}}Y|G_t] = 1_{\{\tau > t\}}E^P[1_{\{t < \tau \leq s\}}e^{\Gamma_t}Y|F_t].$$

(iii) If in addition $Y$ is $F_s$-measurable, then

$$E^P[1_{\{\tau > s\}}Y|G_t] = 1_{\{\tau > t\}}E^P[e^{\Gamma_s}e^{-\Gamma_s}Y|F_t].$$
Proposition 3.6.  

(i) Let \( h : \mathbb{R}_+ \to \mathbb{R} \) be a bounded, continuous function. Then for any \( t < s \leq \infty \)

\[
\mathbb{E}^\mathbb{P} \left[ \mathbf{1}_{\{t < \tau \leq s\}} h(\tau) \bigg| \mathcal{G}_t \right] = \mathbf{1}_{\{\tau > t\}} e^{\Gamma_t} \mathbb{E}^\mathbb{P} \left[ \int_{[t,s]} h(u) dF_u \bigg| \mathcal{F}_t \right]. \tag{3.7}
\]

(ii) Let \( Z \) be a bounded \( \mathbb{F} \)-predictable process. Then for any \( t < s \leq \infty \)

\[
\mathbb{E}^\mathbb{P} \left[ \mathbf{1}_{\{t < \tau \leq s\}} Z_\tau \bigg| \mathcal{G}_t \right] = \mathbf{1}_{\{\tau > t\}} e^{\Gamma_t} \mathbb{E}^\mathbb{P} \left[ \int_{[t,s]} Z_u dF_u \bigg| \mathcal{F}_t \right]. \tag{3.8}
\]

Proof. See Proposition A.8 in Appendix A. \( \Box \)

3.2 Construction of the Default Process

In most intensity-based models, the starting point is the specification of the hazard (rate) process. Of course, this process has no meaning if there is no underlying default (jump) process. Therefore, one has to construct a default process that ‘matches’ with the specified intensity process.

The two most common ways to specify the random default time \( \tau \) for a given default intensity process, are

1. The canonical construction of a default time.
2. The first jump time of a doubly stochastic Poisson process.

We shall briefly explain both methods, but for more details we refer to Bielecki and Rutkowski (2002) [6].

3.2.1 Canonical Construction of Default Time

Let \( \Gamma \) be a given \( \mathbb{F} \)-adapted increasing process that is absolutely continuous with respect to the Lebesgue measure on some filtered probability space \((\Omega, \mathbb{F}, \mathbb{P})\). Furthermore, let \( \Gamma \) be such that \( \Gamma_0 = 0 \) and \( \Gamma_\infty = +\infty \). We can then write

\[
\Gamma_t = \int_0^t \gamma_u du, \quad \forall t \in \mathbb{R}_+, \tag{3.9}
\]

where \( \gamma \) is a non-negative, \( \mathbb{F} \)-progressively measurable process.
Chapter 3. Intensity-Based Models

Assume now that $\phi$ is a uniform($0, 1$)-random variable on some probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, and let us define the product space $(\Omega, \mathcal{G}, \mathbb{P})$ with $\Omega = \hat{\Omega} \times \hat{\Omega}$, $\mathcal{G} = \mathcal{F}_\infty \times \hat{\mathcal{F}}$ and $\mathbb{P} = \hat{\mathbb{P}} \times \hat{\mathbb{P}}$. We define the random time $\tau$ on this product space by

$$
\tau = \inf \left\{ t \in \mathbb{R}_+ : e^{-\Gamma_t} \leq \phi \right\} = \inf \left\{ t \in \mathbb{R}_+ : \Gamma_t \geq -\ln(\phi) \right\}.
$$

(3.10)

Lastly, we set $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$, where $\mathcal{H}_t$ is the natural filtration of the process $1_{\{\tau \leq t\}}$. By construction, we now have that Condition 1 is satisfied. The following result shows that also Condition 2 is satisfied and that, under this construction of $\tau$, the process $\Gamma$ is both the $\mathbb{F}$-hazard process and $\mathbb{F}$-martingale hazard process of $\tau$ under $\mathbb{P}$:

**Proposition 3.7.** Let the process $\Gamma$ and the random time $\tau$ be given by (3.9) and (3.10) respectively. Then

(i) $\Gamma$ is the $\mathbb{F}$-hazard process of $\tau$.

(ii) $\Gamma$ is the $\mathbb{F}$-martingale hazard process of $\tau$.

**Proof.** We recall from Definition 3.1 that the $\mathbb{F}$-hazard process of the random time $\tau$ under $\mathbb{P}$ is given by the process $\Phi$, with $\Phi_t = -\ln(1 - F_t)$, where $F_t = \mathbb{P}(\tau \leq t | \mathcal{F}_t)$. We have

$$
1 - F_t = \mathbb{P}(\tau > t | \mathcal{F}_t) = \mathbb{E}_\mathbb{P} \left[ \mathbb{P}(\tau > t | \mathcal{F}_\infty) | \mathcal{F}_t \right] = e^{-\Gamma_t}.
$$

The last step follows, since, by definition of $\tau$, we have $\{\tau > t\} = \{e^{-\Gamma_t} > \phi\}$, and therefore, $\mathbb{P}(\tau > t | \mathcal{F}_\infty) = e^{-\Gamma_t}$. We observe that $F$ is an $\mathbb{F}$-adapted, continuous increasing process and we can thus conclude that $\Gamma_t = -\ln(1 - F_t)$ is the $\mathbb{F}$-hazard process of $\tau$.

To see that $\Gamma$ is also the $\mathbb{F}$-martingale hazard process of $\tau$, it is enough to check that Condition 2 is satisfied, since we can then apply Proposition 3.3. Let us fix $t$ and consider any $u \leq t$. We get

$$
\mathbb{P}(\tau \leq u | \mathcal{F}_t) = \mathbb{E} \left[ \mathbb{P}(\tau \leq u | \mathcal{F}_\infty) | \mathcal{F}_t \right] = 1 - e^{-\Gamma_u} = \mathbb{P}(\tau \leq u | \mathcal{F}_\infty).
$$

Condition 2 is thus satisfied, and, therefore, we can conclude that $\Gamma$ is also the $\mathbb{F}$-martingale hazard process of $\tau$. \(\blacksquare\)

The procedure described above, allows us to define a default time $\tau$ for a given hazard process in such a way that Conditions 1 and 2 are satisfied. By virtue of Lemma B.2, we then also have that every $\mathbb{F}$-martingale is a $\mathcal{G}$-martingale.
3.2. Construction of the Default Process

3.2.2 Doubly Stochastic Poisson Processes

An alternative construction of the default time for a given hazard rate process is to define a Poisson process, which has as intensity process the given hazard rate process. Such a process is called a doubly stochastic Poisson process. The results of this approach are qualitatively very similar to those of the canonical approach, but since this approach is very popular in the literature (see, e.g., Lando (1998) [44] or Duffie and Singleton (1999) [25]), we decided to spend a few words on it.

First, let us define a Poisson process $\mathcal{N}$ with deterministic and constant intensity $\lambda > 0$:

**Definition 3.8.** A non-decreasing, integer-valued $\mathcal{G}$-adapted process $\mathcal{N}$ defined on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ is a Poisson process with intensity $\lambda > 0$ with respect to $\mathcal{G}$, if

(i) $N_0 = 0$,
(ii) For any $s < t$, the increment $N_t - N_s$ is independent of $\mathcal{G}_s$,
(iii) For all $k = 0, 1, \ldots$, and $s < t$ we have

$$\mathbb{P}\{N_t - N_s = k | \mathcal{G}_s\} = \mathbb{P}\{N_t - N_s = k\} = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda (t-s)}.$$

The Poisson process of Definition 3.8 is called a time-homogeneous Poisson process, since the probability law of the increment $N_{t+h} - N_{s+h}$ is equal to the probability law of the increment $N_t - N_s$ for all $h, \tilde{h} \geq -s$. This means that the law of an increment is determined by the time-difference of the increment and not the time itself. Some useful properties of the time-homogeneous Poisson process are:

**Proposition 3.9.** Let $\mathcal{N}$ be a Poisson process as defined in Definition 3.8. Then the following statements hold:

(i) The compensated process $\hat{\mathcal{N}}_t = N_t - \lambda t$ follows a $\mathcal{G}$-martingale.
(ii) The times between any two jumps are identically independent exponentially distributed random variables with parameter $\lambda$.
(iii) Let $W$ be a Brownian motion defined on the same probability space $(\Omega, \mathcal{G}, \mathbb{P})$ as the Poisson process $\mathcal{N}$. Then the processes $W$ and $\mathcal{N}$ are independent.

A Poisson process defined by 3.8 has a constant intensity. Of course, one could also think of the situation where the intensity is given by a time-dependent function. In this case we can define a so-called time-inhomogeneous Poisson process.

**Definition 3.10.** Let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ be any non-negative, locally integrable function such that $\int_0^\infty \lambda_u du = \infty$. The $\mathcal{G}$-adapted process $\mathcal{N}$ defined on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ is a time-inhomogeneous Poisson process with intensity function $\lambda$, if
(i) \( N_0 = 0 \),
(ii) For every \( 0 \leq s < t \) the increment \( N_t - N_s \) is independent of \( \mathcal{G}_s \), and has the Poisson law with parameter \( \Lambda(t) - \Lambda(s) \), where \( \Lambda(t) = \int_0^t \lambda_u \, du \).

Most properties of time-homogeneous Poisson processes extend to the class of time-inhomogeneous Poisson processes. In particular, a time-inhomogeneous Poisson process and a Brownian motion with respect to \( \mathcal{G} \) follow independent processes under \( \mathbb{P} \).

A further extension of Poisson processes is when we not only assume that the intensity is time-dependent, but also stochastic. In this case we have a so-called doubly stochastic Poisson process, also referred to as a conditional Poisson process. Let us be given a probability space \((\Omega, \mathcal{G}, \mathbb{P})\) and a sub-filtration \( \mathcal{F} \) of \( \mathcal{G} \). Furthermore, let \( \Phi \) be an \( \mathcal{F} \)-adapted, right-continuous, increasing process, with \( \Phi_0 = 0 \) and \( \Phi_\infty = \infty \). Then \( \Phi \) is referred to as the hazard process and usually we have that \( \Phi_t = \int_0^t \phi_u \, du \) for some \( \mathcal{F} \)-progressively measurable process \( \phi \) with locally integrable sample paths. The process \( \phi \) is called the intensity process. Now the conditional Poisson process is defined as follows:

**Definition 3.11.** A \( \mathcal{G} \)-adapted process \( N \) defined on the probability space \((\Omega, \mathcal{G}, \mathbb{P})\) is an \( \mathcal{F} \)-conditional Poisson process with respect to \( \mathcal{G} \), associated with the hazard process \( \Phi \), if for any \( 0 \leq s < t \) and every \( k = 0, 1, \ldots \)

\[
\mathbb{P}\{N_t - N_s = k|\mathcal{G}_s \vee \mathcal{F}_\infty\} = \frac{(\Phi_t - \Phi_s)^k}{k!}e^{-(\Phi_t - \Phi_s)}
\]

Definition 3.11 thus says that if the sample path of the hazard process \( \Phi \) is known (by conditioning on \( \mathcal{F}_\infty \)), then the process \( N \) follows exactly a time-inhomogeneous Poisson process with respect to \( \mathcal{G} \). We also have that an \( \mathcal{F} \)-conditional Poisson process with respect to \( \mathcal{G} \), is an \( \mathcal{F} \)-conditional Poisson process with respect to \( \mathbb{P}^N \vee \mathcal{F} \).

We still have to show that, given a probability space \((\Omega, \mathcal{G}, \mathbb{P})\), two filtrations \( \mathcal{F} \) and \( \mathcal{G} \), and a hazard process \( \Phi \), we can construct an \( \mathcal{F} \)-conditional Poisson process. Following Bielecki and Rutkowski (2002) [6], we assume that the probability space \((\Omega, \mathcal{G}, \mathbb{P})\), with filtration \( \mathcal{F} \) is large enough to accommodate for a Poisson process \( \tilde{N} \) with constant intensity \( \lambda = 1 \), independent of \( \mathcal{F} \), and an \( \mathcal{F} \)-adapted hazard process \( \Phi \) (we can always enlarge a probability space to accommodate for this). The following result tells us how to construct a conditional Poisson process.

**Proposition 3.12.** Let \( \tilde{N} \) be a Poisson process with constant intensity \( \lambda = 1 \), independent of the reference filtration \( \mathcal{F} \). Let \( \Phi \) be an \( \mathcal{F} \)-adapted, right-continuous, increasing process. Then the process \( N_t = \tilde{N}_{\Phi_t}, t \in \mathbb{R}_+ \), follows an \( \mathcal{F} \)-conditional Poisson process with hazard process \( \Phi \) with respect to the filtration \( \mathcal{G} = \mathcal{F}^N \vee \mathcal{F} \).  

28
3.3. Pricing of Defaultable Contingent Claims

Proof. By Definition 3.11, it suffices to check
\[ \mathbb{P}\{N_t - N_s = k | \mathcal{G}_s \vee \mathcal{F}_\infty\} = \frac{(\Phi_t - \Phi_s)^k}{k!} e^{-(\Phi_t - \Phi_s)}. \]

Note that \( \mathcal{G}_s \vee \mathcal{F}_\infty = \mathcal{F}_s^N \vee \mathcal{F}_\infty \), and therefore we need to check
\[ \mathbb{P}\{N_t - N_s = k | \mathcal{F}_s^N \vee \mathcal{F}_\infty\} = \mathbb{P}\{\tilde{N}_{\Phi_t} - \tilde{N}_{\Phi_s} = k | \mathcal{F}_{\Phi_s}^N \vee \mathcal{F}_\infty\} = \frac{(\Phi_t - \Phi_s)^k}{k!} e^{-(\Phi_t - \Phi_s)}. \]

The first equality follows by definition of \( \tilde{N} \). The second equality follows from the independence of \( \tilde{N} \) and \( \mathcal{F}_s \).

In most financial models, only the first jump of the Poisson process matters, and we have that the default time is defined as \( \tau = \tau_1 \). For this definition of the default time, we have that Condition 2 holds, since, for any \( t \in \mathbb{R}_+ \) and \( u \geq t \), we have
\[ \mathbb{P}\{\tau \leq t | \mathcal{F}_u\} = \mathbb{P}\{N_t \geq 1 | \mathcal{F}_u\} = \mathbb{P}\{N_t - N_0 \geq 1 | \mathcal{G}_0 \vee \mathcal{F}_u\} = \mathbb{P}\{\tau \leq t | \mathcal{F}_\infty\}. \]

the last equality follows from the fact that for every \( 0 \leq s < t \leq u \), and every \( k = 0, 1, \ldots \), we have
\[ \mathbb{P}\{N_t - N_s = k | \mathcal{G}_s \vee \mathcal{F}_\infty\} = \mathbb{P}\{N_t - N_s = k | \mathcal{F}_s\} \text{ (time-inhomogeneous Poisson process property)} \]
\[ = \mathbb{P}\{N_t - N_s = k | \mathcal{G}_s \vee \mathcal{F}_u\} \]
\[ = \frac{(\Phi_t - \Phi_s)^k}{k!} e^{-(\Phi_t - \Phi_s)}. \]

We again conclude that, by modeling the default time as the first jump time of the conditional Poisson process, constructed as in Proposition 3.12, we have that Conditions 1 and 2 hold.

3.3 Pricing of Defaultable Contingent Claims

In most applications, intensity-based models are used for pricing purposes. In this case, one starts with modeling the hazard (rate) process \( \Gamma \) under a risk-neutral measure \( \mathbb{Q} \) instead of
the objective measure $\mathbb{P}$. In this section we will give some results on the pricing of defaultable contingent claims.

Let us assume that we constructed a random process $\tau$ on a probability space $(\Omega, \mathcal{G}, \mathbb{Q})$, such that $\Gamma$ is its $\mathbb{F}$-hazard process under $\mathbb{Q}$ using the canonical approach. In general, a defaultable claim can be characterized by the tuple $DCT = (X, A, \tilde{X}, Z, \tau)$, where

- $X$ is the promised claim at maturity time $T$,
- $A$ is the process that describes the promised dividends,
- $\tilde{X}$ is the recovery payoff received at time $T$ if default occurs prior to or at time $T$,
- $Z$ is the recovery payoff process that describes the recovery payoff at the time of default if it occurs prior to or at time $T$,
- $\tau$ is the (random) default time.

In most practical situations, one either has that $Z = 0$ or $\tilde{X} = 0$. In the former case, we have a defaultable claim with recovery at maturity, and in the latter case, we have a defaultable claim with recovery at default. In the case of credit default swap pricing and bond pricing, we usually have recovery at default and no dividend payments. Therefore, we will focus on defaultable claims given by the tuple $DCT = (X, 0, 0, Z, \tau)$.

On the technical side, we assume that the process $Z$ is $\mathbb{F}$-predictable and that the random variable $X$ is $\mathcal{G}_T$-measurable. Furthermore, we assume that suitable integrability conditions hold such that all expectations below are well-defined.

Let $S_t$ denote the pre-default value of a defaultable claim $(X, 0, 0, Z, \tau)$, then it is clear that we can write

$$S_t = B_t \mathbb{E}^\mathbb{Q} \left[ 1_{\{t < \tau \leq T\}} B^{-1}_\tau Z_\tau + B^{-1}_T X 1_{\{T < \tau\}} \right] | \mathcal{G}_t] , \quad (3.11)$$

where $B_t = e^{\int_0^t r_u \, du}$, with $r_u$ the short-term interest rate process. We see that $S_t$ is the sum of the (conditional) expected payoff of the contingent claim $X$ at $T$ and the expected recovery payoff at $\tau \leq T$. The following result gives the representation of the pre-default value after switching the conditioning filtration.

**Proposition 3.13.** The pre-default value process $S_t$ of the defaultable claim $(X, 0, 0, Z, \tau)$ admits the following representation for $t \in [0, T]$:

$$S_t = 1_{\{\tau > t\}} B_t \mathbb{E}^\mathbb{Q} \left[ \int_{[t, T]} B^{-1}_u e^{\Gamma_t - \Gamma_u} Z_u d\Gamma_u + B^{-1}_T X e^{\Gamma_t - \Gamma_T} \right] | \mathcal{F}_t] . \quad (3.12)$$

**Proof.** First we consider
3.3. Pricing of Defaultable Contingent Claims

\[ J_t(Z) = B_t \mathbb{E}^Q \left[ \mathbbm{1}_{\{t < \tau \leq T\}} B^{-1}_\tau Z_\tau \bigg| \mathcal{G}_t \right] . \]

By applying formula (3.8), we obtain

\[ J_t(Z) = -\mathbbm{1}_{\{\tau > t\}} e^{\Gamma_\tau} B_t \mathbb{E}^Q \left[ \int_{[t, T]} B^{-1}_u Z_u dG_u \bigg| \mathcal{F}_t \right] . \]

Since the survival process \( G \) is continuous, the hazard process \( \Gamma \) is an increasing continuous process. We have \( dG_u = -e^{-\Gamma_u} d\Gamma_u \), and therefore we get

\[ J_t(Z) = \mathbbm{1}_{\{\tau > t\}} B_t \mathbb{E}^Q \left[ \int_{[t, T]} B^{-1}_u Z_u e^{\Gamma_{u-t}} d\Gamma_u \bigg| \mathcal{F}_t \right] . \]

Now consider

\[ K_t = B_t \mathbb{E}^Q \left[ B^{-1}_T X \mathbbm{1}_{\{T < \tau\}} \bigg| \mathcal{G}_t \right] . \]

We get by (3.6) that

\[ K_t = \mathbbm{1}_{\{\tau > t\}} B_t \mathbb{E}^Q \left[ B^{-1}_T X e^{\Gamma_{T-t}} \bigg| \mathcal{F}_t \right] . \]

Clearly, we have \( S_t = J_t(Z) + K_t \), and therefore the result follows. \( \Box \)

If we assume that the default time admits the stochastic intensity process \( \gamma \), then formula (3.12) becomes

\[ S_t = \mathbbm{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \int_{[t, T]} e^{-\int_{t}^{u} (r_v + \gamma_v) dv \gamma_u Z_u du} \bigg| \mathcal{F}_t \right] + \mathbbm{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_{t}^{T} (r_v + \gamma_v) dv X} \bigg| \mathcal{F}_t \right]. \] (3.13)

3.3.1 Recovery Modeling

Until now, we have not specified any specific recovery rule or process. We merely denoted the recovery process \( Z \) as an \( \mathbb{F} \)-predictable stochastic process. In the literature, several recovery models have been proposed: zero recovery (ZR), recovery of treasury (RT), recovery of par (RP) and recovery of market value (RMV) (Schönbucher, 2003 [52]). Some of these recovery
rules are more appropriate in the one case, whereas other rules are more appropriate in the other case.

In this section, we will show how the choice of the recovery model influences the pricing formulas and, thereby, the computed value of a defaultable claim. We illustrate these effects by computing the value of a defaultable zero-coupon bond under all of the recovery rules. Note that a defaultable zero-coupon bond is a defaultable claim with $X = 1$ and we will denote its value, given that no default has occurred until time $t$, by $\bar{B}(t, T)$. We will assume that the default time $\tau$ admits a stochastic intensity process $\gamma$ and that, therefore, the pre-default value is given by formula (3.13).

**Zero Recovery**

The zero recovery rule is the simplest of all and just states that there is no recovery ($Z = 0$). This model is often the most unrealistic, but it can be useful as a benchmark or as an intermediary result in other recovery rules. The price of a defaultable zero-coupon bond, using the general formula (3.13), is simply given by

$$\bar{B}(t, T) = 1_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_t^T (r_u + \gamma_u) du} \mid \mathcal{F}_t \right].$$

(3.14)

Note that, if the short-term interest rate process $r$ is independent of the default intensity process $\gamma$, the price of the defaultable zero-coupon bond is given by

$$\bar{B}(t, T) = 1_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_t^T r_udu} \mid \mathcal{F}_t \right] \mathbb{E}^Q \left[ e^{-\int_t^T \gamma_u du} \mid \mathcal{F}_t \right] = 1_{\{\tau > t\}} B(t, T) \mathbb{E}^Q \left[ e^{-\int_t^T \gamma_u du} \mid \mathcal{F}_t \right],$$

where $B(t, T)$ is the risk-free zero-coupon bond price at time $t$.

**Recovery of Treasury**

The recovery of treasury assumption states that the recovery payment is a fractional amount of the market value of an equivalent default-free asset. In the zero-coupon bond case, this thus means that the recovery value of a defaultable zero-coupon bond is given by a fraction of the value of an equivalent default-free bond. Of course, in general, an equivalent default-free asset does not necessarily have to exists, it is sufficient if the price can be determined.

---

5 Although this research is about CDSs, we will illustrate the different recovery rules with pricing formulae for zero-coupon bonds here. The pricing of a CDS will be postponed until chapter 5.
3.3. Pricing of Defaultable Contingent Claims

Let \( \bar{P}_{RT}(t) \) be the price of a defaultable claim under the recovery of treasury assumption. Let us furthermore denote the price of a defaultable claim under the zero recovery rule by \( \bar{P}(t) \) and the price of the equivalent default-free asset by \( P(t) \). We assume that, in the case of default before maturity, one recovers a fraction of \( c \) times the value of the default-free claim. We then get that the price of the defaultable claim under the RT assumption is given by

\[
\bar{P}_{RT}(t) = (1 - c) \bar{P}(t) + cP(t).
\]

This is easy to see, since in all cases we get at least \( c \) times the promised payoff. The other part (a fraction of \( 1 - c \)) is only paid in survival. This immediately shows the main advantage of the RT assumption: we can incorporate positive recovery, without incurring additional complexity in the problem. We only need the price of an equivalent default-free asset (which in most cases is straightforward to compute), and the price of the claim under a zero recovery assumption.

The price of a defaultable zero-coupon bond under the RT assumption is given by

\[
\bar{B}_{RT}(t, T) = (1 - c) \bar{B}(t, T) + cB(t, T),
\]

where \( \bar{B}(t, T) \) and \( B(t, T) \) are as described above. Of course, in a more general set-up, the fraction \( c \) can be made time-dependent or even stochastic.

Recovery of Par

The (fractional) recovery of par assumption states that the recovery payment is a fractional amount of a 'legal claim' \( V \), which can be defined at all times. This legal claim follows naturally in most situations. For example, in the cases of a (zero-)coupon bond and a credit default swap, this legal claim will be the par (or face) value of the contract.

The idea behind this recovery rule is that of a liquidation under the supervision of a bankruptcy court. In most cases, the recovery payments investors receive, are distributed in terms of a fraction of a legal claim. In this case, claimants of the same seniority class will receive the same proportion of their claim as payoff.

The pricing of a defaultable zero-coupon bond under the assumption of recovery of par is described in Duffie (1998) [19]. He describes the valuation under discrete-time recovery and continuous-time recovery. In the discrete-time recovery, it is assumed that the recovery payment is received at the first date after default among a pre-specified list of times \( T_1, T_2, \ldots, T_n = T \), with \( T_i < T_{i+1} \), and \( T \) the maturity time. The recovery fraction of the par value is considered to be a random variable \( W \). Under the assumption that \( W \) is independent of the short-term interest rate process \( r \) and the default time \( \tau \), and \( E[\bar{W}] = \bar{w} \), we get that the price of a defaultable zero-coupon bond is given by
\[ \tilde{B}_{RP}(t, T) = \mathbb{1}_{\{\tau > t\}} \left( V^*(t) + \sum_{\{i : t \leq T_i \leq T\}} V_i(t) \right), \tag{3.16} \]

where

\[ V^*(t) = \mathbb{E}^Q \left[ \mathbb{1}_{\{\tau > T\}} e^{-\int_t^T r_u du} \bigg| \mathcal{F}_t \right] = \mathbb{E}^Q \left[ e^{-\int_t^T (r_u + \gamma_u) du} \bigg| \mathcal{F}_t \right], \]

and

\[ V_i(t) = \mathbb{E}^Q \left[ \mathbb{1}_{\{T_{i-1} \leq \tau < T_i\}} e^{-\int_t^{T_i} r_u du} W \bigg| \mathcal{F}_t \right] = \mathbb{E}^Q \left[ e^{-\int_t^{T_i} (r_u + \gamma_u) du} \bigg| \mathcal{F}_t \right] - \mathbb{E}^Q \left[ \mathbb{1}_{\{\tau > T_i\}} e^{-\int_t^{T_i} r_u du} W \bigg| \mathcal{F}_t \right] = \bar{w} \left( \mathbb{E}^Q \left[ e^{-\int_t^{T_{i-1}} (r_u + \gamma_u) du} - \int_t^{T_{i-1}} r_u du \bigg| \mathcal{F}_t \right] - \mathbb{E}^Q \left[ e^{-\int_t^{T_i} (r_u + \gamma_u) du} \bigg| \mathcal{F}_t \right] \right). \]

In the continuous-time recovery case, we can view \( W \) as a stochastic process, where \( W_\tau \) is the fraction of par recovered at default. We can use formula (3.13) with \( Z_u = \bar{W}_u \), where \( \bar{W}_u \) is the risk-neutral compensator of \( W \). Intuitively, \( \bar{W}_t \) is the expected recovery in the next time-instance \( dt \), given \( \mathcal{F}_t \) (Duffie, 1998 [19]). Dropping the assumption that \( W \) is independent from \( r \) and \( \tau \), we get

\[ \tilde{B}_{RP}(t, T) = \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \int_t^T e^{-\int_t^u (r_v + \gamma_v) dv} \gamma_u W_{u} du \bigg| \mathcal{F}_t \right] + \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_t^T (r_u + \gamma_u) du} \bigg| \mathcal{F}_t \right]. \tag{3.17} \]

**Recovery of Market Value**

The recovery of market value (RMV) assumption states that if the default happens at \( t = \tau \), the value of the recovery payment of an asset is given by a fraction of its pre-default value. Following Duffie and Singleton (1997,1999) [24] [25], we have that the recovery process satisfies \( Z_t = K_t S_{t-} \), where \( K_t \) is a given \( \mathbb{F} \)-predictable process, representing the recovery rate, and \( S \) is the pre-default value of the defaultable claim at time \( t- \). We have the following, rather elegant, pricing formula for a defaultable claim (Bielecki & Rutkowski, 2002 [6]):
3.4. Change of Probability Measure

Proposition 3.14. Let the pre-default price of a defaultable claim be given by formula (3.13), with recovery process \( Z_t = K_t S_t \) for some \( \mathbb{F} \)-predictable process \( K \). Then \( S_t \) is given by

\[
S_t = 1_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_t^T (r_u + (1-K_u)\gamma_u) du} X \mid \mathcal{F}_t \right].
\] (3.18)

The price of a defaultable zero-coupon bond under the RMV assumption is then clearly given by

\[
\bar{B}_{RMV}(t, T) = 1_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_t^T (r_u + (1-K_u)\gamma_u) du} \mid \mathcal{F}_t \right].
\] (3.19)

In the RMV case, we, in general, thus have nicer pricing formulas than in the RP case. In the latter case, we have to sum (or integrate) over all possible default times in our pricing formula, which makes the formulas computationally more demanding. A disadvantage of the RMV model is, however, that we cannot calibrate the default intensity \( \gamma_t \) apart from the \((1 - K_t)\)-term. So we can only make statements about the ‘expected loss’ and not about the probability of default itself. In the RP model, on the other hand, we can calibrate the default intensity separately and, therefore, this approach is more suitable for risk management purposes. This, in combination with the fact that CDS contracts already specify a recovery payment similar to recovery of par, makes us choose the RP rule in our credit-liquidity model in chapter 5.

3.4 Change of Probability Measure

For pricing purposes, one works under a risk-neutral measure \( \mathbb{Q} \) that is equivalent to the objective or real-world probability measure \( \mathbb{P} \). Usually, one models the dynamics of the default intensity process under \( \mathbb{Q} \), and by calibrating model-implied pricing formulas to observed market prices, one obtains the risk-neutral behavior of the intensity process. For risk management purposes, however, one would like to know the real-world default probabilities, and hence, the real-world behavior of the intensity process. In this section we will, therefore, discuss the issue of measure changes within intensity-based models.

3.4.1 Default Risk Compensation for Investors

Although we have seen the analogy between intensity-based models of credit risk and default-free short-term interest rate models, the issue of measure changes is not so straightforward in the former type of models as it is in the latter. In standard short-rate interest rate models, the short-rate is modeled as a diffusion process under \( \mathbb{Q} \) and, via Girsanov’s theorem, a change
Chapter 3. Intensity-Based Models

to the equivalent real-world probability measure $\mathbb{P}$ induces a change in the drift term of the stochastic differential equation associated with the short-rate process $r$. This implies that the process $r$ takes on the same values under $\mathbb{P}$ as under $\mathbb{Q}$, but that their dynamics under both measures differ. This change of dynamics reflects the risk premium investors demand for changes in interest rate risk.

When considering models, such as intensity-based models, that incorporate default risk, one also has to think about the compensation that investors demand for bearing this default risk. Jarrow, Lando and Yu (2005) [38] show that the compensation for the default risk can be split into two parts: first, investors require compensation for the variation in default risk over time, which they call the default risk premium. Second, investors require compensation for the (timing of the) default event itself, which they call the default event risk premium.

The default risk premium is, just as in short-rate models, captured by a drift change in the process dynamics. The default event risk premium, on the other hand, is captured by a difference between the values of the default intensity process under $\mathbb{P}$ and under $\mathbb{Q}$ (i.e., $\lambda_t$ takes on different values under $\mathbb{P}$ and under $\mathbb{Q}$). For notational convenience, we will denote the default intensity process under $\mathbb{P}$ by $\lambda^P$ and the default intensity process under $\mathbb{Q}$ by $\lambda^Q$. The existence of a default event risk premium thus implies that $\lambda^P_t \neq \lambda^Q_t$ for all $t \in \mathbb{R}_+$.

Before explaining these notions further, let us see what the consequences of this extra default event risk premium are. Suppose that we model the $\mathbb{P}$-default intensity as a Cox-Ingersoll-Ross (CIR) process under $\mathbb{P}$. We then have

$$d\lambda^P_t = (\alpha - \beta \lambda^P_t)dt + \sigma \sqrt{\lambda^P_t} dW^P_t.$$ 

Suppose now that we have an equivalent measure $\mathbb{Q}$, and let us assume that $W^P$ and $W^Q$ are related through $dW^P_t = dW^Q_t - \frac{\nu}{\sigma} \sqrt{\lambda^P_t} dt$. We then get the following behavior of $\lambda^P$ under $\mathbb{Q}$:

$$d\lambda^P_t = (\alpha - (\beta + \nu) \lambda^P_t)dt + \sigma \sqrt{\lambda^P_t} dW^Q_t.$$ 

We see that we still denote the process by $\lambda^P$, but that we changed to a $\mathbb{Q}$-Brownian motion. We will, therefore, call this the Q-dynamics of the $\mathbb{P}$-intensity process. Suppose now that the risk premium of the default event itself is given by some positive constant factor $\mu$, such that $\lambda^Q_t = \mu \lambda^P_t$ (i.e., $d\lambda^Q_t = \exp \left( \int_t^T \log \left( \frac{\lambda^Q_s}{\lambda^P_s} \right) dH_s - \int_t^T \lambda^P_s \left( \frac{\lambda^Q_s}{\lambda^P_s} - 1 \right) ds \right)$). We then get the following Q-dynamics of the $\mathbb{Q}$-intensity process:

$$d\lambda^Q_t = (\mu \alpha - (\beta + \nu) \lambda^Q_t)dt + \sigma \sqrt{\mu} \sqrt{\lambda^Q_t} dW^Q_t.$$ 

Lastly, we can also have the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-intensity process: 36
### 3.4. Change of Probability Measure

\[ d\lambda_t^Q = (\mu \alpha - \beta \lambda_t^Q)dt + \sigma \sqrt{\mu} \lambda_t^Q dW_t^P. \]

In total, we thus have four different possibilities. We see that if \( \mu = 1 \) in the above, the \( Q \)-intensity and the \( P \)-intensity processes are the same (\( \lambda_t^Q = \lambda_t^P \)), but the dynamics under \( Q \) and \( P \) differ, because of the drift change. In this case we have a full analogy with the short-rate models. If, on the other hand, \( \mu \neq 1 \), not only the drift dynamics under both measures differ, but the processes themselves too (\( \lambda_t^Q \neq \lambda_t^P \)).

The question that now arises is if it is, in general, likely that \( \mu \neq 1 \)? If this is not the case, we can just use the machinery of the short-rate models. If it is the case, however, we will have to treat measure changes differently from short-rate models. Jarrow, Lando and Yu (2005) \[38\] show that, using empirical results obtained in Duffee (1999) \[18\], it is most likely that there is a risk premium on the default process itself. They also give the following stylized example to give more intuition on the risk premium on the default process itself:

**Example 1.** Consider a firm whose \( P \)-default intensity is given by \( \lambda_t^P \). We assume a risk-free short-rate \( r(t) \), and we assume that the firm issues a bond promising to pay a continuous coupon flow of \( r(t) + \lambda_t^P(u) \) up to a maturity date \( T \) and a lump sum payment of 1 at \( T \). Furthermore, we assume that there is no recovery at default. The coupon payments of the bond are thus continuously adjusted by a step-up provision that adjusts the coupon to reflect the instantaneous default intensity under the objective measure.

First, suppose that there is no risk premium on the default event itself. We then have that \( Q \)-intensity process is the same as the \( P \)-intensity process, but the dynamics of the intensity processes differ under both measures as they have different drifts. We get that the price of the claim is given by

\[
P(0, T) = \mathbb{E}^Q \left[ \int_0^T (r(t) + \lambda_t^P(u)) du \right] + \mathbb{E}^Q \left[ e^{-\int_0^T (r(t)+\lambda_t^Q(u)) du} \right]
\]

\[
= \mathbb{E}^Q \left[ \int_0^T (r(t) + \lambda_t^Q(t)) e^{-\int_0^t (r(u)+\lambda_t^Q(u)) du} dt \right] + \mathbb{E}^Q \left[ e^{-\int_0^T (r(u)+\lambda_t^Q(u)) du} \right]
\]

\[
= \mathbb{E}^Q \left[ \int_0^T -\frac{\partial}{\partial t} e^{-\int_0^t (r(u)+\lambda_t^Q(u)) du} dt \right] + \mathbb{E}^Q \left[ e^{-\int_0^T (r(u)+\lambda_t^Q(u)) du} \right]
\]

\[
= \mathbb{E}^Q \left[ -e^{-\int_0^T (r(u)+\lambda_t^Q(u)) du} \bigg|_{t=0}^{t=T} \right] + \mathbb{E}^Q \left[ e^{-\int_0^T (r(u)+\lambda_t^Q(u)) du} \right]
\]

\[
= 1 - \mathbb{E}^Q \left[ e^{-\int_0^T (r(u)+\lambda_t^Q(u)) du} \right] + \mathbb{E}^Q \left[ e^{-\int_0^T (r(u)+\lambda_t^Q(u)) du} \right]
\]

\[
= 1.
\]
So regardless of how the drift is changed in the $Q$-intensity, the price of the above product will be 1. We see that the payment of the objective default intensity is enough to compensate for the default risk, no matter how risk-averse the agents are with respect to changes in default risk.

If, on the other hand, we assume that there is a risk premium for the default event, say $\lambda^Q = \mu \lambda^P$ for some positive constant $\mu$ ($\mu > 1$ if agents are risk-averse), then the $Q$-intensity is a different process than the $P$-intensity and we get that the price of the claim is given by

$$P(0, T) = \mathbb{E}^Q \left[ \int_0^T (r(t) + \lambda^P(t))e^{-\int_0^t (r(u) + \mu \lambda^P(u))du}dt \right] + \mathbb{E}^Q \left[ e^{-\int_0^T (r(u) + \mu \lambda^P(u))du} \right].$$

We see that the price is decreasing in $\mu$, meaning that the more risk-averse the agents are toward the default event, the less they are willing to pay for a claim that steps up the coupon payments by an amount equal to the objective default intensity.

The above discussion is still a bit heuristic on the meaning of a default risk premium and a default event risk premium and, therefore, following Yu (2002) [60], we will show that these risk premia can be explained by looking at the expected returns.

Let us first consider the pricing of a zero-coupon bond in which the short-rate is modelled as a CIR process under the real-world probability measure $\mathbb{P}$, that is,

$$dr_t = \kappa(\theta - r_t)dt + \sigma \sqrt{r_t}dW^P_t,$$

and where the risk-neutral short-rate process is assumed to be given by

$$dr_t = (\kappa + \nu) \left( \frac{\kappa \theta}{\kappa + \nu} - r_t \right) dt + \sigma \sqrt{r_t}dW^Q_t,$$

where $\kappa$, $\theta$, $\nu$ and $\sigma$ are constants and $W^P$ and $W^Q$ are $\mathbb{P}$- and $\mathbb{Q}$-Brownian motions, respectively (i.e., the change of measure is given by $d\mathbb{P} = E^P \left( e^{\int_0^T \sigma \sqrt{r_s}dW^Q_s} \right)$). It can be shown (this is the topic of chapter 4) that the zero-coupon bond price is given by

$$P(t, T) = e^{-A(\theta, \kappa, \nu, \sigma, T-t)-B(\theta, \kappa, \nu, \sigma, T-t)r_t},$$

where, without going into details here, one can show that the functions $A$ and $B$ satisfy a certain system of ordinary differential equations and can be solved analytically in terms of the
3.4. Change of Probability Measure

process parameters and the time till maturity (note that, for notational convenience, we will drop the function arguments of $A$ and $B$).

By applying Itô’s lemma to the above expression, we get the following dynamics of $P(t, T)$:

\[
\frac{dP(t, T)}{P(t, T)} = (r_t + B\nu r_t)dt - B\sigma\sqrt{r_t}dW^p_t.
\]

The instantaneous expected return is thus

\[
r_t + B\nu
\]

and $\nu$ can be interpreted as the market price of interest rate risk.

Let us now consider the situation of a defaultable zero-coupon bond and let us assume a recovery of market value rule. From Proposition 3.14, we know that the price of this bond can be computed in a similar way as the price of a default-free bond with an adjusted short-rate $R_t = r_t + (1 - K_t)\lambda^Q_t = r_t + s^Q_t$, where $r_t$ is the standard short-rate and $s^Q_t$ is the risk-neutral mean loss-rate. Let us assume that the dynamics of $s^Q_t$ are given by

\[
ds^Q_t = (\kappa' + \nu') \left( \frac{\kappa' \theta'}{\kappa' + \nu'} - s^Q_t \right) dt + \sigma' \sqrt{s^Q_t} dW^Q_t,
\]

and, furthermore, let us assume that the $P$-dynamics of $s^Q_t$ are given by

\[
ds^Q_t = \kappa'(\theta' - s^Q_t) dt + \sigma' \sqrt{s^Q_t} dW^P_t,
\]

where $W^P$ is a $P$-Brownian motion that is independent of $W^P$ and $\tilde{W}^Q$ is a $Q$-Brownian motion that is independent of $W^Q$. The defaultable zero-coupon bond price is now given by

\[
\bar{P}(t, T) = e^{-A - A'T - Br_t - B's^Q_t},
\]

where $A, A', B$ and $B'$ are again functions satisfying systems of ordinary differential equations, which we will not consider here.

By similar arguments as above, it is tempting to conclude that the instantaneous expected return is given by

\[
r_t + s^Q_t + B\nu r_t + B'\nu' s^Q_t,
\]

since the bond price dynamics are given by
\[
\frac{d\bar{P}(t,T)}{\bar{P}(t,T)} = \left( r_t + s_t^Q + B\nu r_t + B'\nu' s_t^Q \right) dt - B\sigma\sqrt{r_t}dW_t - B'\sigma'\sqrt{s_t^Q}dW'.
\] (3.22)

This reasoning, however, is flawed, since if we assume that the market prices of risk are zero, i.e., \( \nu = \nu' = 0 \), it follows from (3.21) that the expected return is \( r_t + s_t^Q \). We would, however, expect an instantaneous expected return of \( r_t \), since we are pricing risk-neutrally.

To see what is missing, we give the following heuristic argument. Let \( \lambda^P \) be the default intensity process under \( P \). The likelihood of a default event in the time-interval \( (t, t + \Delta t) \) is then given by \( \lambda^P_t \Delta t \). The expected return of the bond on this interval is thus given by

\[
\text{Expected return} = (1 - \lambda^P_t \Delta t) \cdot \frac{\Delta \bar{P}(t,T)}{\bar{P}(t,T)} \bigg|_{\text{no default}} + \lambda^P_t \Delta t \cdot \frac{\Delta \bar{P}(t,T)}{\bar{P}(t,T)} \bigg|_{\text{default}}
\]

\[
= (1 - \lambda^P_t \Delta t) \cdot \frac{\Delta \bar{P}(t,T)}{\bar{P}(t,T)} \bigg|_{\text{no default}} + K_t\bar{P}(t-,T) - \bar{P}(t-,T) \lambda^P_t \Delta t
\]

\[
\approx \frac{\Delta \bar{P}(t,T)}{\bar{P}(t,T)} \bigg|_{\text{no default}} - (1 - K_t)\lambda^P_t \Delta t.
\]

If we now let \( \Delta t \to 0 \), we have that the first term in the last equation gives (3.21). The full instantaneous expected return is thus given by

\[
r_t + B\nu r_t + B'\nu' s_t^Q + s_t^Q - s_t^P,
\] (3.23)

where \( s_t^P = \lambda^P_t(1 - K_t) \) is the mean loss-rate under the physical measure.

We thus see that the instantaneous expected return consists of 3 components. The first component is the return on an otherwise equivalent default-free bond (the first two terms in (3.23)). The second component consists of the third term in (3.23) and can be viewed as the compensation for variation in default risk. Even if \( \lambda^Q = \lambda^P \) (and hence \( s^Q = s^P \)), the second term needs to be accounted for. The parameter \( \nu' \) can thus be interpreted as a risk premium on default risk.

In the situation that the default event risk cannot be diversified away and, hence, \( \lambda^P \neq \lambda^Q \), one has to consider the third component, which consists of \( s_t^Q - s_t^P \). As we will show below, we have that intensities of a default process under a change of measure are related through \( \lambda_t^Q = \mu_t \lambda_t^P \). The compensation for the default event thus becomes
3.4. Change of Probability Measure

We can thus interpreted $\mu_t$ as the risk premium on the default event.

We will now show that the $\mathbb{P}$- and $\mathbb{Q}$-default intensities are related through a process $\mu_t$ such that $\lambda_t^Q = \mu_t \lambda_t^P$. Let us assume that we have a probability measure $\mathbb{Q}$, equivalent to the objective measure $\mathbb{P}$ on $(\Omega, \mathcal{G}_T)$ for some fixed $T > 0$. Let us furthermore assume that

- Conditions 1 and 2 hold,
- the filtration $\mathbb{F} = \mathbb{F}^W$ is generated by a Brownian motion,
- the $\mathbb{F}$-hazard process $\Gamma$ of $\tau$ is a continuous, increasing process under $\mathbb{Q}$.

Under these assumptions, we have by Proposition 3.3 that the $\mathbb{F}$-hazard process is equal to the $\mathbb{F}$-martingale hazard process under $\mathbb{P}$, and that the Brownian motion remains a Brownian motion with respect to the enlarged filtration $\mathbb{G}$ (by the martingale invariance property).

Let the Radon-Nikodým density process $\eta_t, t \leq T$, be given by

$$\eta_t := \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{G}_t} = \mathbb{E}^\mathbb{P}[X | \mathcal{G}_t], \quad \mathbb{P}\text{-a.s., } t \leq T, \quad (3.24)$$

where $X$ is $\mathcal{G}_T$-measurable, integrable, random variable, such that $\mathbb{P}\{X > 0\} = 1$ and $\mathbb{E}^\mathbb{P}[X] = 1$. Obviously, the process $\eta$ follows a $\mathbb{G}$-martingale. The following result gives a martingale representation under the above-mentioned conditions:

**Proposition 3.15.** Under the conditions stated in this section, we have for any $\mathbb{G}$-martingale $N$ that we can write

$$N = N_0 + \int_0^t \xi_u^N dW_u + \int_{[0,t]} \zeta_u^N d\hat{M}_u,$$

where $\xi^N$ and $\zeta^N$ are $\mathbb{G}$-predictable stochastic processes and $\hat{M}_t = H_t - \Gamma_{t\wedge \tau}$.

**Proof.** See Corollary A.14 in Appendix A. \hfill \Box

In light of the above result, we can write the Radon-Nikodým process $\eta$ as follows:
\[ \eta_t = 1 + \int_0^t \xi_u dW_u + \int_{[0,t]} \zeta_u d\hat{M}_u, \]

where \( \xi \) and \( \zeta \) are \( \mathbb{G} \)-predictable processes and \( \hat{M}_t = H_t - \Gamma_{t \wedge \tau} \).

A famous result by Jacod (1979) \cite{Jacod1979}, tells us that for a strictly positive martingale, the process of left-hand limits is strictly positive as well, and therefore we can rewrite the above formula as follows:

\[ \eta_t = 1 + \int_{[0,t]} \eta_u - \xi_u \eta_t - \zeta_u \eta_t - dW_u + \int_{[0,t]} \eta_u - (\beta_u dW_u + k_u d\hat{M}_u), \quad (3.25) \]

where \( \beta_t = \frac{\xi_t}{\eta_t} \) and \( k_t = \frac{\zeta_t}{\eta_t} \) are \( \mathbb{G} \)-predictable processes. We are now able to prove the following result:

**Proposition 3.16.** Let \( \mathbb{Q} \) be an equivalent probability measure to \( \mathbb{P} \) on \((\Omega, \mathcal{F}_T)\). Let the Radon-Nikodým density process be given by (3.24) with \( \eta \) satisfying (3.25). Then the process

\[ W^*_t = W_t - \int_0^t \beta_u du, \quad \forall \ t \in [0,T], \quad (3.26) \]

follows a Brownian motion with respect to \( \mathbb{G} \) under \( \mathbb{Q} \), and the process \( M^*_t, t \in [0,T] \), given by

\[ M^*_t = \hat{M}_t - \int_0^{t \wedge \tau} k_u d\Gamma_u = H_t - \int_0^{t \wedge \tau} (1 + k_u) d\Gamma_u, \quad (3.27) \]

follows a \( \mathbb{G} \)-martingale under \( \mathbb{Q} \). Furthermore, the \( \mathbb{G} \)-martingales \( W^* \) and \( M^* \) are mutually orthogonal under \( \mathbb{Q} \).

**Proof.** We will first show that \( W^* \) is a Brownian motion under \( \mathbb{Q} \). We will start by showing that \( \eta W^* \) is a \( \mathbb{G} \)-(local)-martingale under \( \mathbb{P} \) and consequently \( W^* \) is a \( \mathbb{G} \)-(local)-martingale under \( \mathbb{Q} \). For \( t \leq T \), we have by Itô’s lemma

\[ d(\eta_t W^*_t) = W^*_t d\eta_t + \eta_t dW^*_t + d[W^*, \eta]_t. \]

We note that \( \eta_t dW^*_t = \eta_t dW_t + \eta_t \beta_t dt \) and
3.4. Change of Probability Measure

\[ d(W^*, \eta)_t = d \left[ W_t - \int_0^t \beta_u du, 1 + \int_0^t \eta_u - (\beta_u dW_u + k_u d\hat{M}_u) \right] \]

\[ = d \left[ W_t, \int_0^t \eta_u - \beta_u dW_u \right] \]

\[ = \eta_t - \beta_t d[W, W]_t \]

\[ = \eta_t - \beta_t dt. \]

We conclude that \( d(\eta_t W^*_t) = W^*_t d\eta_t + \eta_t dW_t \) is a \( \mathcal{G} \)-(local)-martingale under \( \mathbb{P} \), and hence \( W^* \) is a \( \mathcal{G} \)-(local)-martingale under \( \mathbb{Q} \). The quadratic covariation \( [W^*, W^*_t] = [W, W]_t = t \) under \( \mathbb{Q} \), and we have that \( W^* \) is continuous. By Levy’s characterization of the Brownian motion (Spreij, 2012, proposition 7.14 [53]), we conclude that \( W^* \) is a Brownian motion under \( \mathbb{Q} \).

In a similar way we show that \( d(\eta_t M^*_t) \) is a \( \mathcal{G} \)-martingale under \( \mathbb{P} \) and, consequently, \( M^* \) is a \( \mathcal{G} \)-martingale under \( \mathbb{Q} \). We get

\[ d(\eta_t M^*_t) = M^*_t d\eta_t + \eta_t dM^*_t + d[M^*, \eta]_t. \]

We note that \( \eta_t - dM^*_t = \eta_t - d\hat{M}_t - \eta_t - k_t d\Gamma_{t \wedge \tau} \) and

\[ d[M^*, \eta]_t = d \left[ H_t - \int_0^t (1 + k_u) d\Gamma_{t \wedge \tau}, 1 + \int_0^t \eta_u - (\beta_u dW_u + k_u d\hat{M}_u) \right] \]

\[ = d \left[ H_t, \int_0^t \eta_u - k_u d\hat{M}_u \right] \]

\[ = d \left[ H_t, \int_0^t \eta_u - k_u d(H_t - \Gamma_{t \wedge \tau}) \right] \]

\[ = \eta_t - k_u d[H, H]_t \]

\[ = \eta_t - k_u dH_t. \]

We get \( d(\eta_t M^*_t) = M^*_t d\eta_t + \eta_t - d\hat{M}_t - \eta_t - k_t d\Gamma_{t \wedge \tau} + \eta_u - k_u dH_t \), and rearranging gives

\[ d(\eta_t M^*_t) = M^*_t d\eta_t + \eta_t - (1 + k_t) d\hat{M}_t, \]

which is a \( \mathcal{G} \)-martingale under \( \mathbb{P} \), and therefore it follows that \( M^* \) is a \( \mathcal{G} \)-martingale under \( \mathbb{Q} \). The orthogonality of \( W^* \) and \( M^* \) under \( \mathbb{Q} \) follows from Itô’s lemma and the fact that \( [W^*, M^*] = 0 \) (\( W^* \) is continuous and \( M^* \) is a process of finite variation).
Chapter 3. Intensity-Based Models

In view of Proposition 3.16, it is tempting to conjecture that the $\mathbb{F}$-hazard process $\Gamma^*$ of $\tau$ under $\mathbb{Q}$ is given by $\Gamma^*_t = \int_0^t (1 + k_u) d\Gamma^*_u = \int_0^T (1 + \tilde{k}_u) d\Gamma^*_u$ (where $\tilde{k}_u = k_{u \wedge \tau}$). We would also like to conclude that for a random time $\tau$ that admits an $\mathbb{F}$-intensity process $\gamma$ under $\mathbb{P}$, the $\mathbb{F}$-intensity process $\gamma^*$ under $\mathbb{Q}$ is given by $\gamma^*_t = (1 + \tilde{k}_t) \gamma_t$. In general, however, these relations are not valid for the $\mathbb{F}$-hazard processes (see for a counterexample Kusuoka (1999) [43]). If, however, we make the additional assumption that the process $\kappa$ is $\mathbb{F}$-adapted, then these statements are true for the $\mathbb{F}$-martingale hazard process.

Fortunately, in most practical applications, we indeed have that the filtration $\mathbb{F}$ is generated by a Brownian motion, and that there is equivalence between the hazard process and the martingale hazard process. Furthermore, by the canonical construction of the default time, the other assumptions of this section hold and, therefore, we can conclude that in most practical applications the default intensity processes are related through $\lambda^Q_t = (1 + \tilde{k}_t) \lambda^P_t = \mu_t \lambda^P_t$.

### 3.4.2 Rabobank’s Model for the Default Event Risk Premium

The Rabobank already has a model to compute the risk premium on the default event for a specific country. We will not go into details on this model, but we briefly want to mention the existence of it, since we will combine the outcomes of our model with this model in order to compute the actual default probabilities.

The Rabobank has formulated a theoretical model and argues that the $\mathbb{P}$-default intensity and $\mathbb{Q}$-default intensity are related as follows:

$$\lambda^P_t = K \lambda^Q_t,$$

where $K$ is a (country-specific) constant. The constants $K$ are estimated from default data and global market data. In order to compute actual default probabilities, we will use the following relationship:

$$\mathbb{P}\{ \tau > T | \mathcal{F}_t \} = \mathbb{E}^P \left[ e^{- \int_t^T \lambda^P(s) ds} \bigg| \mathcal{F}_t \right] = \mathbb{E}^P \left[ e^{- \int_t^T K \lambda^Q(s) ds} \bigg| \mathcal{F}_t \right].$$

We see in the final term that the expectation is taken with respect to the real-world probability measure $\mathbb{P}$ and that the default intensity process is the $\mathbb{Q}$-default intensity process. In order to be able to compute this expression, we thus need to know the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-default intensity process (see the discussion at the start of this section). By calibrating time series of model-implied premia to observed premia, we will generate a (discrete) time series of the $\mathbb{Q}$-default intensity and from this time series we are able to extract the $\mathbb{P}$-dynamics by means of maximum likelihood estimation. This will be explained in more detail in chapters 5, 6 and 8.
3.5 Chapter Summary

In this chapter, we introduced the class of intensity-based models of default risk. This class of models does not try to explain default risk by economic fundamentals, but it just imposes that entities follow a certain default process in which the probability of default is influenced by the default intensity. Typically, this default intensity follows a stochastic (diffusion) process.

We started this chapter with a formal introduction to the notions of the hazard- and martingale hazard processes associated with a default process and we showed how these concepts are related to each other. Another important feature of intensity-based models that we discussed was the role of the filtrations within this setting. We showed the so-called “filtration switching formula”, which allows us to change the conditioning filtration in conditional expectations. This result is useful, since it allows us to compute certain expectations more easily, for example in the context of pricing of defaultable claims.

Furthermore, we showed how one can actually construct a default process that matches a given (martingale) hazard process, since one usually starts with modelling the hazard process. We discussed the so-called canonical approach and the notion of a doubly stochastic Poisson process. Both options lead to similar results in most applications and for our research the results are even the same.

We also showed how one can use intensity-based models to construct pricing formulas for defaultable claims. An important component of the pricing formulas is the specification of the recovery value in the case of default. We discussed the models that are standard in the literature and we argued that the recovery of par assumption is the most suitable for CDS pricing.

We ended this chapter with a discussion on the topic of measure changes, which is more complicated than in, for example, standard short-rate models. In models with default risk, investors typically do not only require risk premia for variations in default risk, but also for the default event itself. The risk premia on variations in default risk are called the default risk premia and are captured by a change in the drift term of the default intensity process. This is similar to the risk premia investors require on interest rate risk in short-rate models. The risk premia on the default event, however, is captured by a difference between the default intensity process under the real-world measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$. We illustrated how one can interpret these risk premia in terms of expected returns. Lastly, we mentioned the model that Rabobank uses to estimate the risk premium on the default intensity and we explained how our model will connect to Rabobank’s model.
4

Affine Processes

In chapter 3, we introduced the class of intensity-based models and we showed how these models can be used to price defaultable claims. We stated that the starting point of these models is the modeling of the (martingale) hazard process $\Lambda$, which is often given by $\Lambda_t = \int_0^t \lambda_u du$ for some non-negative, $\mathbb{F}$-progressively measurable process $\lambda$ (commonly denoted as the default intensity process). Until now, however, we did not specify how to model the hazard process, we merely derived some results for hazard processes in general.

We already noted the analogy between short-rate interest rate models and intensity-based models. Let us recall that (assuming that the default time admits a stochastic intensity process $\lambda$) the probability of survival to a future time $s$, given that there is no default until time $t < s$, is given by

$$\mathbb{P}\{\tau > s | \mathcal{G}_t\} = \mathbb{1}_{\\{\tau > t\}} \mathbb{E}_t^\mathbb{P}\left[ e^{-\int_t^s \lambda_u du} \bigg| \mathcal{F}_t \right]. \quad (4.1)$$

The evaluation of this expression is computationally equivalent to computing a zero-coupon bond price, if $\lambda$ is thought of as the short-rate process. From the short-rate interest rate literature (see for example Brigo and Mercurio (2006) [9]), we know that a convenient way to model the short-rate process is by means of an affine process, since then expressions of the form (4.1) can be computed analytically or by numerically solving a system of ordinary differential equations.

By modeling the hazard rate process as an affine process, we are thus able to compute (4.1) (and similar expressions) explicitly, which is also convenient for pricing defaultable claims. In this chapter, we will, therefore, introduce the class of affine processes and state some useful results. We, by no means, aim to give a complete overview on this topic, and for further
details and references we refer to Duffie, Filipović and Schachermayer (2003) [21] and Filipović (2009) [27].

4.1 Set-Up and Definitions

Affine processes are commonly used in financial models. Their analytical tractability and flexibility to capture empirical observations, like stochastic-volatility and (time-varying) mean-reversion, makes them very suitable for the modeling of financial processes (Duffie, Filipović and Schachermayer (2003) [21]). Especially since the pathbreaking studies of Vacićek (1977) [56] and Cox, Ingersoll and Ross (1985) [16], who assumed that the short-term interest rate follows a Gaussian and square-root diffusions, respectively, the use of affine processes in finance has gained a lot of attention. Affine processes are mainly used in models of term-structures and bond prices (see for example Duffie and Kan (1996) [22], and Dai and Singleton (2000) [17]), but their use in credit risk models has also increased (see for example Duffie (2005) [20], and Duffie and Singleton (1999) [25]).

To define an affine process, we consider the following set-up. Let \( W \) be a \( d \)-dimensional Brownian motion on the probability space \( \Omega, \mathcal{F}, \mathbb{P} \), and let \( X \subset \mathbb{R}^d \) be a closed state space with non-empty interior. Let us assume that for every \( x \in X \) there exists a unique solution of the stochastic differential equation

\[
\frac{dX_t}{dt} = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x,
\]

where \( b : \mathcal{X} \to \mathbb{R}^d \) is continuous, and \( \sigma : \mathcal{X} \to \mathbb{R}^{d\times d} \) is measurable and such that the diffusion matrix \( a(x) = \sigma(x)\sigma(x)^T \) is continuous in \( x \in \mathcal{X} \).

**Definition 4.1.** The process \( X \) satisfying (4.2) is called affine if for all \( t \leq T, \quad u \in i\mathbb{R}^d, \quad x \in \mathcal{X} \), its conditional characteristic function is of the form

\[
\mathbb{E} \left[ e^{u^T X_T} \bigg| \mathcal{F}_t \right] = e^{\phi(T-t,u) + \psi(T-t,u)^T X_t},
\]

where \( \phi(t, u) \) and \( \psi(t, u) \) are \( \mathbb{C} \)- and \( \mathbb{C}^d \)-valued functions, respectively.

We will focus specifically on the class of regular affine processes, since in this case the functions \( \phi \) and \( \psi \) are characterized by a system of ordinary differential equations. Roughly speaking, an affine process is regular, if the coefficients \( \phi(\cdot, u) \) and \( \psi(\cdot, u) \) of the characteristic function are differentiable and if their derivatives are continuous at zero (Duffie, 2005 [20]).

The following result gives necessary and sufficient conditions for a process \( X \) to be affine:

\[\text{Keller-Ressel, Schachermayer and Teichmann (2011) [40] show that all affine processes are also regular.}\]
4.1. Set-Up and Definitions

Theorem 4.2. Suppose $X$ is affine. Then the diffusion matrix $a(x)$ and drift $b(x)$ are affine in $x$, and are of the form

$$a(x) = a + \sum_{i=1}^{d} x_i \alpha_i, \quad b(x) = b + \sum_{i} x_i \beta_i = b + \mathcal{B} x, \quad (4.4)$$

where $a$ and $\alpha_i$ are $d \times d$-matrices, and $b$ and $\beta_i$ are $d$-vectors. We denote by $\mathcal{B}$ the $d \times d$ matrix with as $i$th column vector $\beta_i$. Moreover, $\phi$ and $\psi$ of (4.3) solve the following system of Riccati equations:

$$\partial_t \phi(t, u) = \frac{1}{2} \psi(t, u)^\top a \psi(t, u) + b^\top \psi(t, u), \quad \phi(0, u) = 0,$$

$$\partial_t \psi_i(t, u) = \frac{1}{2} \psi(t, u)^\top \alpha_i \psi(t, u) + \beta_i^\top \psi(t, u), \quad 1 \leq i \leq d, \quad \psi(0, u) = u. \quad (4.5)$$

Conversely, if the diffusion matrix $a(x)$ and drift $b(x)$ are of the form (4.4) and if there exists a solution $(\phi, \psi)$ of the Riccati equations (4.4) such that $\phi(t, u) + \psi(t, u)^\top x$ has a non-positive real part for all $t \geq 0, u \in i\mathbb{R}^d$ and $x \in \mathcal{X}$. Then $X$ is affine and its conditional characteristic function is of the form (4.3).

In the light of Theorem 4.2, we see that the parameters $a, \alpha_i, b$ and $\beta_i$ in (4.4) should be chosen in such a way that $X$ does not leave the set $\mathcal{X}$. Furthermore, since $a(x)$ is a diffusion matrix it should be symmetric and positive semi-definite for all $x \in \mathcal{X}$.

Until now, we did not specify $\mathcal{X}$, but from now on (unless stated differently), we will assume that $\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n$, where $m, n \geq 0$ and $m + n = d$. This state space is called the canonical state space in the literature and captures many applications in finance literature. For notational convenience, let us denote the index sets $I = \{1, \ldots, m\}$ and $J = \{m+1, \ldots, m+n\}$. For any vector $\mu$ and matrix $\nu$, and index sets $M, N$, we denote by $\mu_M = (\mu_i)_{i \in M}$ the sub-vector, and by $\nu_{MN} = (\nu_{ij})_{i \in M, j \in N}$ the sub-matrix. The following result gives admissibility conditions on the parameters of (4.4) (Filipović, 2009 [27]).

Theorem 4.3. The process $X$ on the canonical state space $\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n$ is affine if and only if $a(x)$ and $b(x)$ are of the form (4.4) for parameters $a, \alpha_i, b, \beta_i$ which are admissible in the following sense:

---

2A Riccati equation is a first-order ordinary differential equation that is quadratic in the unknown $\mathbb{R}$-valued function $y(x)$:

$$y'(x) = q_0(x) + q_1(x)y(x) + q_2(x)y^2(x).$$

---
Chapter 4. Affine Processes

$a, \alpha_i$ are symmetric positive semi-definite,

\[ a_{II} = 0 \quad (\text{and thus } a_{IJ} = a_{JI}^\top = 0). \]

\[ \alpha_j = 0 \quad \text{for all } j \in J, \]

\[ \alpha_{i,kl} = \alpha_{i,lk} = 0 \quad \text{for } k \in I \setminus i, \quad \text{for all } 1 \leq i, l \leq d, \]

\[ b \in \mathbb{R}_+^m \times \mathbb{R}^n, \]

\[ B_{IJ} = 0, \]

\[ B_{II} \] has non-negative off-diagonal elements.

(4.6)

In this case, the system of Riccati equations (4.5) simplifies to

\[ \partial_t \phi(t,u) = \frac{1}{2} \psi_J(t,u)^\top a_{JJ} \psi_J(t,u) + b^\top \psi(t,u), \quad \phi(0,u) = 0, \]

\[ \partial_t \psi_i(t,u) = \frac{1}{2} \psi(t,u)^\top \alpha_i \psi(t,u) + \beta_i^\top \psi(t,u), \quad i \in I, \]

\[ \partial_t \psi_J(t,u) = B_{JJ}^\top \psi_J(t,u), \]

\[ \psi(0,u) = u, \]

(4.7)

and there exists a unique global solution $(\phi(\cdot,u),\psi(\cdot,u)) : \mathbb{R}^+ \to \mathbb{C}_-^m \times i\mathbb{R}^n$ for all initial values $u \in \mathbb{C}_m \times i\mathbb{R}^n$.

Before we continue, we will give some examples of commonly-used affine processes.

**Example 2.** The Ornstein-Uhlenbeck (OU) process

\[ dX_t = (\alpha - \beta X_t)dt + \sigma dW_t \]  

is an affine process with $d = 1$ and $\mathcal{X} = \mathbb{R}$. Vaciček (1977) [56] used this process to model the short-term interest rate.

**Example 3.** The Feller process (CIR process)

\[ dX_t = \beta(\theta - X_t)dt + \sigma \sqrt{X_t}dW_t = (\alpha - \beta X_t)dt + \sigma \sqrt{X_t}dW_t \]

is an affine process with $d = 1$ and $\mathcal{X} = \mathbb{R}_+$. Cox, Ingersoll and Ross (1985) [16] were the first to use this process to model interest rates. Since then, this process is very popular in financial models, mainly because it lives on $\mathbb{R}_+$. Under the Feller condition $(\alpha > \frac{1}{2} \sigma^2)$, we even have that the process remains strictly positive (Schönbucher, 2003 [52]).
Example 4. The Arithmetic Brownian motion

\[ dX_t = \mu dt + \sigma dW_t \quad (4.10) \]

is an affine process with \( d = 1 \) and \( \mathcal{X} = \mathbb{R} \).

Example 5. Consider the following system of stochastic differential equations:

\[
\begin{align*}
\frac{dX^1_t}{dX^2_t} &= \left( r - \frac{1}{2} \beta X^2_t \right) dt + \sqrt{X^2_t} dW^1_t, \\
\frac{dX^2_t}{dX^2_t} &= \left( \theta - \kappa X^2_t \right) dt + \sigma \sqrt{X^2_t} dW^2_t,
\end{align*}
\]

where \( r \) is a constant and \( W^1 \) and \( W^2 \) are correlated Brownian motions with correlation coefficient \(-1 \leq \rho \leq 1\). We have that \((X^1, X^2)\) is affine with state space \( \mathbb{R} \times \mathbb{R}_+ \). Heston (1993) [32] used this set-up to model stock prices with stochastic volatility.

Example 6. The multivariate CIR process, where each factor follows an independent CIR process of the form (4.9) is also affine. In the bivariate case, for example, we have

\[
\begin{pmatrix}
\frac{dX^1_t}{dX^2_t} \\
\frac{dX^2_t}{dX^2_t}
\end{pmatrix} = \begin{pmatrix}
\alpha_1 & \beta_1 \\
\alpha_2 & \beta_2
\end{pmatrix} \begin{pmatrix}
X^1_t \\
X^2_t
\end{pmatrix} dt + \begin{pmatrix}
\sigma_1 & 0 \\
0 & \sigma_2
\end{pmatrix} \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix} \begin{pmatrix}
X^1_t \\
X^2_t
\end{pmatrix} \begin{pmatrix}dW^1_t \\
\sigma \sqrt{X^2_t} dW^2_t
\end{pmatrix}.
\]

(4.12)

4.2 Affine Processes and Bond Pricing

Apart from their flexibility in modeling, affine processes can also provide nice analytical solutions to pricing problems. Consider, for example, a risk-neutral pricing problem, in which we want to compute the discounted expected payoff of a \( T \)-claim with payoff of the form \( f(X(T)) \).

Assuming all integrability conditions are met, the price at time \( t \leq T \) of this claim is then given by

\[
\pi(t) = \mathbb{E} \left[ e^{-\int_t^T r(s) ds} f(X(T)) \bigg| F_t \right].
\]

(4.13)

In an affine set-up, we can, under some additional conditions, either find analytical or a numerically tractable formulas for (4.13). We will just focus on the computation of a zero-coupon bond price (that is \( f = 1 \) in (4.13)), since we will encounter expressions of this type later on.
Let us first consider the following “extended transform” of the affine process $X$ on $\mathbb{R}_+^m \times \mathbb{R}^n$ with admissible parameters $a, \alpha_i, b, \beta_i$ as given in (4.6):

$$
E \left[ e^{-\int_t^T r(s) \, ds} e^{u^\top X(T)} \, \bigg| \mathcal{F}_t \right],
$$

where

$$
r(t) = c + \gamma^\top X(t),
$$

with $c \in \mathbb{R}$ and $\gamma \in \mathbb{R}^d$. The following result is a special case of the more general result derived by Duffie, Pan and Singleton (2000) [23], who show the explicit solution of an even more general transform of processes of the more general class of affine jump-diffusion processes.

**Theorem 4.4.** Let $E \left[ e^{-\int_0^T r(s) \, ds} \right] < \infty$ for all $x \in \mathbb{R}_+^m \times \mathbb{R}^n$ and $0 < T < \tau$, where $r$ is given by (4.15). Then there exists a unique solution $(\phi(\cdot, u), \psi(\cdot, u)) : [0, \tau] \to \mathbb{C} \times \mathbb{C}^d$ of the following system of ordinary differential equations:

$$
\begin{align*}
\partial_t \phi(t, u) &= \frac{1}{2} \psi_J(t, u)^\top a_J \psi_J(t, u) + b^\top \psi(t, u) - c, \quad \phi(0, u) = 0, \\
\partial_t \psi_i(t, u) &= \frac{1}{2} \psi(t, u)^\top \alpha_i \psi(t, u) + \beta_i^\top \psi(t, u) - \gamma_i, \quad i \in I, \\
\partial_t \psi_J(t, u) &= B^\top J \psi_J(t, u) - \gamma_J, \\
\psi(0, u) &= u,
\end{align*}
$$

for $u = 0$. Moreover, (4.14) admits the following representation:

$$
E \left[ e^{-\int_t^T r(s) \, ds} e^{u^\top X(T)} \, \bigg| \mathcal{F}_t \right] = e^{\Phi(T-t, u) + \Psi(T-t, u)^\top X(t)}
$$

for all $x \in \mathbb{R}_+^m \times \mathbb{R}^d$ and $u \in \mathcal{S}(\mathcal{D}_R(\tau))^3$.

**Proof.** Let us consider the process $X' = (X, Y)$ on enlarged state space $\mathbb{R}_+^m \times \mathbb{R}^{n+1}$, where

$$
Y(t) = y + \int_0^t \left( c + \gamma^\top X(s) \right) ds, \quad y \in \mathbb{R}.
$$

\footnote{The set $\mathcal{D}_R(\tau)$ is the $\tau$-section of the maximal domain for the system of Riccati equations (4.16). The set $\mathcal{S}(\mathcal{D}_R(\tau))$ is the set of those elements in $\mathbb{C}^d$, whose real parts are in $\mathcal{D}_R(\tau)$ (see Filipović (2009) [27]).}

52
4.2. Affine Processes and Bond Pricing

We clearly see that, since $X$ is an affine diffusion process and $dY(t) = c + \gamma^T X(t)dt$, the process $X'$ is again a diffusion process with diffusion matrix $a' + \sum_{i \in I} x_i \alpha'_i$ and drift $b' + \mathcal{B}'x'$, where

$$a' = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \quad \alpha'_i = \begin{pmatrix} \alpha_i & 0 \\ 0 & 0 \end{pmatrix}, \quad b' = \begin{pmatrix} b \\ c \end{pmatrix}, \quad \mathcal{B}' = \begin{pmatrix} \mathcal{B} & 0 \\ \gamma^T & 0 \end{pmatrix}.$$ 

From Theorem 4.3, we see that the parameters $a', \alpha'_i, b', \beta'_i$ form admissible parameters, and, therefore, we can conclude that the process $X'$ is an affine diffusion process on $\mathbb{R}_+^m \times \mathbb{R}^{n+1}$.

Again by Theorem 4.3, we know that the system of Riccati equations satisfies (4.7). Let $u \in \mathbb{C}^m \times i\mathbb{R}^n$ and $v \in i\mathbb{R}$. We have $\psi'(t, u, v) \in \mathbb{C}^m \times i\mathbb{R}^{n+1}$. Let us denote the vector of the first $d$ entries of $\psi'(t, u, v)$ by $\psi'_{\{1, \ldots, d\}}(t, u, v)$ and the $(d+1)$th entry by $\psi'_{d+1}(t, u, v)$. Then (4.7) can be written as

$$\begin{align*}
\partial_t \phi'_i(t, u, v) &= \frac{1}{2} \psi'_{it}(t, u, v) a_{ij} \psi'_{jt}(t, u, v) + b^T \psi'_{\{1, \ldots, d\}}(t, u, v) + c \psi'_{d+1}(t, u, v), \\
\phi'(0, u, v) &= 0, \\
\partial_t \psi'_{i}(t, u, v) &= \frac{1}{2} \psi'_{it}(t, u, v) \alpha'_i \psi'_{jt}(t, u, v) + \beta^T \psi'_{\{1, \ldots, d\}}(t, u, v) + \gamma_i \psi'_{d+1}(t, u, v), \quad \forall i \in I, \\
\partial_t \psi'_{j}(t, u, v) &= \mathcal{B}'_{ij} \psi'_{j}(t, u, v) + \gamma_j \psi'_{d+1}(t, u, v), \\
\partial_t \psi'_{d+1}(t, u, v) &= 0, \\
\psi'(0, u, v) &= \begin{pmatrix} u \\ v \end{pmatrix}.
\end{align*}$$

(4.18)

By Theorem 4.3, we know that there exists a unique $(\mathbb{C}_- \times \mathbb{C}^m \times i\mathbb{R}^{n+1})$-valued solution $(\phi'(-, u, v), \psi'(-, u, v))$ of (4.18) for all $(u, v) \in \mathbb{C}^m \times i\mathbb{R}^n \times i\mathbb{R}$. Now we note that $\psi'_{d+1}(t, u, v) = v$ for all $t$, since its initial value is $v$ and its derivative with respect to $t$ is zero. If we now replace $\psi'_{d+1}(t, u, v)$ by $v$ in (4.18), then we have that the conditional characteristic function of $X'$ is given by

$$\mathbb{E} \left[ e^{u^T X(T) + vY(T)} \bigg| \mathcal{F}_t \right] = e^{\phi'_i(T-t, u, v) + \psi'_{\{1, \ldots, d\}}(T-t, u, v) + \gamma_i \psi'_{d+1}(T-t, u, v)} X(t)X(t) + v Y(t),$$

(4.19)

for all $(u, v) \in \mathbb{C}^m \times i\mathbb{R}^n \times i\mathbb{R}$ and $t \leq T$. Setting $v = -1$ and $\Phi(t, u) = \phi(t, u, -1)$ and $\Psi(t, u) = \psi'_{\{1, \ldots, d\}}(t, u, -1)$, we get that the left-hand side of (4.19) is given by
Chapter 4. Affine Processes

\[ \mathbb{E} \left[ e^{u^\top X(T) + v Y(T)} \bigg| \mathcal{F}_t \right] = \mathbb{E} \left[ e^{u^\top X(T) - \left( y + \int_0^T (c + \gamma^\top X(s)) ds \right)} \bigg| \mathcal{F}_t \right] \\
= \mathbb{E} \left[ e^{-\int_0^T r(s) ds e^{u^\top X(T) - y}} \bigg| \mathcal{F}_t \right]. \]

The right-hand side of (4.19) is given by

\[
e^{\psi(T-t,u,v)^\top X(t) + v Y(t)} = e^{\Phi(t,u) + \Psi(t,u) - Y(t)} = e^{\Phi(t,u) + \Psi(t,u) - \left( y + \int_0^T (c + \gamma^\top X(s)) ds \right)} = e^{\Phi(t,u) + \Psi(t,u) - \left( y + \int_0^T r(s) ds \right)}.
\]

We conclude that the result is proved. \[ \square \]

A straightforward consequence of Theorem 4.4 is the bond price formula for affine processes. This is a result we will use multiple times below (see for example Appendices D, E and F).

**Proposition 4.5.** Let \( X \) be a regular affine diffusion process on \( \mathbb{R}_+^m \times \mathbb{R}_+^n \) with admissible parameters as specified in Theorem 4.3. Let \( r \) be the short-rate process as defined in (4.15). Then we have that the \( T \)-bond price at time \( t \leq T \) is given by

\[
\mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T r(u) du} \bigg| \mathcal{F}_t \right] = e^{-A(T-t) - B(T-t)^\top X(t)}, \quad (4.20)
\]

where we set \( A(t) = -\Phi(t,0) \) and \( B(t) = -\Psi(t,0) \), with \( \Phi \) and \( \Psi \) as defined in Theorem 4.4.

### 4.3 Existence of Affine Processes

All results derived above assumed the existence of a unique solution of the stochastic differential equation (4.2) on some state space \( \mathcal{X} \subset \mathbb{R}^d \). In general, the matrix \( \sigma \) in (4.2) does not satisfy the Lipschitz-continuity condition in \( x \), and therefore, the existence of a unique solution is non-trivial (Filipović, 2009 [27]). By Duffie and Kan (1996) [22], we know that we can write (4.2), under some non-degeneracy conditions, as

\[
dX(t) = (aX(t) + b)dt + \Sigma D(X(t))dW(t), \quad (4.21)
\]
where \(a, \Sigma \in \mathbb{R}^{d \times d}, b \in \mathbb{R}^d\) and \(D\) is a \(d \times d\)-diagonal matrix with \(D_{ii}(x) = \sqrt{\alpha_i + \beta_i^\top x}\), with \(\alpha_i \in \mathbb{R}\) and \(\beta_i \in \mathbb{R}^d\). Duffie and Kan (1996) [22] note that the coefficient vectors \(\beta_i\) generate stochastic volatility. For a solution of (4.21) to exist, we need that all diagonal elements of \(D\) are non-negative. Let the domain of non-negative volatilities be denoted by

\[\mathcal{D} = \{x \in \mathbb{R}^d : \alpha_i + \beta_i^\top X \geq 0, 1 \leq i \leq d\}.\] (4.22)

Then any solution of (4.21) must be such that the process \(X\) remains in \(\mathcal{D}\). Duffie and Kan (1996) [22] provide the following conditions under which a unique solution exists:

**Condition 3.** For all \(1 \leq i \leq d\):

1. For all \(x\) such that \(D_{ii}(x) = 0\), the inequality \(\beta_i^\top (ax + b) > \frac{1}{2} \beta_i^\top \Sigma \Sigma^\top \beta_i\) holds true.
2. For all \(1 \leq j \leq d\), if \((\beta_j^\top \Sigma)^I_j \neq 0\), then \(v_i = kv_j\) for some positive scalar \(k\).

**Theorem 4.6.** Under Condition 3 there exists a unique strong solution to the stochastic differential equation (4.21) that stays in \(\mathcal{D}\), provided that \(X(0) \in \mathcal{D}\).

### 4.4 Application to Credit Risk Modeling

In the introduction of this chapter, we saw that we could compute the survival probability in an intensity-based model by means of an expression similar to that of a zero-coupon bond price. Therefore, it is convenient to model the default intensity \(\lambda(t)\) as

\[\lambda(t) = \lambda(X(t)) = \rho_0 + \rho^\top X(t) = \rho_0 + \sum_{i=1}^d \rho_i X_i(t),\] (4.23)

where \(\rho_0 \in \mathbb{R}, \rho = (\rho_1, \ldots, \rho_d) \in \mathbb{R}^d\) and \(X\) a regular affine diffusion process. In this case the survival probabilities can be calculated using Proposition 4.5.

Note that the default intensity is not well-defined if the process becomes negative. Therefore, we need to ensure that \(\lambda \geq 0\). In a one-factor model, we can simply define \(X\) as a CIR process and let the parameters \(\rho_0\) and \(\rho\) be non-negative. In multi-factor models, however, the non-negativity requirement puts strong restrictions on the state space and the parameters. To see this, note that, in the canonical state space set-up, the factors corresponding to the index set \(J = \{m + 1, \ldots, d\}\) can take values in \(\mathbb{R}\). Non-negativity can, in general, thus not be ensured in such a multi-factor model, unless we take the state space to be \(\mathbb{R}_+^d\) and let \(\rho_0\) and \(\rho\) be non-negative.
Considering the affine stochastic differential equation representation result (4.21), we can find another way to ensure the non-negativity of $\lambda$. Consider the open domain $\mathcal{C}$ of non-negative intensities:

$$\mathcal{C} = \{ x \in \mathbb{R}^d : \rho_0 + \sum_{i=1}^d \rho_i x_i \geq 0 \}. \quad (4.24)$$

A natural way to ensure $\lambda(t) \geq 0$ is to require that $\mathcal{C} \subset \mathcal{D}$, where $\mathcal{D}$ is defined in (4.22). So, if we choose positive $\theta_1, \ldots, \theta_d$, and if we set

$$\rho_0 = \sum_{i=1}^d \theta_i \alpha_i, \quad \rho_i = \sum_{j=1}^d \theta_i \beta_j^{(i)}, \quad (4.25)$$

where $\beta_j^{(i)}$ is the $j$th entry of vector $\beta_i$. Then we have

$$\lambda(t) = \sum_{i=1}^d \theta_i \alpha_i + \sum_{i=1}^d \sum_{j=1}^d \theta_i \beta_j^{(i)} X_i(t) = \sum_{i=1}^d \theta_i \left( \alpha_i + \beta_i^T X(t) \right) = \sum_{i=1}^d \theta_i D_{ii}(X(t)) \geq 0. \quad (4.26)$$

In practical applications, one often needs to make a trade-off. Ensuring non-negativity of $\lambda$ imposes severe restrictions on the parameters and can therefore result in a bad fit of the model. Of course, if one does not explicitly impose the non-negativity of $\lambda$, the probability of $\lambda$ being negative should be small. As we will explain in the next chapter, we will not strictly impose the non-negativity of the default intensity in our model, but, as we will see, in practical situations this will give no problems.

### 4.5 Chapter Summary

In this chapter, we introduced the class of affine processes. The major attraction of these processes is that many expressions that are encountered in pricing formulas can be reduced to a system of ordinary differential equations or can even be solved analytically if one assumes that the relevant stochastic processes are affine. This is also true for intensity-based models, in which the default intensity can be modeled as an affine process.

Many expressions within intensity-based models are similar to expressions encountered in short-rate interest rate models. Especially, if one thinks of the default intensity process as the short-rate process, the so-called zero-coupon bond price formula is encountered many times. We
4.5. Chapter Summary

have shown how an affine modeling structure gives rise to a solution of the bond price formula in terms of a system of ordinary differential equations. Typically, this system has to be solved numerically, but in some cases, it can be solved analytically.

The non-negativity constraint of the default intensity process imposes severe restrictions on the modeling possibilities and, therefore, there is often a trade-off between ensuring that this constraint is fulfilled and relaxing this constraint, which allows for more modeling freedom.
The Credit-Liquidity Model

Having built up all the ingredients in the previous chapters, we are able to introduce our ‘credit-liquidity’ model, which is an intensity-based model that also incorporates liquidity effects. We want to achieve two goals. Firstly, we want to measure the size of the credit and liquidity components in the CDS premia. Secondly, we want to obtain the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-default intensity process of the CDS reference entity (see section 3.4 for a discussion of this notion), since this, in combination with Rabobank’s model of the default event risk premium, is sufficient to compute the actual default probabilities (see formula (3.29)).

To achieve these goals, we will, following Bühler and Trapp (2008) [11], model the CDS bid price separately from the CDS ask price by introducing a bid-liquidity discount factor and an ask-liquidity discount. In this way, we explicitly take into account the information of the bid-ask spread, which is, in general, thought of as the main indicator of liquidity problems in a market. The inclusion of a liquidity discount factor within the intensity-based framework is also proposed by other authors (see, e.g., Duffie & Singleton (1997) [24], Dunbar (2007) [26], Chen, Fabozzi and Sverdlove (2010) [14]), and can be thought of as the “fractional carrying costs” of holding an instrument (Duffie & Singleton, 1997 [24]).

By calibrating the model-implied bid/ask premia to the observed bid/ask premia, we are able to decompose the CDS premium into a liquidity component and a credit component. Furthermore, the calibration of the model gives us the time series of the $\mathbb{Q}$-default intensity process, which can be used to obtain the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-default intensity process.

The data description, calibration procedure and the discussion of the results are the topics of the next chapters. In this chapter, we will start with a short review of the literature on CDS liquidity. After that, we will explain the model set-up and discuss the different assumptions we make. In the last section of this chapter, we will discuss how we will use the model to achieve our goals.
5.1 CDS Liquidity Literature

Since the liquidity literature is vast, we do not try to give a complete overview, and we will focus on some studies that explicitly investigate CDS liquidity. We refer to Amihud, Mendelson and Pedersen (2005) [4] for a thorough survey of both theoretical and empirical papers that study the impact of liquidity on the prices of assets such as stocks and bonds. For a survey on market microstructure, a branch of literature that investigates pricing mechanisms and the causes of illiquidity, we refer to Madhavan (2000) [46].

Given the clear indications of liquidity issues in the CDS market, it is surprising that not much research has been done on this topic. Especially the sovereign CDS market has not received a lot of attention in the academic literature. Of course, in total there are still more studies than we can discuss here, but at least we will be able to discuss most of the standard references.

To the best of our knowledge, Chen Cheng and Wu (2005) [13] are the first to study CDS liquidity. They use an intensity-based model to study the dynamic interaction between interest rate, credit and liquidity risk. They calibrate model-implied CDS premia to a large dataset of corporate CDS quotes. Since their dataset does not contain bid-ask information, they estimate liquidity by the frequency of price updates. They find that there is a large liquidity factor in the CDS quotes and that the buyers of CDSs receive heavy discounts as compensation for liquidity premia.

Tang and Yan (2008) [54] also use corporate CDS data to study the effects of liquidity on the CDS premium. They find that both search costs and adverse selection play an important role in affecting liquidity and, in turn, the CDS premia. They estimate that liquidity premia and liquidity risk premia together account for about 20% of the CDS spread.

Bongaerts, de Jong and Driessen (2011) [8] develop a Liquidity Capital Asset Pricing Model (LCAPM) that extends the model proposed by Acharya and Pedersen (2005) [1]. Their goal is to investigate the effects of liquidity on CDS prices. Using a large dataset of corporate CDS bid/ask quotes, they find, using a two-step regression, that credit and liquidity risk are significant determinants of the CDS premium and that the expected liquidity premium can be ascribed to the protection seller. Furthermore, they find that the expected liquidity premium exceeds the effect of liquidity risk.

Another intensity-based model is proposed by Bühler and Trapp (2008) [11]. They model the bid and ask premium separately by incorporating a bid-liquidity discount factor and an ask-liquidity discount factor, respectively. Furthermore, they incorporate the liquidity of the related bond market, since they argue that the bond liquidity has an impact on the recovery rate and, therefore, on the default payment. Opposed to the other papers, they, realistically, do not assume that liquidity and credit effects are independent. They use 5 year corporate CDS and corporate bond data of different maturities to calibrate their model. Their approach
5.2 Model Set-Up

allows for a natural decomposition of the CDS premium into a credit component and a liquidity component and they find that, on average, the liquidity component explains 4% of the CDS premium. Brigo, Predescu and Capponi (2010) [10] conducted a literature review on CDS liquidity modeling and they state that the study of Bühler and Trapp (2008) [11] “provides the most general and a more realistic reduced form framework by incorporating correlation between liquidity and credit, liquidity spillover effects between bonds and CDS contracts and asymmetric liquidity effects on the Bid and Ask CDS premium rates.”

All of the above studies focus on corporate CDSs, and to the best of our knowledge, not much work has been done on liquidity effects in the sovereign CDS market. Calice, Chen & Williams (2011) [12] study liquidity spillovers in sovereign bond and CDS markets and their main finding is that CDS liquidity can have a substantial impact on the bond spreads. The only study that we found that explicitly decomposes sovereign CDS premia into credit and liquidity components is the study by Badaoui, Cathcart and El-Jahel (2013) [5]. In their study, they extend the Bühler and Trapp model and apply it to sovereign CDS and bond data. They find that approximately 44% of the sovereign CDS premium can be attributed to liquidity risk and 56% to credit risk.

Although all of the above studies use different methodologies and data, they all conclude that CDS premia are not pure measures of credit risk, and that there is a significant liquidity factor in the CDS premia.

5.2 Model Set-Up

In this section, we will discuss our ‘credit-liquidity’ model. The structure of our model is similar to that of the model of Bühler and Trapp (2008) [11], but, while they investigate the interplay between the corporate bond and CDS market, we will focus on the sovereign CDS market. We choose to base our model on that of Bühler and Trapp, since this model allows for a specific decomposition of the CDS premium into a credit and a liquidity part, which is one of the things we also want to investigate. Furthermore, we think that their model is the most realistic, since it is the only model that explicitly incorporates correlation between the liquidity and credit components.

We will assume an arbitrage-free economy in which only credit default swaps and default-free zero coupon bonds are traded. On the technical side, this means that we assume that we are given the probability space $(\Omega, \mathcal{G}, Q)$, endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, such that $\mathcal{F}_t \subset \mathcal{G}_t$ for all $t \in \mathbb{R}_+$. The probability measure $Q$ denotes the risk-neutral probability measure, equivalent to the objective probability measure, which, by the first fundamental theorem of asset pricing, implies that the model is arbitrage-free. We assume that the probability space is sufficiently rich to be able to construct a default time $\tau : \Omega \rightarrow \mathbb{R}_+$ on it for an $\mathbb{F}$-given hazard
process using the canonical approach (see section 3.2 on the construction of a random time). We will, furthermore, assume that \( Q\{\tau = 0\} = 0 \) and \( Q\{\tau > t\} > 0 \) for all \( t \in \mathbb{R}_+ \) and we denote by \( H \) the process given by \( H_t = 1_{\{\tau \leq t\}} \) for \( t \in \mathbb{R}_+ \). This process is right-continuous and we denote the filtration generated by the process \( H \) as \( \mathbb{H} \), with \( \mathbb{H}_t = \sigma(\{\tau \leq u\} : u \leq t) \). We will set \( \mathbb{G} = \mathbb{F} \vee \mathbb{H} \), i.e., \( \mathbb{G}_t = \mathbb{F}_t \vee \mathbb{H}_t \) for all \( t \in \mathbb{R}_+ \). Note that, since the default time is constructed by the canonical approach, Condition 2 of chapter 3 is satisfied.

The liquidity in the credit default swap market can change over time, but we will assume that there are no liquidity effects in the default-free zero coupon bond market and, therefore, the prices of these bonds are not influenced by liquidity. The consequence of this assumption is that the effects of liquidity on the CDS premia are measured relative to the default-free zero coupon bonds and, therefore, this assumption allows us to circumvent the problem of specifying a perfectly liquid instrument in comparison to which all CDS are traded with a discount\(^1\) (Bühler & Trapp, 2008 [11]).

Since the bid and ask premia of a CDS contract differ, and since the underlying credit risk should be the same, we argue that the difference between the two is caused by a different liquidity discount factor. By modeling the bid and the ask premia separately by means of different liquidity discount factors, we use the information of the bid-ask spread on the liquidity level in the market. In general, the use of bid-ask spreads in studies on liquidity effects is very common (see, for example, Amihud & Mendelson (1986) [3]).

We will thus encounter the following 'discount' factors in our model:

\[
\begin{align*}
\tilde{D}(t,T) &= e^{-\int_t^T r(s)ds} \\
\tilde{P}(t,T) &= e^{-\int_t^T \lambda(s)ds} \\
\tilde{L}^{bid/ask}(t,T) &= e^{-\int_t^T \gamma^{bid/ask}(s)ds}.
\end{align*}
\]

Here \( r(t) \) denotes the instantaneous default-free interest rate process, \( \lambda(t) \) the default intensity process and \( \gamma^{bid/ask}(t) \) the liquidity intensity rate for the bid premium (\( \gamma^{bid} \)) and ask premium (\( \gamma^{ask} \)), which we assume to be different\(^2\). We assume that \( r(t), \lambda(t) \) and \( \gamma^{bid/ask}(t) \) follow stochastic \( \mathbb{F} \)-adapted processes and, therefore, the above discount factors are stochastic as well.

\(^1\)The word “discount” is a bit unfortunate, since it is ambiguous whether the CDS premium is higher because of liquidity effects or lower. If sellers dominate the market, they could demand a risk premium for bearing liquidity risk and, therefore, the CDS premium is set at a higher level than it would be in a perfectly liquid market. If, on the other hand, buyers dominate the market, they could ask for a discount for trading in an illiquid market and, therefore, the CDS premium is set at a lower level than in the perfectly liquid case.

\(^2\)Note that the name liquidity intensity may be misleading in the sense that it is not the intensity process of an underlying jump process (like the default intensity). The liquidity discount factors have, therefore, more in common with the interest rate discount factor than with the default process discount factor.
5.2. Model Set-Up

The precise structure of the stochastic components will be given later, but first we will deduce the CDS bid and ask pricing formulas.

5.2.1 Model-Implied Bid and Ask Premia

To model the default payment (floating) leg of the CDS contract, we have to choose how to model the recovery payment (see section 3.3). Since CDS contracts actually specify a recovery of par payoff, this recovery model is the obvious choice. In principle, the recovery rate can be time-varying and stochastic, but to simplify the expressions, we will assume that the recovery rate is an exogenously given constant. This assumption is not so strict as it looks, since the quotes in the market are also based on models that use a fixed recovery rate\(^3\). Apart from the CDS documentation that uses fixed recovery rates, Houweling and Vorst (2005) \cite{houweling_vorst_2005} show that misspecification of the recovery rate has a negligible effect on the pricing of CDS and, therefore, taking a constant recovery rate is justified.

We denote the set of times \(T_1, T_2, \ldots, T_n\) as the dates on which the protection buyer pays the premium to the protection seller. We will make a simplifying assumption that, if a default occurs on a non-reference date, the protection payment is paid at the first time \(T_i\) following the default\(^4\). This assumption allows us to ignore accrual interest payments of the protection buyer to the protection seller and, furthermore, it allows us to focus only on the reference dates, which makes the pricing formulas computationally less expensive\(^5\).

Let \(F\) denote the notional value of the CDS contract, \(s_{\text{bid/ask}}\) the annualized bid and ask premia in percentages (normally they are quoted in basis points), \(t = T_0, T_n\) the maturity of the contract and \(\delta_i = T_i - T_{i-1}\) the year fraction between \(T_i\) and \(T_{i-1}\). The discounted value of the floating leg of the CDS is then given by

\[
\Pi_{\text{float}}(t, [T_1, \ldots, T_n], F, R) = (1 - R) \cdot F \cdot \sum_{i=1}^{n} \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \bar{D}(t, T_i). \tag{5.1}
\]

Expression (5.1) can be interpreted as follows: if a default occurs between \(T_{i-1}\) and \(T_i\) (this happens with probability \(\bar{P}(t, T_{i-1}) - \bar{P}(t, T_i)\)), the protection seller pays an amount of 1 minus

\(^3\)Since 2009, CDS contracts have standardized premia. The quotes in the market are, however, still those premium values that give the contracts initial values of zero. The conversion of the standardized premia to the classic premia is done by the ISDA CDS Standard Model, which specifies a recovery rate of 40\% for developed countries and of 25\% for developing countries. Since everyone in the market is using the ISDA CDS Standard Model, credit default swaps cannot be used to determine the loss given default (LGD) measure of credit risk.

\(^4\)In practice, the default payment is also not made immediately after the default event, since the level of the default payment has to be specified by procedures that might take some time (e.g. bankruptcy court decisions).

\(^5\)In a continuous-time framework, one has to integrate over all possible default times and, in general, this integral has to be solved numerically. In a discrete-time framework, however, we can just sum over the reference dates (Duffie, 1998 [19]).
the recovery rate times the notional value of the contract at time $T_i$ to the protection buyer. This payment has to be discounted by the interest rate ($\bar{D}(t, T_i)$).

Since the CDS premium is agreed upon by both the protection buyer and seller, we will assume that all liquidity effects can be incorporated into the fixed leg. We will, therefore, model the fixed leg of the CDS separately for the bid and the ask side. The discounted value of the fixed CDS leg for the protection seller at time $t$ is given by

$$\Pi_{fix}^{ask}(t, [T_1, \ldots, T_n], s^{ask}, F, R) = s^{ask} \cdot F \cdot \sum_{i=1}^{n} \bar{P}(t, T_i) \bar{D}(t, T_i) \bar{L}^{ask}(t, T_i) \delta_i. \quad (5.2)$$

The discounted value of the fixed CDS leg for the protection buyer at time $t$ is given by

$$\Pi_{fix}^{bid}(t, [T_1, \ldots, T_n], s^{bid}, F, R) = (1 - R) \cdot F \cdot \sum_{i=1}^{n} \bar{P}(t, T_i) \bar{D}(t, T_i) \bar{L}^{bid}(t, T_i) \delta_i. \quad (5.3)$$

Expression (5.2) can be interpreted as follows: if there is no default before time $T_i$ (this happens with probability $\bar{P}(t, T_i)$), the protection seller wants to receive a payment of $\delta_i \cdot s^{ask} \cdot F$ at time $T_i$, and to value this payment today, it has to be discounted by the interest and the liquidity discount factors ($\bar{D}(t, T_i)$ and $\bar{L}^{ask}(t, T_i)$). A similar interpretation follows for (5.3), but here we reason from the buyer’s perspective, who wants to make a payment of $\delta_i \cdot s^{bid} \cdot F$ at $T_i$. All these expressions coincide with those in Brigo and Mercurio (2006) [9], except that we have added the liquidity discount factors in the fixed leg.

Let us now consider the total value of the CDS contract. We denote by $\Pi_{CDS}^{ask}$ the discounted CDS value as viewed by the protection seller and by $\Pi_{CDS}^{bid}$ the discounted CDS value as viewed by the protection buyer. We get

$$\Pi_{CDS}^{ask}(t, [T_1, \ldots, T_n], s^{ask}, F, R) = s^{ask} \cdot F \cdot \sum_{i=1}^{n} \bar{P}(t, T_i) \bar{D}(t, T_i) \bar{L}^{ask}(t, T_i) \delta_i$$

$$- (1 - R) \cdot F \cdot \sum_{i=1}^{n} \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \bar{D}(t, T_i),$$

$$\Pi_{CDS}^{bid}(t, [T_1, \ldots, T_n], s^{bid}, F, R) = (1 - R) \cdot F \cdot \sum_{i=1}^{n} \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \bar{D}(t, T_i)$$

$$- s^{bid} \cdot F \cdot \sum_{i=1}^{n} \bar{P}(t, T_i) \bar{D}(t, T_i) \bar{L}^{bid}(t, T_i) \delta_i.$$
5.2. Model Set-Up

Until now, we did not take into account that the interest rate, the default intensity and the liquidity intensities are all stochastic. By risk-neutral pricing, we can, therefore, compute the value of a CDS contract by taking conditional expectations:

\[
CDS_{\text{ask}}(t, [T_1, \ldots, T_n], s_{\text{ask}}, F, R) = \mathbb{E}^Q \left[ \Pi_{\text{CDS}}(t, [T_1, \ldots, T_n], s_{\text{ask}}, F, R) \bigg| \mathcal{G}_t \right],
\]

\[
CDS_{\text{bid}}(t, [T_1, \ldots, T_n], s_{\text{bid}}, F, R) = \mathbb{E}^Q \left[ \Pi_{\text{CDS}}(t, [T_1, \ldots, T_n], s_{\text{bid}}, F, R) \bigg| \mathcal{G}_t \right].
\]

By Lemma 3.4, we can switch to the filtration \(\mathcal{F}_t\) and we get:

\[
CDS_{\text{ask}}(t, [T_1, \ldots, T_n], s_{\text{ask}}, F, R) = 1_{\{\tau > t\}} \mathbb{E}^Q \left[ \Pi_{\text{CDS}}(t, [T_1, \ldots, T_n], s_{\text{ask}}, F, R) \bigg| \mathcal{F}_t \right],
\]

\[
CDS_{\text{bid}}(t, [T_1, \ldots, T_n], s_{\text{bid}}, F, R) = 1_{\{\tau > t\}} \mathbb{E}^Q \left[ \Pi_{\text{CDS}}(t, [T_1, \ldots, T_n], s_{\text{bid}}, F, R) \bigg| \mathcal{F}_t \right].
\]

The bid and the ask premia \(s_{\text{ask}}\) and \(s_{\text{bid}}\) are by definition those premia that make \(CDS_{\text{ask}}\) and \(CDS_{\text{bid}}\) equal to zero and, therefore, we have the following formulas for \(s_{\text{ask}}\) and \(s_{\text{bid}}\):

\[
s_{\text{ask}} = \frac{(1 - R) \cdot F \cdot \sum_{i=1}^{n} \mathbb{E}^Q \left[ \left( \tilde{P}(t, T_{i-1}) - \tilde{P}(t, T_i) \right) \tilde{D}(t, T_i) \bigg| \mathcal{F}_t \right]}{F \cdot \sum_{i=1}^{n} \delta_i \mathbb{E}^Q \left[ \tilde{D}(t, T_i) \tilde{P}(t, T_i) \tilde{L}_{\text{ask}}(t, T_i) \bigg| \mathcal{F}_t \right]},
\]

\[
s_{\text{bid}} = \frac{(1 - R) \cdot F \cdot \sum_{i=1}^{n} \mathbb{E}^Q \left[ \left( \tilde{P}(t, T_{i-1}) - \tilde{P}(t, T_i) \right) \tilde{D}(t, T_i) \bigg| \mathcal{F}_t \right]}{F \cdot \sum_{i=1}^{n} \delta_i \mathbb{E}^Q \left[ \tilde{D}(t, T_i) \tilde{P}(t, T_i) \tilde{L}_{\text{bid}}(t, T_i) \bigg| \mathcal{F}_t \right]}.\]

Expressions (5.4) and (5.5) differ only by the liquidity discount factors \(\tilde{L}_{\text{bid}}\) and \(\tilde{L}_{\text{ask}}\). To rule out arbitrage opportunities, we need that \(s_{\text{ask}} \geq s_{\text{bid}}\), which implies that \(\tilde{L}_{\text{ask}} \leq \tilde{L}_{\text{bid}}\) or, stated differently, \(\gamma_{\text{ask}} \geq \gamma_{\text{bid}}\). This inequality is not strictly imposed by our model, but, as we will see later, it is never violated in our empirical study.

Bühler and Trapp (2008) [11] arrive at similar bid and ask spread formulas, but they do not give a clear economic interpretation of the liquidity discount factors. We will, therefore, elaborate on this a bit. First we note that, if the CDS market would be perfectly liquid, we would have no liquidity discount factors and, therefore, putting \(\gamma_{\text{bid}} = \gamma_{\text{ask}} = 0\) in (5.4) and (5.5) gives
the CDS bid and ask premia (which are equal) in that case. In the perfectly liquid market, the CDS premium, therefore, reflects the premium based on credit risk only. We already noted that, to prevent arbitrage opportunities, we need $\bar{L}^{\text{ask}} \leq \bar{L}^{\text{bid}}$ and in the perfectly liquid market we have $\bar{L}^{\text{ask}} = \bar{L}^{\text{bid}} = 1$. There are now three cases to consider:

1. $\bar{L}^{\text{ask}} \leq 1 \leq \bar{L}^{\text{bid}}$,
2. $\bar{L}^{\text{ask}} \leq \bar{L}^{\text{bid}} \leq 1$,
3. $1 \leq \bar{L}^{\text{ask}} \leq \bar{L}^{\text{bid}}$.

In the first case, we have $\gamma^{\text{ask}} \geq 0$ and $\gamma^{\text{bid}} \leq 0$. A decrease in liquidity on the ask side of the market, induces an increase in $\gamma^{\text{ask}}$ and a decrease in liquidity on the bid side induces a decrease $\gamma^{\text{bid}}$ ($\gamma^{\text{bid}}$ becomes more negative). So the closer $\gamma^{\text{ask}}$ and $\gamma^{\text{bid}}$ are to zero, the higher the liquidity in the respective markets is. The consequence of this is that a decrease in liquidity in one (or both) sides of the market, would, ceteris paribus, induce a wider bid-ask spread. In principle, this is an observation that one would expect. We have that the CDS premium that would be agreed upon in the perfectly liquid market lies somewhere between the bid and ask premia of the illiquid market. The height of the agreed CDS premium in the illiquid market may thus be higher or lower than the CDS premium in the perfectly liquid market, depending on whether the buyer or the sellers are dominant. An economic interpretation of the behavior of the liquidity intensities would be that a decrease in liquidity on the ask side makes protection sellers demand a higher compensation for supplying the remaining liquidity on the sell-side of the market. If they, in a later stadium, would want to hedge their positions by buying CDSs, the lower sell-side liquidity would make that harder and, therefore, they demand a risk premium for selling in the first place. On the other hand, a decrease in liquidity on the buy-side could give the remaining buyers a better bargaining position and allow them to negotiate a lower premium.

In the second case, both $\gamma^{\text{ask}}$ and $\gamma^{\text{bid}}$ are positive and a decrease in liquidity induces an increase in the intensities. In this case, the CDS premium in the perfectly liquid market is even lower than the bid price in the illiquid market, meaning that the CDS premium that is agreed upon by protection buyers and sellers in the illiquid market is always higher than the CDS premium in the perfectly liquid market (since the agreed premium lies somewhere between the bid and the ask premium). The liquidity risk premium is thus claimed by the protection sellers in this case, since they get a higher CDS premium than they would have had in a perfectly liquid market. That the liquidity premiums go to the protection sellers is in line with Bongaerts, De Jong and Driessen (2011) [8]. An economic interpretation of the behavior of the ask liquidity intensity is again that a lower liquidity on the sell-side of the market makes the remaining sellers ask a higher compensation for liquidity risk. The effect of a decrease in liquidity on the buy-side now has a very different interpretation than in the first case. A decrease in the liquidity on the buy-side increases the bid premium as the buyers that are still in the market are those buyers that really want to buy protection and, therefore, are willing to bid higher.
5.2. Model Set-Up

In the third case, both $\gamma_{ask}$ and $\gamma_{bid}$ are negative and a decrease in liquidity induces a decrease in the intensities (they become more negative). In a similar argument as above, we can now state that the CDS premium in the illiquid market will be lower than in the perfectly liquid market, meaning that the protection buyers claim the liquidity risk premium in terms of a discount relatively to the pure credit risk CDS premium of the perfectly liquid market. In this case, a lower liquidity on the sell-side of the market induces lower ask premia, as the sellers that remain in this market are more eager to sell and, therefore, ask lower premia. A decrease in liquidity on the buy-side of the market means that the remaining buyers have more bargaining power and are, therefore, able to set lower bid premia.

It could also happen that in some periods buyers dominate the market and in other periods sellers dominate the market and that, therefore, a regime-switch could happen. In general, we can conclude that the further the values of $\gamma_{ask/bid}$ are away from zero, the lower the liquidity is in the corresponding market. The signs of $\gamma_{bid/ask}$ depends on the (relative) market power the buyers and sellers have.

A last note we want to make regarding expressions (5.4) and (5.5) is that, ceteris paribus, an increase in credit risk widens the bid-ask spread, as it has more effect on the ask spread than on the bid spread. This is in line with empirical observations that a larger bid-ask spread is associated with higher levels of the CDS premium (IMF, 2013 [33]).

5.2.2 Specification of the Stochastic Components

We will assume that the risk-free rate is independent of the default and liquidity intensities and is given by the USD swap rates. The assumption of independence between the risk-free rate and the default intensity is standard in the academic literature and amongst practitioners. It allows one to separate the calibration of the interest rate part to interest rate data and the default intensity part to the credit market data. Brigo and Mercurio (2006) [9] show with a Monte Carlo simulation that this assumption is not harmful for pricing purposes. The independence between interest rates and the liquidity intensities maintains the advantage of being able to analyse the interest rate part separately. The bid and ask premia formulas of (5.4) and (5.5) now become

\[ s_{ask} = \frac{(1 - R) \cdot \sum_{i=1}^{n} \mathbb{E}^Q \left[ \bar{D}(t, T_i) \right] \mathbb{F}_t \mathbb{E}^Q \left[ \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \right] \mathbb{F}_t}{\sum_{i=1}^{n} \delta_i \mathbb{E}^Q \left[ \bar{D}(t, T_i) \right] \mathbb{F}_t \mathbb{E}^Q \left[ \bar{P}(t, T_i) \bar{L}_{ask}(t, T_i) \right] \mathbb{F}_t} \]  

(5.6)

and
\[ s_{\text{bid}} = \frac{(1 - R) \cdot \sum_{i=1}^{n} \mathbb{E}^Q \left[ \tilde{D}(t, T_i) \right] \mathbb{F}_t \left[ \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \right] \mathbb{F}_t}{\sum_{i=1}^{n} \delta_i \mathbb{E}^Q \left[ \tilde{D}(t, T_i) \right] \mathbb{F}_t \left[ \bar{P}(t, T_i) \tilde{L}_{\text{bid}}(t, T_i) \right] \mathbb{F}_t} \], \quad (5.7)

where the interest rate discount factors \( \mathbb{E}^Q \left[ \tilde{D}(t, T_i) \right] \mathbb{F}_t \) will be computed from discount curves constructed from the USD swap rate curve

We do, however, assume that the liquidity and default intensities are correlated, and we suggest the following dependence structure:

\[
\begin{pmatrix}
\frac{d \lambda(t)}{dt} \\
\frac{d \gamma_{\text{bid}}(t)}{dt} \\
\frac{d \gamma_{\text{ask}}(t)}{dt}
\end{pmatrix} =
\begin{pmatrix}
1 & g_{\text{bid}} & g_{\text{ask}} \\
f_{\text{bid}} & 1 & \omega_{\text{ask,bid}} \\
f_{\text{ask}} & \omega_{\text{bid,ask}} & 1
\end{pmatrix}
\begin{pmatrix}
\frac{dx(t)}{dt} \\
\frac{dy_{\text{bid}}(t)}{dt} \\
\frac{dy_{\text{ask}}(t)}{dt}
\end{pmatrix}. \quad (5.8)
\]

The factors \( x(t), y_{\text{bid}}(t) \) and \( y_{\text{ask}}(t) \) are assumed to be independent and we can think of these factors as the \textit{pure} default and liquidity intensities. The intensities on the left hand side of (5.8) then represent the (full) correlated intensities. We will denote the components of the factor matrix as the \textit{correlation factors}, since they induce a correlation structure in the model.

The above dependence structure thus implies that

\[
\begin{align*}
\lambda(t) &= x(t) + g_{\text{bid}} y_{\text{bid}}(t) + g_{\text{ask}} y_{\text{ask}}(t), \\
\gamma_{\text{bid}}(t) &= f_{\text{bid}} x(t) + y_{\text{bid}}(t) + \omega_{\text{ask,bid}} y_{\text{ask}}(t), \\
\gamma_{\text{ask}}(t) &= f_{\text{ask}} x(t) + \omega_{\text{bid,ask}} y_{\text{bid}}(t) + y_{\text{ask}}(t).
\end{align*}
\]

The pure credit risk factor \( x \) may thus influence the liquidity intensities \( \gamma^l \) through the factors \( f_l, l \in \{\text{bid, ask}\} \). Depending on the sign of \( \gamma^l \), a positive value of \( f_l \), may either imply that a higher (pure) default intensity increases liquidity (if \( \gamma^l \) is negative) or decreases liquidity (if \( \gamma^l \) is positive). The factors \( g_{\text{ask}} \) and \( g_{\text{bid}} \) indicate the effects the liquidity intensities have on the default intensity. We would expect that these factors are small, as the liquidity of a CDS market should not influence the default probability very much. The factors \( \omega_{\text{bid,ask}} \) and \( \omega_{\text{ask,bid}} \) imply a direct link between the liquidity of the sell- and buy-sides. It allows us to determine if liquidity in one side of the market influences the liquidity on the other side of the market directly. Of course the liquidities also indirectly influence each other by their dependence on \( x(t) \). We will assume that the correlation factors are constant over time.

\(^{6}\)The construction of the discount curve is done by bootstrapping the USD swap rates and, therefore, we do not need to model the interest rates explicitly.
5.2. Model Set-Up

Instead of modeling $\lambda$, $\gamma^{bid}$ and $\gamma^{ask}$, we will model the pure intensities $x$, $y^{bid}$ and $y^{ask}$. We will assume that the pure default intensity $x$ follows a CIR process:

$$dx(t) = (\alpha - \beta x(t))dt + \sigma \sqrt{x(t)}dW_x(t).$$

For the liquidity intensities, we will use two options. In the first option, we will model the pure liquidity intensities $y^{bid}$ and $y^{ask}$ as an Arithmetic Brownian motion without drift (also referred to as Gaussian processes):

$$dy^l(t) = \sigma^l dW_{y^l}(t) \text{ for } l \in \{bid, ask\}.$$

In the second model option, we will model the pure liquidity intensities as Ornstein-Uhlenbeck (OU) processes with mean-reversion levels of zero and the same mean-reversion speed parameter $\eta$:

$$dy^l(t) = \eta y^l(t)dt + \sigma^l dW_{y^l}(t) \text{ for } l \in \{bid, ask\}.$$

In both cases, we have that $W_x, W_{y^{bid}}$ and $W_{y^{ask}}$ are independent $\mathcal{Q}$-Brownian motions and we will, from now on, assume that the filtration $\mathcal{F}$ is generated by the 3-dimensional Brownian motion $W = (W_x, W_{y^{bid}}, W_{y^{ask}})$.

We choose to model the intensities in this way for multiple reasons. First of all, modeling the pure default intensity by a CIR process assures non-negativity of $x$. Of course, the actual default intensity is $\lambda$, which may become negative in the above set-up, as the factors $y^{bid}$ and $y^{ask}$ can take on both negative and positive values due to their Gaussian nature. We conjecture, however, that the correlation factors $g^{bid}$ and $g^{ask}$ will be very small, since the liquidity of a credit default swap should not have a large influence on the default probability of a sovereign and that, therefore, the default intensity $\lambda$ will be almost entirely driven by $x$. The non-negativity of $x$ then induces the non-negativity of $\lambda$.

Modeling the pure liquidity intensities as driftless Brownian motions is in line with Longstaff, Mithal and Neis (2005) [45] and Badaoui, Cathcart and El-Jahel (2013) [5]. The big advantage of modeling the pure liquidity intensities in this way, is that we only have two liquidity-related parameters, namely $\sigma^{bid}$ and $\sigma^{ask}$. As we will see in chapter 6, the running time of our calibration procedure increases exponentially in the number of process parameters, thereby making it impossible to run useful calibrations for even a modest number of process parameters. Of course, economically speaking, the use of a driftless Arithmetic Brownian motion is less appealing, as it suggests that the variance of the liquidity processes increases over time. Because of the shortcomings of the driftless Arithmetic Brownian motion, we also model the liquidity intensities as Ornstein-Uhlenbeck processes to see if these processes give better results.
We assume that both $y^{bid}$ and $y^{ask}$ have mean-reversion levels of zero and the same mean-reversion speeds. This is again done to keep the total number of process parameters low\(^7\). We conjecture that the OU process is a bit more stable than the driftless Arithmetic Brownian motion and, therefore, we hope to get more stable results.

In both cases, $y^{bid}$ and $y^{ask}$ can be either positive or negative. This is not a problem per se, since, as we discussed above, the liquidity intensities $\gamma^{bid}$ and $\gamma^{ask}$ can be either positive or negative (depending on the relative market power of the buyers and sellers), but, of course, we do not want $\lambda$ to become negative.

Lastly, all the processes described above fall into the the class of affine processes. In combination with the above defined dependence structure, this allows us to derive analytical expressions for (5.6) and (5.7). To see this, note that there are still two types of expectations that need to be computed in these formulas, namely

\[
\mathbb{E}
\left[
\tilde{P}(t,T_i)\bigg| \mathcal{F}_t\right] = \mathbb{E}
\left[
\mathbb{E}
\left[
\left. e^{-\int_t^{T_i} \lambda(s)ds} \right| \mathcal{F}_t\right] \right],
\]

\[
\mathbb{E}
\left[
\tilde{P}(t,T_i)\tilde{L}^l(t,T_i)\bigg| \mathcal{F}_t\right] = \mathbb{E}
\left[
\mathbb{E}
\left[
\left. e^{-\int_t^{T_i} \lambda(s)ds} e^{-\int_t^{T_i} \gamma^l(s)ds} \right| \mathcal{F}_t\right] \right],
\]

for $l \in \{bid,ask\}$. We will show that we can write these expectations as the product of independent expectations, which can be solved analytically by their affine structure. Let us first prove the following result:

**Lemma 5.1.** Consider two independent Brownian motions $W_t$ and $V_t$ with natural filtrations $\mathcal{F}_t^W$ and $\mathcal{F}_t^V$ respectively. For $s \leq t$ and functions $f$ and $g$, we have

\[
\mathbb{E}
\left[
\left. f(W_t)g(V_t) \right| \mathcal{F}_s^W \vee \mathcal{F}_s^V\right] = \mathbb{E}
\left[
\left. f(W_t) \right| \mathcal{F}_s^W \vee \mathcal{F}_s^V\right] \mathbb{E}
\left[
\left. g(V_t) \right| \mathcal{F}_s^W \vee \mathcal{F}_s^V\right]. \tag{5.9}
\]

\(^7\)As the results of our calibration will show, the bid and ask liquidity intensities move together strongly and, therefore, the assumption of the same mean-reversion speed is not unreasonable. The assumption of a zero mean-reversion level is more debatable, but it implies that, in the long-run, the market moves towards perfect liquidity.
5.3. How to Use the Model?

Proof.

\[ \mathbb{E} \left[ f(W_t)g(V_t) \mid \mathcal{F}_t^W \vee \mathcal{F}_s^V \right] = \mathbb{E} \left[ \mathbb{E} \left[ f(W_t)g(V_t) \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \]

(tower property)

\[ = \mathbb{E} \left[ f(W_t) \mathbb{E} \left[ g(V_t) \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \]

(measurability)

\[ = \mathbb{E} \left[ f(W_t) \mathbb{E} \left[ g(V_t) \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \]

(independency)

\[ = \mathbb{E} \left[ f(W_t) \mathbb{E} \left[ g(V_t) \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \]

(taking-out-what-is-known)

\[ = \mathbb{E} \left[ f(W_t) \mathbb{E} \left[ g(V_t) \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \mid \mathcal{F}_s^W \vee \mathcal{F}_s^V \right] \]

(independency).

Lemma 5.1 can easily be extended to more than two independent Brownian motions and it allows us to write

\[ \mathbb{E} \left[ e^{-\int_t^T \lambda(s) ds} \mid \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^T \left( x(s) + g_{\text{bid}}y_{\text{bid}}(s) + g_{\text{ask}}y_{\text{ask}}(s) \right) ds} \mid \mathcal{F}_t \right] \]

\[ = \mathbb{E} \left[ e^{-\int_t^T x(s) ds} \mid \mathcal{F}_t \right] \mathbb{E} \left[ e^{-\int_t^T g_{\text{bid}}y_{\text{bid}}(s) ds} \mid \mathcal{F}_t \right] \mathbb{E} \left[ e^{-\int_t^T g_{\text{ask}}y_{\text{ask}}(s) ds} \mid \mathcal{F}_t \right] \]

and

\[ \mathbb{E} \left[ e^{-\int_t^T \lambda(s) ds} e^{-\int_t^T \gamma'(s) ds} \mid \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^T (1+f_{\text{bid}}) x(s) ds} \mid \mathcal{F}_t \right] \mathbb{E} \left[ e^{-\int_t^T (1+g_{\text{bid}}) y'(s) ds} \mid \mathcal{F}_t \right] \times \mathbb{E} \left[ e^{-\int_t^T (g_k,\omega_{k,l}) y_{k,l}(s) ds} \mid \mathcal{F}_t \right], \]

for \( l, k \in \{\text{bid, ask}\} \). Every expectation in the above two expressions is of the form (4.20), and can be solved analytically. The computations are done in Appendices D, E, F and the results are summarized in Appendix C.

5.3 How to Use the Model?

By calibrating formulas (5.6) and (5.7) to daily observed bid and ask premia, we find the parameter values of the process dynamics, and we obtain discrete time series of \( x(t), y_{\text{bid}}(t) \) and \( y_{\text{ask}}(t) \) (i.e., we find the daily values of \( (x(t), y_{\text{bid}}(t), y_{\text{ask}}(t)) \) for days \( t = 0, 1, \ldots T \), with
Chapter 5. The Credit-Liquidity Model

T the last day of the sample). The calibration procedure will be explained in the next chapter and the results will be given in chapter 7.

In this section, we will explain how we will use the model to reach the two goals we have set, namely to decompose the CDS spread in credit and liquidity components and to find the \( \mathbb{P} \)-dynamics of the \( \mathbb{Q} \)-default intensity process. We will thus work under the assumption that we successfully calibrated our model.

5.3.1 Decomposition of the CDS Premium

A standard measure of liquidity effects in the literature on market microstructure is to consider the size of the bid-ask spread relative to the mid premium. In our model set-up, however, looking at the (absolute or relative) size of the bid-ask spread is not a suitable measure for liquidity effects. To see this, let us look at formulas (5.6) and (5.7). We see that, ceteris paribus, an increase in credit risk widens the bid-ask spread. Therefore, the bid-ask spread is not a pure measure of liquidity risk in our model and we need to find another way to distillate the credit and liquidity effects. Luckily, our model allows for a natural decomposition of the CDS (mid) premium into a pure credit risk premium, a pure liquidity risk premium and a correlation risk premium.

The pure credit risk premium component will be given by considering a perfectly liquid market. In a perfectly liquid CDS market, we do not need liquidity discount factors and, therefore, the CDS premium is given by taking the calibrated default intensity parameters and intensities and by setting the liquidity intensities to zero (which implies that \( L^{\text{ask}} = L^{\text{bid}} = 1 \) and that the correlation factors don’t play a role in formulas (5.6) and (5.7)). In this case, the model-implied bid and ask premia coincide as there are no liquidity costs. The corresponding premium will be denoted by \( s^{\text{cred}} \) and we will refer to it as the pure CDS credit premium.

If we now also use the calibrated liquidity parameters and intensities in formulas (5.6) and (5.7), but still ignore the effect of the correlation factors (by setting them equal to zero), we have that the model-implied bid and the ask premia will, in general, have different values. We will denote these values by \( \tilde{s}^{\text{bid}} \) and \( \tilde{s}^{\text{ask}} \), respectively. A natural measure of the liquidity premium is to use the corresponding mid premium and subtract from this the pure CDS credit premium. We thus get \( s^{\text{liq}} = \tilde{s}^{\text{mid}} - s^{\text{cred}} \), where \( \tilde{s}^{\text{mid}} = \frac{\tilde{s}^{\text{ask}} + \tilde{s}^{\text{bid}}}{2} \).

In a next step, we also include the correlation factors and we find the full model-implied bid and ask premia, which we will denote by \( \tilde{s}^{\text{bid}} \) and \( \tilde{s}^{\text{ask}} \). We will define the correlation premium as \( s^{\text{corr}} = \tilde{s}^{\text{mid}} - s^{\text{liq}} - s^{\text{cred}} \), where \( \tilde{s}^{\text{mid}} = \frac{\tilde{s}^{\text{ask}} + \tilde{s}^{\text{bid}}}{2} \).

All-in-all, we thus decompose the CDS mid premium into three components: a pure credit risk component, a pure liquidity risk component, and a correlation risk component. Note that by decomposing the mid premium, we implicitly assume that the actual agreed CDS premium is
5.3. How to Use the Model?

The mid premium. This assumption is standard in the literature, but need not be true in reality. It depends on the bargaining between the buyers and the sellers what CDS premium is agreed on. We know, however, that it must lie somewhere between the bid and the ask premium and, therefore, the mid premium seems like a logical choice.

By taking the mid premium we do not, however, assume that the buyers and sellers have equal market power, since it depends on the signs of the liquidity intensities which side of the market gets compensated for the liquidity risk. If, for example, both the bid and the ask liquidity intensities are positive, both the bid and the ask premia will be above the pure credit risk premium and, therefore, also the mid premium will be above the pure credit risk premium. This means that sellers receive a higher premium payment than in a perfectly liquid market and, therefore, they have more market power than the buyers. In the above decomposition, we can thus also have negative liquidity and correlation premia. If both the liquidity intensities are negative, both the bid and the ask premia lie below the pure credit risk premium and, therefore, \( s^{liq} \) is negative, meaning that the buyer gets compensated for the liquidity risk by means of a discount.

The above decomposition works fine in the case that we model the liquidity intensities as driftless Arithmetic Brownian motions, but when we model the liquidity intensities as Ornstein-Uhlenbeck processes, we cannot take all correlation factors equal to zero as there will be a division through zero somewhere in the formulas (see the bond price formulas in Appendix C). In this model variant, we can, therefore, only decompose the CDS mid premium into a pure credit premium (which is the same as described above) and another premium, which consists of the pure liquidity and correlation premia combined. This part of the premium is defined as the full mid premium minus the pure credit premium.

5.3.2 Finding the \( \mathbb{P} \)-Dynamics of the \( \mathbb{Q} \)-Default Intensity Process

The second goal of our research is to find the \( \mathbb{P} \)-dynamics of the \( \mathbb{Q} \)-default intensity process. This notion was discussed in section 3.4 and originated from the observation that investors do not only require a compensation for changes in the default environment, but also for the default event itself (Jarrow, Lando & Yu, 2005 [38]). Because of the risk premium on the default event itself, the \( \mathbb{Q} \)-default intensity is different from the \( \mathbb{P} \)-default intensity. Rabobank modeled the difference between the \( \mathbb{Q} \)- and \( \mathbb{P} \)-default intensities by means of a constant (country-specific) factor with which the \( \mathbb{Q} \)-default intensity has to be multiplied in order to obtain the \( \mathbb{P} \)-default intensity. Letting this country-specific constant be denoted by \( K \), we get for the real-world survival probabilities the following expression:

\[
\mathbb{P} \left\{ \tau > T \mid \mathcal{F}_t \right\} = \mathbb{E}^\mathbb{P} \left[ e^{-\int_t^T \lambda^\mathbb{P}(s)ds} \bigg| \mathcal{F}_t \right] = \mathbb{E}^\mathbb{P} \left[ e^{-\int_t^T K\lambda^\mathbb{Q}(s)ds} \bigg| \mathcal{F}_t \right].
\] (5.10)
We thus see that the expectation is taken with respect to the real-world probability measure \( \mathbb{P} \), but the default intensity process on the right-hand side is the \( \mathbb{Q} \)-default intensity process.

Let us recall that the \( \mathbb{Q} \)-default intensity, which we from now on will denote by \( \lambda^Q \), in our model is given by

\[
\lambda^Q(t) = x(t) + g_{bid}y^{bid}(t) + g_{ask}y^{ask}(t).
\]

If we assume for a moment that the liquidity intensities are modeled Ornstein-Uhlenbeck processes with mean-reversion level of zero and mean-reversion speed \( \eta \), we get the following dynamics (the case with Gaussian processes for the liquidity intensities is similar):

\[
d\lambda^Q(t) = dx(t) + g_{bid}dy^{bid}(t) + g_{ask}dy^{ask}(t)
= (\alpha - \beta x(t))dt + \sigma \sqrt{x(t)}dW_x(t) + g_{bid} \left( -\eta y^{bid}(t)dt + \sigma^{bid}dW_{y^{bid}}(t) \right)
+ g_{ask} \left( -\eta y^{ask}(t)dt + dW_{y^{ask}}(t) \right),
\]

where \( W_x, W_{y^{bid}} \) and \( W_{y^{ask}} \) are independent \( \mathbb{Q} \)-Brownian motions, which we, from now on, will give a superscript \( \mathbb{Q} \). Let us assume that all risk factors related to the default event itself are captured by \( x \). We then have that the pure \( \mathbb{P} \)-default intensity differs from the pure \( \mathbb{Q} \)-default intensity. We will, from now, on denote by \( x^Q \) the pure \( \mathbb{Q} \)-default intensity and by \( x^P \) the pure \( \mathbb{P} \)-default intensity (note that \( y^{bid} \) and \( y^{ask} \) are the same process under \( \mathbb{P} \) and \( \mathbb{Q} \), but with different drift terms).

Let us look at the case that we model the liquidity intensities as Ornstein-Uhlenbeck processes. We consider the change of measure from \( \mathbb{Q} \) to \( \mathbb{P} \) and we assume that the \( \mathbb{Q} \)-Brownian motions and \( \mathbb{P} \)-Brownian motions are related as follows:

\[
dW^Q_{y^{bid}}(t) = dW^P_{y^{bid}}(t) - (\delta^0_{bid} + \delta^1_{bid} y^{bid}(t))dt,
\]
\[
dW^Q_{y^{ask}}(t) = dW^P_{y^{ask}}(t) - (\delta^0_{ask} + \delta^1_{ask} y^{ask}(t))dt,
\]
\[
dW^Q_x(t) = dW^P_x(t) - \left( \frac{\delta^x_0}{\sqrt{x^Q(t)}} + \frac{\nu}{\sigma} \sqrt{x^Q(t)} \right) dt,
\]

We get that the \( \mathbb{P} \)-dynamics of the pure \( \mathbb{Q} \)-default intensity are of the form.

74
5.3. How to Use the Model?

\begin{align*}
\text{dx}^Q(t) &= \left(-\delta_0^Q \sigma - (\beta + \nu)x^Q(t)\right)dt + \sigma \sqrt{x^Q(t)}dW^P_x(t) \\
&= (\tilde{\alpha} - \tilde{\beta}x^Q(t))dt + \sigma \sqrt{x^Q(t)}dW^P_x(t),
\end{align*}

and the \(\mathbb{P}\)-dynamics of the liquidity processes are of the form

\begin{align*}
\text{dy}^l(t) &= \left[-\delta_0^l \sigma - (\eta + \delta_1^l \sigma^l)y(t)\right]dt + \sigma^l \text{dy}^P_y(t) \\
&= [\theta_l - \kappa_l y^l(t)]dt + \sigma^l \text{dy}^P_y(t),
\end{align*}

for \(l \in \{\text{bid}, \text{ask}\}\). Hence, under these affine measure changes, the \(\mathbb{P}\)-dynamics of the processes are of the same class as the \(Q\)-dynamics. We, therefore, get that the \(\mathbb{P}\)-dynamics of \(\lambda^Q\) are given by

\begin{align*}
\text{d}\lambda^Q(t) &= (\tilde{\alpha} - \tilde{\beta}x^Q(t))dt + \sigma \sqrt{x^Q(t)}dW^P_x(t) + g^\text{bid} \left((\theta^\text{bid} + \kappa^\text{bid} y^\text{bid}(t))dt + \sigma^\text{bid} dW^P_{y^\text{bid}}(t)\right) \\
&\quad + g^\text{ask} \left((\theta^\text{ask} + \kappa^\text{ask} y^\text{ask}(t))dt + \sigma^\text{ask} dW^P_{y^\text{ask}}(t)\right),
\end{align*}

where \(W^P_x, W^P_{y^\text{bid}}\) and \(W^P_{y^\text{ask}}\) are independent \(\mathbb{P}\)-Brownian motions and all the drift terms follow from Girsanov's theorem. The independence of the \(\mathbb{P}\)-Brownian motions follows from the multivariate Lévy’s characterization theorem, which is stated as follows:

**Theorem 5.2.** Let \(X\) be a \(d\)-dimensional continuous local martingale, with \(X_0 = 0\). Suppose that \([X]_t = tI\), where \(I\) is the \(d \times d\)-identity matrix and \([X]_t\) is the (co)variation matrix of \(X_t\). Then the components \(X^i_t, i = 1, \ldots, d\), are independent Brownian motions.

**Proof.** Let \(u \in \mathbb{R}^d\) and define \(f(t, x) = \exp \left(\frac{i}{2}u^\top x + \frac{1}{2}u^\top \mathbb{I}tu\right)\). Applying Itô’s lemma to \(f(t, X_t)\) gives

\[\text{df}(t, X_t) = \frac{1}{2}u^\top u f(t, X_t) + \sum_{j=1}^{d} i u_j f(t, X_t) dX^j_t - \frac{1}{2}u^\top u f(t, X_t) \]

\[= \sum_{j=1}^{d} i u_j f(t, X_t) dX^j_t.\]
Or stated differently,

\[
f(t, X_t) = 1 + \sum_{j=1}^{d} \int_{0}^{t} iu_j f(s, X_t) dX_t^j.
\]

We observe that \( f(t, X_t) \) is also a continuous local martingale and therefore there exists a localizing sequence of stopping times \( T^n \) such that for \( t > s \) we have

\[
E \left[ e^{iu\top X_t \wedge T^n + \frac{1}{2} u\top \mathbb{1}(t\wedge T^n)u} \Big| \mathcal{F}_s \right] = e^{iu\top X_s \wedge T^n + \frac{1}{2} u\top \mathbb{1}(s\wedge T^n)u}.
\]

The sequence \( e^{iu\top X_s \wedge T^n + \frac{1}{2} u\top \mathbb{1}(s\wedge T^n)u} \) is bounded and converges almost surely (and in \( L_1 \)) to \( e^{iu\top X_s + \frac{1}{2} u\top \mathbb{1}s} \). Furthermore, the conditional expectation converges in \( L_1 \) to the conditional expectation of the limit and therefore we have

\[
E \left[ e^{iu\top X_t + \frac{1}{2} \mathbb{1}tu} \Big| \mathcal{F}_s \right] = e^{iu\top X_s + \frac{1}{2} \mathbb{1}s}.
\]

Now we get

\[
E \left[ e^{iu\top (X_t - X_s)} \Big| \mathcal{F}_s \right] = e^{-\frac{1}{2} u\top \mathbb{1}(t-s)u}.
\]

Let \( Y \) be an arbitrary \( \mathbb{R}^d \)-valued, \( \mathcal{F}_s \)-measurable random variable and let \( \psi_Y(u) \) denote its characteristic function with \( u \in \mathbb{R}^d \). We get that the joint characteristic function of \( X_t - X_s \) and \( Y \) is given by

\[
\psi_{Y, X_t - X_s}(u, v) = E \left[ e^{iu\top Y + iv\top (X_t - X_s)} \right] = E \left[ e^{iu\top Y} e^{iv\top (X_t - X_s)} \right] = E \left[ e^{iu\top Y} E \left[ e^{iv\top (X_t - X_s)} \Big| \mathcal{F}_s \right] \right] = E \left[ e^{iu\top Y} \right] e^{-\frac{1}{2} v\top \mathbb{1}(t-s)v} = \psi_Y(u) \psi_{X_t - X_s}(v).
\]
5.3. How to Use the Model?

We can thus conclude that \((X_t - X_s)\) is independent from \(\mathcal{F}_s\), since \(Y\) was chosen arbitrarily.

We note that \(\psi(X_t - X_s)\) is the characteristic function of a multivariate normally distributed random vector with mean zero and covariance matrix \(I(t - s)\).

Because of the properties of the multivariate normal distribution, we know that each component of \((X_t - X_s)\), \((X^i_t - X^i_s)\), \(i = 1, \ldots, d\), is normally distributed with mean zero and variance \((t - s)\) and is independent from \((X^j_t - X^j_s)\), \(j \neq i\). Since this is true for all \(t > s\), we conclude that each component \(X^i\) is a Brownian motion and independent from \(X^j, j \neq i\).

So, since \(W^Q_x, W^Q_{y_{\text{bid}}}\) and \(W^Q_{y_{\text{ask}}}\) are independent \(Q\)-Brownian motions we have

\[ [W^Q_x]_t = [W^Q_{y_{\text{bid/ask}}}]_t = t \]

and

\[ [W^Q_x, W^Q_{y_{\text{ask/bid}}}]_t = [W^Q_{y_{\text{ask}}}, W^Q_{y_{\text{bid}}}]_t = 0 \text{ for all } t > 0. \]

After changing the measure from \(Q\) to \(P\), the \(P\)-Brownian motions \(W^P_x, W^P_{y_{\text{bid}}}\) and \(W^P_{y_{\text{ask}}}\) have the same covariations under \(P\) as they had under \(Q\) and, hence, they are again independent Brownian motions, but now under \(P\).

So if we make assumptions on the change of measure, similar to the example above, we know that the pure intensities under \(P\) are of a similar form as under \(Q\) (this is also possible in the Arithmetic Brownian motion liquidity set-up). Furthermore, as we have showed above, the pure intensities are also independent under \(P\) (as their driving Brownian motions are independent) and, therefore, \(\lambda^Q\) has the same structure under \(P\) as under \(Q\): it is again the weighted sum of the pure intensities with the correlation factors as weights. We can now analyze the \(P\)-dynamics of the pure intensities separately and by adding them (and multiplying by the corresponding correlation factors) we get the \(P\)-dynamics of the correlated default intensity \(\lambda^Q\).

The question that now arises is how we can find the \(P\)-dynamics of the pure liquidity intensities and of the \(Q\)-default intensity? To see this, we note that the calibration gives us time series for pure liquidity intensities and the pure \(Q\)-default intensity. By definition, if we estimate the process dynamics parameters of a stochastic process based on a historical time series, we are considering the real-world behavior of the process. Therefore, having the time series of the pure intensities gives us a way to extract the \(P\)-dynamics of the intensities (see, for example, Brigo and Mercurio (2006) [9] for an application of this observation to short-rate interest rate models).

Since we have modeled the pure default intensity as a CIR process and the pure liquidity intensities either as driftless Arithmetic Brownian motions or Ornstein-Uhlenbeck processes,
we know their transition densities and, therefore, we can derive exact maximum likelihood estimators for the drift parameters. In chapter 8, we will discuss maximum likelihood estimators based on historic observations of CIR, OU and Arithmetic Brownian motion processes, and we will perform a Monte Carlo study to test the maximum likelihood estimators. After that, we will apply the results to the obtained time series of the pure intensities, and we will derive the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-default intensity process.

5.4 Chapter Summary

In this chapter, we presented the credit-liquidity model, which is an intensity-based model that incorporates liquidity effects by means of extra liquidity discount factors. We assumed that the CDS bid and ask market have different liquidity discount factors by modeling separate bid and ask liquidity intensities. We were able to derive formulas for both the CDS bid and ask premia, which can later be calibrated to observed bid and ask premia.

Concerning the stochastic components of our model, we assumed that risk-free interest rates are independent from the credit and liquidity factors and that interest rate discount factors can be obtained from discount curves based on USD swap rates. This allows us to ignore the modeling of an interest rate process. Our model does, however, explicitly model the default and liquidity intensities and allows for a correlation structure between them. We model all intensities as affine combinations of affine diffusions, which, in combination with the proposed correlation structure, allows us to compute the formulas for the bid and ask premia completely analytically.

The two objectives of our study are to assess the impact of liquidity on the CDS premium and to compute the implied default probability of the reference entity. We showed that our model allows for a natural decomposition of the CDS mid premium into a credit and a liquidity component. Furthermore, by calibrating the model-implied bid and ask premia to observed bid and ask premia, we can generate time series of the $\mathbb{Q}$-default and liquidity intensities from which, under some assumptions on the Radon-Nikodým process, we can obtain the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-default intensity process. In combination with Rabobank’s model of the default event risk premium, we can compute real-world default probabilities from these dynamics.
6

Data Description and Calibration Procedure

We will use our model to investigate the credit and liquidity components of CDSs on the Turkish and Brazilian governments, which are both relevant for Rabobank. An advantage of looking at emerging countries such as Turkey and Brazil is that both their government debt and CDSs are denominated in USD and, therefore, there are no currency issues in our analysis. In this chapter, we will discuss what data we will use for calibrating our model and we will also explain how this calibration will be done. The results will be given in the next chapter.

6.1 Data Description

By the assumption we made on the independence of the interest rates from the liquidity and credit components, we are able to calibrate the interest rate part separately from the credit and liquidity parts. We will assume that the risk-free discount factors can be obtained from bootstrapping the USD swap rate curve.

For each (business) day in the period 01-01-2008 until 28-02-2014, we have data on the term structure of the USD swap rates (with maturities ranging from 1 day up to 60 years). For each of these days, we construct a (slightly more than) 10 year discount curve on a daily basis by bootstrapping the term structure of the swap rate data and by interpolating the obtained discount factors. The terms \( \mathbb{E}^Q \left[ \hat{D}(t, T_i) | \mathcal{F}_t \right] \) in formulas (5.6) and (5.7) can now be obtained.

1CDSs on European countries, for example, are typically denominated in USD, whereas the underlying government debt is denominated in Euros. This complicates the analysis, since foreign exchange risk has to be accounted for. Especially in distressed economies this risk cannot be ignored.
by looking at the discount curve of day \( t \) and taking the value of this discount curve at maturity \( T_i \).

In order to calibrate the credit and liquidity components, we will use bid and ask data of the 2, 3, 5 and 10 year CDSs. As can be seen from Figure 2.9, these maturities, in general, provide the highest liquidity in the sovereign CDS market and, therefore, we conjecture that using these maturities gives the most reliable results. By using CDSs with different maturities, we assume that (at least the overlapping parts) of the CDSs are influenced by the same credit risk and liquidity risk factors. This seems reasonable, since, for example, the default risk in the first two years should be the same for the 2 year CDS as for the 10 year CDS on the same reference entity.

We thus obtain the bid and ask quotes for the 2, 3, 5 and 10 year CDSs on the Turkish and Brazilian governments for the period 01-01-2008 until 28-02-2014. If, on a certain day, a quote is missing, we drop the entire day from our sample. In this way, we obtain 1483 days of data on the Turkish CDSs and 1571 days of data on the Brazilian CDSs, where each day of data thus consists of 4 bid and 4 ask quotes. Note that a business year consists of roughly 250 days and, therefore, we almost have full data for every business day (there are 1584 business days in this period).

Figures 6.1 and 6.2 show the levels of the mid premia of the 2, 3, 5 and 10 year CDSs on the Turkish and Brazilian governments, respectively, for the period 01-01-2008 until 28-02-2014. We see that, during this whole period, the CDS term structure is increasing for both sovereigns. Furthermore, the behaviour of the mid premia is very similar across the different maturities: they all move up and down together. It can also be seen that, during 2008 and the first half of 2009 (the first 400 days), there is a peak in the CDS premia and the behaviour is much more volatile than in the rest of the period. This is mainly due to the start of the global financial crisis.

Figures 6.3 and 6.4 show the development of bid-ask spreads of the 5 year CDSs on the Turkish and Brazilian governments respectively (the bid-ask spreads of the other maturities give similar pictures). It can be seen that in times of high CDS (mid) premia, the bid-ask spreads are relatively high too (see, for example, the first 500 days). From 2010 onwards, the size of the bid-ask spread seems to stabilize for both countries.

Table 6.1 gives some summary statistics of the mid premia and the bid-ask spreads for the whole sample period. Table 6.2 gives the same summary statistics, but with the data from 2008 and the first half of 2009 excluded. As can be seen from these tables, the bid-ask spreads and mid premia behave less volatile for the subsample period compared to the full sample period. Furthermore, we see that the CDS premia on the Brazilian government are lower than on the Turkish government.
6.1. Data Description

Figure 6.1: CDS mid premia for 2, 3, 5 and 10 years CDS on Turkish government for the period 01-01-2008 until 28-02-2014.

Figure 6.2: CDS mid premia for 2, 3, 5 and 10 years CDS on Brazilian government for the period 01-01-2008 until 28-02-2014.
Chapter 6. Data Description and Calibration Procedure

Figure 6.3: Bid-ask spread in basis points of the 5 year CDS on the Turkish government.

Figure 6.4: Bid-ask spread in basis points of the 5 year CDS on the Brazilian government.
### 6.1. Data Description

**Summary Statistics Mid Premia and Bid-Ask Spreads 01-01-2008 until 28-02-2014**

#### Turkey (1483 complete days of data)

<table>
<thead>
<tr>
<th>Contract Length</th>
<th>2Y</th>
<th>3Y</th>
<th>5Y</th>
<th>10Y</th>
<th>2Y</th>
<th>3Y</th>
<th>5Y</th>
<th>10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>168.68</td>
<td>187.69</td>
<td>226.19</td>
<td>259.64</td>
<td>8.64</td>
<td>10.33</td>
<td>5.45</td>
<td>9.02</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>102.53</td>
<td>97.97</td>
<td>90.28</td>
<td>83.04</td>
<td>8.48</td>
<td>7.64</td>
<td>5.44</td>
<td>6.12</td>
</tr>
<tr>
<td>Max</td>
<td>807.52</td>
<td>814.21</td>
<td>824.61</td>
<td>819.65</td>
<td>66.18</td>
<td>63.84</td>
<td>63.52</td>
<td>50.00</td>
</tr>
<tr>
<td>Min</td>
<td>45.73</td>
<td>71.35</td>
<td>112.55</td>
<td>145.30</td>
<td>0.03</td>
<td>0.34</td>
<td>0.06</td>
<td>0.98</td>
</tr>
</tbody>
</table>

#### Brazil (1571 complete days of data)

<table>
<thead>
<tr>
<th>Contract Length</th>
<th>2Y</th>
<th>3Y</th>
<th>5Y</th>
<th>10Y</th>
<th>2Y</th>
<th>3Y</th>
<th>5Y</th>
<th>10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>103.78</td>
<td>122.76</td>
<td>157.54</td>
<td>191.80</td>
<td>6.63</td>
<td>6.29</td>
<td>4.17</td>
<td>5.87</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>65.05</td>
<td>67.63</td>
<td>69.25</td>
<td>67.34</td>
<td>5.08</td>
<td>4.72</td>
<td>2.98</td>
<td>3.19</td>
</tr>
<tr>
<td>Max</td>
<td>532.44</td>
<td>547.63</td>
<td>586.41</td>
<td>601.16</td>
<td>42.36</td>
<td>49.74</td>
<td>36.42</td>
<td>38.15</td>
</tr>
<tr>
<td>Min</td>
<td>42.65</td>
<td>60.00</td>
<td>85.17</td>
<td>115.13</td>
<td>0.07</td>
<td>0.04</td>
<td>0.22</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Table 6.1: Summary statistics of the mid-premia and bid-ask spreads of our dataset. The data ranges from 01-01-2008 until 28-02-2014.
Table 6.2: Summary statistics of the mid-premia and bid-ask spreads of our dataset. The data ranges from 01-06-2009 until 28-02-2014.

### 6.2 Calibration Procedure

In this section, we will discuss how we will calibrate the model parameters to the above described data. First, let us recall the model-implied bid and ask premia formulas

\[
s_{\text{ask}} = (1 - R) \cdot \sum_{i=1}^{n} \mathbb{E}^Q \left[ \mathbb{E}^Q \left[ \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \bigg| \mathcal{F}_t \right] \right] \left[ \bar{D}(t, T_i) \bigg| \mathcal{F}_t \right] \left[ \mathbb{E}^Q \left[ L^\text{ask}(t, T_i) \bigg| \mathcal{F}_t \right] \right] \tag{6.1}
\]

and

\[
s_{\text{bid}} = (1 - R) \cdot \sum_{i=1}^{n} \delta_i \mathbb{E}^Q \left[ \mathbb{E}^Q \left[ \mathbb{E}^Q \left[ \left( \bar{P}(t, T_{i-1}) - \bar{P}(t, T_i) \right) \bigg| \mathcal{F}_t \right] \right] \right] \left[ \bar{D}(t, T_i) \bigg| \mathcal{F}_t \right] \left[ \mathbb{E}^Q \left[ \bar{L}^\text{bid}(t, T_i) \bigg| \mathcal{F}_t \right] \right] \tag{6.2}
\]

As discussed above, the interest rate discount terms \( \mathbb{E}^Q \left[ \bar{D}(t, T_i) \bigg| \mathcal{F}_t \right] \) are obtained directly from the discount curves, which are constructed from USD swap rates. The expressions \( \mathbb{E}^Q \left[ \bar{P}(t, T_i) \bigg| \mathcal{F}_t \right] \) and \( \mathbb{E}^Q \left[ \bar{P}(t, T_i) L^l(t, T_i) \bigg| \mathcal{F}_t \right] \) (\( l \in \{ \text{bid}, \text{ask} \} \)), however, are given by the formulas stated in Appendix C and depend on the process parameters, the correlation factor parameters and the pure intensities \( x(t) \), \( y^{\text{bid}}(t) \) and \( y^{\text{ask}}(t) \).
6.2. Calibration Procedure

Since, by definition, the intensities are not observable, we have to consider \(x(t)\), \(y_{\text{bid}}(t)\) and \(y_{\text{ask}}(t)\) as additional parameters for every day \(t\) in our sample. So, depending on whether we model the liquidity parameters as driftless Arithmetic Brownian motions or as Ornstein-Uhlenbeck processes with the same mean-reversion speeds and zero mean-reversion levels, we have the following parameters we need to calibrate: in the first case we need to calibrate the five process parameters \(\{\alpha, \beta, \sigma, \sigma_{\text{bid}}, \sigma_{\text{ask}}\}\), the six correlation factor parameters \(\{f_{\text{bid}}, f_{\text{ask}}, g_{\text{bid}}, g_{\text{ask}}, \omega_{\text{bid},\text{ask}}, \omega_{\text{ask},\text{bid}}\}\) and the pure intensities \(\{x(t), y_{\text{bid}}(t), y_{\text{ask}}(t)\}\) for every day \(t\) in the sample and in the second case we need to calibrate the six process parameters \(\alpha, \beta, \sigma, \eta, \sigma_{\text{bid}}, \sigma_{\text{ask}}\), the six correlation factor parameters \(\{f_{\text{bid}}, f_{\text{ask}}, g_{\text{bid}}, g_{\text{ask}}, \omega_{\text{bid},\text{ask}}, \omega_{\text{ask},\text{bid}}\}\) and the pure intensities \(\{x(t), y_{\text{bid}}(t), y_{\text{ask}}(t)\}\) for every day \(t\) in the sample. We take the first day of our sample as \(t = 0\) and the last day of our sample as \(T\).

We will propose a grid search procedure to calibrate our model parameters and the (pure) intensity time series. In each grid point, we will fix the values of the the process parameters and, given these values, we will find the correlation factors and the daily values of the pure intensities by least squares methods. We will take the grid point with the lowest objective value (which will be specified below) and create a finer grid around this point and repeat the whole procedure. The details will be given in the next subsection.

6.2.1 Calibration Steps

Let us consider a grid with fixed process parameters in each grid point. By allowing each process parameter to have different values, we have that the grid has a dimension equal to the number of process parameters. In the case of the Arithmetic Brownian motion liquidity intensities, we thus have a 5-dimensional grid and in the case of the Ornstein-Uhlenbeck liquidity intensities, we have a 6-dimensional grid. In each grid point we take the following steps:

1. First, we will set the correlation factor parameters equal to zero, which implies that the liquidity and default intensities are independent and equal to the pure intensities \(\lambda(t) = x(t)\), etc.). Now, given the process parameters and the zero correlation factors, we will determine the values of the intensities cross-sectionally (i.e., day by day) by minimizing the following sums of squares:

\[
\begin{align*}
(x(t), y_{\text{bid}}(t), y_{\text{ask}}(t)) &= \text{argmin} \sum_{i=1}^{8} \left(s_{i}^{\text{mod}}(t) - s_{i}^{\text{obs}}(t)\right)^{2}, \quad \forall t = 0, 1, \ldots T. \tag{6.3}
\end{align*}
\]

Here \(s_{i}^{\text{mod}}(t)\) denotes the model-implied value of premium \(i\) on day \(t\), where \(i\) is the bid or ask premium of one of the 4 maturities (2, 3, 5 and 10 years) and, similarly, \(s_{i}^{\text{obs}}(t)\) denotes the corresponding observed premium at day \(t\). We could thus, for example, take \(i = 1\) as the bid premium of the 2 year CDS, \(i = 2\) as the ask premium of the 2 year CDS, \(i = 3\) as the bid premium of the 3 year CDS, etc.
2. In the second step, we will find the correlation factor parameters by minimizing the following sum of squares:

\[
(f, g, \omega_{bid}, \omega_{ask}) = \arg\min_{T} \sum_{t=0}^{8} \sum_{i=1}^{8} (s_{i}^{mod}(t) - s_{i}^{obs}(t))^{2},
\]

where we keep the process parameters and intensities, which we found in step 1, fixed.

3. We will now repeat steps 1 and 2, but we will not set the correlation factor parameters equal to zero in step 1. Instead, we will take the correlation factor parameters equal to the values found in step 2 of the previous iteration. We will repeat these steps until either a maximum number of iterations or a convergence criterium is reached. Typically, we set the maximum number of iterations equal to 7 and the convergence criterium is that the value of (6.4) does not change much between two iterations.

When we have found the optimum in a certain grid point, we move on to the next grid point and repeat the steps described above. In the end, we take the grid point with the lowest value of the right-hand side (6.4) and we build a finer grid around this point and repeat the same steps.

Bühler and Trapp (2008) [11] and Badaoui, Cathcart and El-Jahel (2013) [5] also use similar grid search methods to calibrate their models, but there are some differences with our proposed algorithm. First of all, both Bühler and Trapp and Badaoui, Cathcart and El-Jahel calibrate their model on bond data and 5 year CDS data, whereas we do not use bond data at all and we use CDSs with a range of different maturities. Bühler and Trapp did not have much choice, since they investigated the corporate CDS market and the only liquidly traded CDSs in that market are the 5 year CDSs (see Figure 2.9). We, on the other hand, use sovereign CDSs, which are (relatively) more actively traded across different maturities. Although we loose the option to analyze liquidity spillovers between the bond and the CDS markets, we think that credit default swaps are better for analyzing credit risk than bonds and, ultimately, this is what we are interested in the most. Furthermore, the above described calibration procedure allows us to calibrate the model for each country separately. Badaoui, Cathcart and El-Jahel (2013) [5], on the other hand, group countries of certain rating classes together in their calibration procedure, but, since we are interested in country-specific default probabilities, this would not be suitable for us.

Another difference is that Bühler and Trapp, for example, find the values of the pure intensities by minimizing the sum of squares of the difference between the model premia and the actual premia in one large optimization, whereas we obtain the intensities cross-sectionally by small daily optimizations, i.e., we determine \(x(t), y^{bid}(t)\) and \(y^{ask}(t)\) by optimizing against quotes on day \(t\), while Bühler and Trapp find all intensities for all days at once by looking at all quotes at once. We think it is better to determine the intensities cross-sectionally, since they are the values of a stochastic process on a certain time and, therefore, they should only depend on data
6.2. Calibration Procedure

on that same time. This approach is in line with Badaoui, Cathcart and El-Jahel (2013) [5]. Both Bühler and Trapp and Badaoui, Cathcart and El-Jahel are particularly vague about how they find the correlation factor parameters\(^2\) and, therefore, we cannot compare this step of our algorithm to theirs. Iterating between steps 1 and 2, however, is more in line with Bühler and Trapp than with Badaoui, Cathcart and El-Jahel, who only compute the correlation factor parameters once per grid point.

6.2.2 Implementation Issues

The proposed grid search algorithm is computationally expensive and, therefore, the implementation is non-trivial. In each grid point we need to run \(T \) times a least squares optimization to find the daily values of the intensities. After that, we need to run one large least squares optimization to find the correlation factor parameters and then we need to repeat these steps multiple times (typically at most seven times). Furthermore, the number of grid points grows exponentially in the number of process parameters. If we want to calibrate the model in which the liquidity intensities are driftless Arithmetic Brownian motions, we have five process parameters \(\{\alpha, \beta, \sigma, \sigma^{bid}, \sigma^{ask}\}\). If we allow each parameter to take 10 different values, we have a grid with \(10^5\) grid points. If, on the other hand, we want to calibrate the Ornstein-Uhlenbeck variant of the model, we have six process parameters, \(\{\alpha, \beta, \sigma, \eta, \sigma^{bid}, \sigma^{ask}\}\), and the grid already consists of \(10^6\) grid points. If each grid point would take only 5 seconds to run, the total time of our calibration would be approximately 1400 hours. This already explains why we wanted the number of process parameters to be low and why we did not take the most general forms of the processes for the liquidity intensities, since then the number of grid points (and also the running time) would explode.

In order to reduce the running time, we have tried to program the calibration procedure in an efficient way. First of all, since the computations in different grid points are independent from, but very similar to, each other, we have implemented the calibration in a parallel setting. In this way, multiple grid points can be dealt with at the same time, thereby reducing the running time by a factor equal to the number of threads. Typically, a decent quad core laptop can run 8 threads at the same time, thereby reducing the total running time by a factor 8.

Another major reduction of the running time was achieved by choosing clever starting points for the optimization procedures. The least squares optimization we used, is based on the Levenberg-Marquardt algorithm, which becomes more efficient when the starting point is chosen closer to the actual minimum, since then the algorithm needs less iterations to converge to the optimal point. If we, for example, look at the values of the intensities, we conjecture that it is likely that intensities on two consecutive days have values that are not very far away from each other.

\(^2\)We tried to contact the authors about their calibration procedures, but Bühler and Trapp did not reply. Badaoui, on the other hand, nicely replied with some general tips and tricks, but did not want to go into detail.
other (unless there is a jump, but, in general, this will not happen frequently). By choosing the optimized values of the intensities of day \( t - 1 \) as starting point of the optimization procedure for the intensities on day \( t \), step 1 of the optimization algorithm in each grid point became much more efficient than by naively taking some random starting point.

Since the Levenberg-Marquardt algorithm is an iterative algorithm, and since we have a large dataset (approximately 1500-8 quotes per country), the pricing formulas are called many times. Within each grid point, some parts of these pricing formulas are, however, independent from the parameters that are optimized and, therefore, have the same value in each function call. By precalculating all these constant terms and looking up the values when needed, we prevent ourselves from calculating the same (sub)expressions every time the price functions are called. This implementation trick also significantly improved the running time of our program.

A last major improvement of our implementation was reached when we implemented the whole procedure in C++. Before this, we implemented the calibration procedure in the Python programming language, but by switching to C++ we reduced the running time by a factor 50.

All-in-all, each grid point still takes roughly 5 seconds to run on a normal laptop and, therefore, we cannot easily run dense grids or extend our model to more general stochastic processes with more parameters. Our approach is, therefore, to run coarse grids and to ‘zoom in’ a few times. By this we mean that, when we find the optimal point in a grid, we can create a finer grid around it (with tighter bounds for the process parameters) and repeat the grid search procedure. In this way one zooms into the area in which the original optimal point was found.

The approach of zooming into the grid is cheaper than running a more dense grid. To see this, consider the following example:

Suppose we want to calibrate the driftless Arithmetic Brownian motion model variant, which has 5 process parameters. If we conjecture that the values of all the process parameters lie between 0 and 1 and if we take steps of 0.1 in each process parameter, we have a grid with \( 11^5 = 161051 \) points. The optimal grid point now has a precision of 0.05 in the process parameters. With this we mean that, if a process parameter has a value of either 0.08 or 0.09 in the true optimal point, the above described procedure will find 0.1 in both cases, since we only look at the values 0, 0.1, 0.2 . . . , 1.

If, on the other hand, we would allow each process parameter to take only 6 different values in the grid, then the grid would have only \( 6^5 = 7776 \) grid points. The precision of the optimal point is now at a level of roughly 0.167/2 \( \approx 0.08 \) by a similar argument as above. If we now construct a tighter grid of the same size around the optimal point that, for example, takes the optimal parameter values plus or minus 0.25 as new bounds, then we have that the optimal point in this smaller grid has a precision of approximately 0.04. We thus see that the total number of grid points we had to consider was \( 7776 \cdot 2 = 15552 \), whereas the precision of the parameters in the final optimal grid point is higher than in the case of the dense grid, which
6.3. Chapter Summary

required the evaluation of more grid points.

The point of the above example is that zooming into the grid, by creating smaller grids around
the optimal point of the previous grid, is computationally cheaper than running a more dense
grid. The danger of this approach is, however, that, in a coarse grid, one can end up in a local
minimum more easily than in a more dense grid. Unless the implementation becomes much
faster we cannot, however, run dense grids and we are restricted to the approach of zooming
into the grid.

A last note that we want to make is that, by making use of CDSs of different maturities, the
decomposition of the CDS mid premium into a credit and a liquidity component, which was
proposed in chapter 5, will probably work best for CDSs with the shortest maturities. To see
this, we note that the credit and liquidity risk of the first two years of, say, a 10 and 5 year CDS
contract should be the same as the credit and liquidity risk of the 2 year CDS itself. We thus
see that the first 2 years overlap for all the 2, 3, 5 and 10 year CDSs, whereas, for example, the
last 5 year of the 10 year CDS do not overlap with any of the others. We, therefore, conjecture
that the 2 year CDS will be calibrated best in the sense that the liquidity and default intensities
make the most sense for this maturity and, therefore, the decomposition into a credit and a
liquidity part too.

6.3 Chapter Summary

In this chapter, we discussed the data that we will use to test our model on and we discussed
how the calibration procedure works and what implementation issues we encountered.

We want to test our model on Turkey and Brazil. As stated before, we will use USD swap
rates to create a USD interest rate discount curve to account for all interest related discount
factors. We will, furthermore, use the bid and ask premia of the 2, 3, 5 and 10 year CDSs
to calibrate our bid and ask premia formulas. We obtained these bid and ask premia for the
period 01-01-2008 until 28-02-2014. We saw that in the start of this period, the CDS (mid)
premia and the bid-ask spreads were relatively high for both countries. This can most likely be
attributed to the beginning of the global financial crisis. We gave some summary statistics on
the mid premia and the bid-ask spreads for the CDSs of all maturities for both the complete
sample as well as the ‘post-crisis’ subsample, which ranges from 01-06-2009 until 28-02-2014.

In order to calibrate the model-implied bid and ask premia to the observed premia, we noted
that we do not only have to find the parameters of the intensity processes and the correlation
factor parameters, but also the (daily) values of the pure default and liquidity intensities. This
means that the number of parameters that have to be calibrated is enormous, since for every
day of data, we have three intensity values that we need to fit.
We proposed a grid search algorithm that allows us to calibrate all the parameters and intensity values in different steps. In each grid point we fix the process parameters and use least squares optimization procedures to find the values of correlation factor parameters and time series of the intensities that give the best fit of our model-implied formulas to the observed premia. We will, in the end, choose the grid point that gives the best fit and we will create a finer grid around this and repeat the procedure (possibly multiple times). In this way, the optimal process parameters have those values that correspond to the best grid point we eventually find.

Since we fix the process parameters in each grid point and since we want to check different values of each process parameter, the grid has a dimension equal to the number of process parameters. This means that the number of grid points one needs to consider grows exponentially with the number of process parameters and, with this, also the computational time. This explains why we chose to model the liquidity intensities with as few process parameters as possible. Furthermore, because of the computational time of calibration procedure, we are not able to consider very dense grids, which has as a danger that we might converge to a local minimum.
Results

In this chapter, we will give the results of our model calibration to the CDS data on the Turkish and Brazilian governments. As we saw in the previous chapter, the start of the crisis significantly impacted the CDS premia and bid-ask spreads and in order to exclude these temporary effects, we will look at the calibration results on the data subsamples from 01-06-2009 until 28-02-2014. We will first show how well the fit of the model to the data is and after that we will discuss the obtained parameter values and intensities. We will end this chapter with the decomposition of the 2 year CDS mid premium into a credit and a liquidity part.

7.1 Calibration Results

As explained in the previous chapter, we calibrated our model-implied formulas for the bid and ask premia on observed bid and ask premia of the 2, 3, 5 and 10 year CDSs of Turkey and Brazil. The objective function that we try to minimize is given by

$$\sum_{t=0}^{T} \sum_{i=1}^{8} (s_{i}^{\text{mod}}(t) - s_{i}^{\text{obs}}(t))^2,$$

(7.1)

where $s_{1}^{\text{mod}}(t)$ denotes the 2 year model-implied bid premium at day $t$, $s_{2}^{\text{mod}}(t)$ the 2 year model-implied ask premium at day $t$, $s_{3}^{\text{mod}}$ the model-implied 3 year bid premium at day $t$, etc., and, similarly, $s_{i}^{\text{obs}}(t)$ denotes the corresponding observed 2 year bid premium at day $t$, $s_{2}^{\text{obs}}$ the observed 2 year ask premium at day $t$, etc.

Table 7.1 gives the values of the objective function for the different model set-ups for Brazil and Turkey.
Chapter 7. Results

Value Objective Function

<table>
<thead>
<tr>
<th></th>
<th>Brazil</th>
<th>Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arithmetic BM</td>
<td>OU</td>
</tr>
<tr>
<td></td>
<td>97998.1</td>
<td>112874</td>
</tr>
</tbody>
</table>

Table 7.1: Value of the objective function (7.1) after calibration.

Recalling that we have 1214 days of observations for Brazil, we see that, using (7.1) and table 7.1, the bid and ask premia are, on average, mispriced with 3.17 basis points in the Arithmetic Brownian motion case and with 3.41 basis points in the OU case. For Turkey we have 1141 days of complete data and this implies that, on average, we misprice the bid and ask premia with 4.48 and 4.04 basis points in the Arithmetic Brownian motion and Ornstein-Uhlenbeck set-ups, respectively.

Tables 7.2 and 7.3 give more details on the model fit for Brazil in the Arithmetic Brownian motion and OU set-ups, respectively. If we look at the average relative errors, we see that, in the Arithmetic Brownian motion case, the average relative error over all premia is 2.05% and, in the OU case, the average relative error over all premia is 2.16%.

| Pricing Errors Brazil with Arithmetic Brownian Motion Liquidity Intensities | Absolute Errors (|model-implied premium - observed premium|) | Relative Errors ((|model-implied premium - observed premium|)/observed premium) |
|-------------------------------------------------|---------------------------------|---------------------------------|
| 2Y Bid | 2Y Ask | 3Y Bid | 3Y Ask | 5Y Bid | 5Y Ask | 10Y Bid | 10Y Ask | 2Y Bid | 2Y Ask | 3Y Bid | 3Y Ask | 5Y Bid | 5Y Ask | 10Y Bid | 10Y Ask |
| Mean   | 3.3291 | 2.8155 | 1.7042 | 1.0160 | 3.4620 | 3.3166 | 0.8713 | 1.0262 |
| Std.   | 3.0431 | 2.2855 | 1.1127 | 0.7814 | 2.8592 | 2.6262 | 0.8598 | 0.9136 |
| Min    | 0.0028 | 0.0059 | 0.0080 | 0.0016 | 0.0011 | 0.0000 | 0.0028 | 0.0013 |

Table 7.2: Absolute and relative pricing errors Brazilian CDS premia using driftless Arithmetic Brownian motion for the liquidity intensities.
7.1. Calibration Results

### Pricing Errors Brazil with Ornstein-Uhlenbeck Liquidity Intensities

<table>
<thead>
<tr>
<th></th>
<th>2Y Bid</th>
<th>2Y Ask</th>
<th>3Y Bid</th>
<th>3Y Ask</th>
<th>5Y Bid</th>
<th>5Y Ask</th>
<th>10Y Bid</th>
<th>10Y Ask</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absolute Errors</strong> (</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.9505</td>
<td>2.3993</td>
<td>0.9311</td>
<td>0.8681</td>
<td>3.3014</td>
<td>2.5510</td>
<td>1.047</td>
<td>0.7948</td>
</tr>
<tr>
<td>Min</td>
<td>0.0025</td>
<td>0.0012</td>
<td>0.0021</td>
<td>0.0039</td>
<td>0.0132</td>
<td>0.0051</td>
<td>0.0025</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2Y Bid</th>
<th>2Y Ask</th>
<th>3Y Bid</th>
<th>3Y Ask</th>
<th>5Y Bid</th>
<th>5Y Ask</th>
<th>10Y Bid</th>
<th>10Y Ask</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relative Errors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0417</td>
<td>0.0294</td>
<td>0.0122</td>
<td>0.0122</td>
<td>0.0344</td>
<td>0.0258</td>
<td>0.0089</td>
<td>0.0069</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0407</td>
<td>0.0269</td>
<td>0.0115</td>
<td>0.0087</td>
<td>0.0225</td>
<td>0.0194</td>
<td>0.0047</td>
<td>0.0046</td>
</tr>
<tr>
<td>Max</td>
<td>0.2489</td>
<td>0.1726</td>
<td>0.0596</td>
<td>0.0548</td>
<td>0.1045</td>
<td>0.0828</td>
<td>0.0224</td>
<td>0.0208</td>
</tr>
<tr>
<td>Min</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 7.3: Absolute and relative pricing errors Brazilian CDS premia using Ornstein-Uhlenbeck processes for the liquidity intensities.

We see that the results in both set-ups are very similar and that there are no differences worth mentioning. Since the fit of the 2 year bid premia is the worst in both set-ups, we take a closer look into these calibration results. Figure 7.1 shows the observed and model-implied 2 year bid premia and the corresponding absolute error in both set-ups. Looking at the graphs on the left-hand side, we see that the model-implied premia follow a very similar pattern as the observed premia in both model set-ups. Looking at the graphs on the right-hand side, we see that the ‘bad’ summary statistics can mainly be attributed to the last part of the sample period, since the absolute errors are large here. The cause of this is that the 2 year bid premium is very volatile in this period and, therefore, more prone to errors.
Figure 7.1: Absolute pricing errors on Brazilian 2Y bid premia in Arithmetic Brownian motion set-up (upper two figures) and in Ornstein-Uhlenbeck set-up (lower two figures).
Tables 7.4 and 7.5 show the summary statistics for the absolute and relative pricing errors for Turkey in both model set-ups. We see an average relative error of 2.44% in the Arithmetic Brownian motion set-up and of 2.14% in the OU set-up. Overall, the results are very similar to that of Brazil.

<table>
<thead>
<tr>
<th>Pricing Errors Turkey with Arithmetic Brownian Motion Liquidity Intensities</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Absolute Errors (</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>**Relative Errors (</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td>Min</td>
</tr>
</tbody>
</table>

Table 7.4: Absolute and relative pricing errors Turkish CDS premia using driftless Arithmetic Brownian motion for the liquidity intensities.
Table 7.5: Absolute and relative pricing errors Turkish CDS premia using Ornstein-Uhlenbeck processes for the liquidity intensities.

Since we are not trying to use our model for trading purposes, the above mentioned pricing errors are not a problem. The errors can mainly be attributed by the fact that our calibration procedure does not allow us to search the whole process parameter space, but only certain points (the values that are fixed in each grid point). Furthermore, since we made some simplifying assumptions to construct the pricing formulas, there is still room for improvement by relaxing these assumptions. For example, we now assumed that the payments made by the protection seller in the case of a default are made at the first premium payment date after the default. In reality, this is of course not the case and one could model this more realistically against the cost of more (computationally) expensive formulas. Typically, the pricing formulas will have expressions of integrals with respect to all the possible default times and, in general, these integrals have to be solved numerically.

Table 7.6 gives the values of the calibrated parameters. Looking at the process parameters, we see that the values of all the volatility parameters and the mean-reversion level parameter $\alpha$ are very similar for both countries and in both model set-ups. The largest differences can be found in the mean-reversion speed parameters $\beta$ and $\eta$. Especially $\eta$ is much higher for Turkey than for Brazil.

More surprisingly are the results on the correlation factor parameters. A first thing that we note is that in all set-ups the values of almost all the factor parameters are very small. This suggests that the correlation factors only play a minor role and that the correlated intensities are almost entirely driven by the corresponding pure intensities. We expected this already for
the default intensity, since we conjectured that the liquidity of the CDSs should have little impact on the default intensity (small \(g_{ask}\) and \(g_{bid}\)), but we would expect a larger impact of the default intensity on the liquidity intensities, since the willingness to buy or sell a credit default swap should inevitably be related to the credit riskiness of the underlying reference entity. Furthermore, we observe that the signs of the correlation factor parameters are not always consistent for the different set-ups. Given that the factor parameters have such low values, we do not consider this a problem, since the effect of the correlation factors is negligible anyway, irrespective of the sign.

The low values of the correlation factor parameters may be explained by the fact that we separate the calibration of these parameters from the calibration of the (pure) intensities and the process parameters in our calibration procedure. This sequential approach is indeed not ideal, but is motivated by the fact that we want to calibrate the intensities in a cross-sectional way (the intensities on a day \(t\) should be calibrated on observations on day \(t\)), whereas the correlation factor parameters should be the constant over time and should, therefore, be calibrated across all observations. We note that Bühler and Trapp (2008) [11] and Badaoui, Cathcart and El-Jahel (2013) [5] also use this sequential approach. Their results on the correlation factors are debatable as well and, therefore, this step in the calibration procedure is open for improvement.

<table>
<thead>
<tr>
<th></th>
<th>Process Parameters</th>
<th>Correlation Factor Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brazil</td>
<td>Turkey</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.015</td>
<td>0.01633</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>(\sigma_{bid})</td>
<td>0.13</td>
<td>0.1233</td>
</tr>
<tr>
<td>(\sigma_{ask})</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>(\eta)</td>
<td>0.001</td>
<td>0.33</td>
</tr>
<tr>
<td>(f_{ask})</td>
<td>0.001</td>
<td>-0.0088</td>
</tr>
<tr>
<td>(f_{bid})</td>
<td>0.0010</td>
<td>0.0016</td>
</tr>
<tr>
<td>(g_{ask})</td>
<td>-0.0078</td>
<td>0.0001</td>
</tr>
<tr>
<td>(g_{bid})</td>
<td>-0.0083</td>
<td>0.0001</td>
</tr>
<tr>
<td>(\omega_{bid,ask})</td>
<td>0.00059</td>
<td>0.0016</td>
</tr>
<tr>
<td>(\omega_{ask,bid})</td>
<td>0.00799</td>
<td>0.0014</td>
</tr>
</tbody>
</table>

Table 7.6: Process and correlation parameter calibration results.
Chapter 7. Results

Table 7.7 gives the calibration results of the pure intensities. We see that for Brazil the results are very similar in both model set-ups. For Turkey, on the other hand, the results for, especially, the liquidity intensities are different in both model set-ups. The default intensities are, however, still roughly the same (although somewhat less than for Brazil) in both model set-ups. It is comforting that the default intensities are not that sensitive to the model specification, since the pure default intensities are, in principle, modeled the same in both model set-ups. This also means that, since the default probability is mainly driven by the pure default intensity, the model-implied default probabilities will be similar in both model set-ups. This is something we would like, since a difference in the modeling of the liquidity intensities should not have a large impact on the implied default probabilities. That the liquidity intensities behave more differently for Turkey than for Brazil in both model set-ups may be caused by the fact that the Turkish bid-ask premia are, on average, wider and more volatile (see table 6.2) and that, therefore, the calibration is more sensitive to the model set-up. Figures 7.2 and 7.3 give the plots of the (correlated) intensities for Brazil and Turkey, respectively. These plots indeed confirm that for Brazil the calibrated intensities are very similar in both model set-ups, whereas for Turkey the intensities are different in both model set-ups.

We see that the pure default intensities are on average higher for Turkey than for Brazil and this nicely coincides with the fact that the CDS premia for Turkey are on average higher than for Brazil. Of course, CDS premia are not pure measures of credit risk (as we will argue in the next section), but in principle the level of the premia should at least be able to reproduce a ranking in credit riskiness.

It is also clear from table 7.7 that the liquidity intensities are always positive, and, in the light of our discussion in chapter 5, this indicates that the liquidity risk premium is attributed to the protection sellers. Intuitively, this is also the most plausible, since market illiquidity gives more risk to protection sellers than protection buyers. Protection buyers know beforehand what the maximum amount of premium payments is they have to pay. Protection sellers, on the other hand, do not know beforehand if and how much they have to pay. If the default probability changes over time, the likelihood of having to make a payment changes and in an illiquid market it is then harder for the protection sellers to hedge their positions. They want to be compensated for this risk and they do this by charging higher premia.

Now, we somewhat artificially use driftless Arithmetic Brownian motions or OU processes with the same speed of mean reversion parameter to model the liquidity intensities, but, since we observe that the liquidity intensities are always positive, it would be nice to consider the model set-up in which we model these liquidity intensities as CIR processes. A model set-up with three CIR processes is, however, currently not feasible, since we then have nine process parameters (three for each CIR process) and, as we discussed in chapter 6, the calibration time grows exponentially in the number of process parameters. If, however, the computational speed of the calibration could be improved significantly, it would certainly be interesting to try this
model set-up, since it allows for more diversification between the bid and ask liquidities and it fulfills the non-negativity constraints.

Lastly, we show in figure 7.4 the differences between the (correlated) ask and bid liquidity intensities for Brazil in both model set-ups. This is done to show that these differences are indeed always positive, meaning that $\gamma_{\text{ask}} > \gamma_{\text{bid}}$. This is needed to assure that the bid premia are not larger than ask premia (and thereby preventing arbitrage opportunities in our model). We explicitly show this, since our model did not strictly impose this restriction beforehand. These results also hold for Turkey.

| Summary Statistics Pure Intensities | Arithmetic Brownian Motion Set-Up | | | Ornstein-Uhlenbeck Set-Up | | |
|---|---|---|---|---|---|---|---|---|---|
| | Brazil | Turkey | Brazil | Turkey | Brazil | Turkey |
| | $x_{\text{bid}}$ | $y_{\text{bid}}$ | $y_{\text{ask}}$ | $x_{\text{bid}}$ | $y_{\text{bid}}$ | $y_{\text{ask}}$ | $x_{\text{bid}}$ | $y_{\text{bid}}$ | $y_{\text{ask}}$ |
| Mean | 0.0065 | 0.4170 | 0.4460 | 0.0046 | 0.4813 | 0.5088 | 0.0081 | 0.7741 | 0.8131 |
| Std. Dev. | 0.0018 | 0.0337 | 0.0335 | 0.0039 | 0.0421 | 0.0421 | 0.0033 | 0.0912 | 0.0913 |
| Max | 0.012 | 0.5308 | 0.5577 | 0.0226 | 0.5630 | 0.5920 | 0.0122 | 0.9441 | 0.9784 |
| Min | 0.0030 | 0.3535 | 0.3875 | 0.0032 | 0.3888 | 0.4137 | 0.0020 | 0.5433 | 0.5753 |

Table 7.7: Summary statistics pure intensities.
Chapter 7. Results

Figure 7.2: Default and bid and ask liquidity intensities for Brazilian CDS.
7.1. Calibration Results

Figure 7.3: Default and bid and ask liquidity intensities for Turkish CDS.
Figure 7.4: Difference between ask and bid liquidity intensities for Brazil.
7.2 Decomposition of the CDS Mid Premium

One of the main goals of our research is to investigate which part of the CDS premium can be attributed to pure credit risk. In this section, we will give the results of the decomposition of the 2 year CDS mid premia into pure credit and liquidity parts by using the decomposition procedure that was explained in chapter 5.

We state the results for the 2 year CDSs only, since we think that the risks in the 2 year CDSs are calibrated best for two reasons: First of all, it is, in general, much harder to estimate risks for longer maturities. We, therefore, think that the market participants have more trouble with valuing risks on a 10 year horizon than on 2 year horizon and that, therefore, the 10 year CDS premia may be more prone to distorting factors. Secondly, since the payment dates of the 2 year CDSs overlap completely with the first payment dates of the 3, 5 and 10 year CDSs, we think that these payment dates are calibrated best in an economic way. With this we mean that, even though the pricing errors need not be the lowest for the 2 year CDS premia, we still think that the underlying default and liquidity components make the most sense for this maturity.

Table 7.8 gives the results of the decompositions. We note that in the Arithmetic Brownian motion set-up, we decompose the premium into a credit, a liquidity and a correlation components, whereas we only decompose the CDS premium into a credit and a liquidity part in the OU set-up. This is done, since we cannot split the liquidity part and the correlation part in the OU case, since this requires that we compute CDS premia under the assumption that the correlation factor parameters are zero, which, as we explained in chapter 5, induces a division through zero in some formulas. So, in the OU case, the liquidity part should be compared with the combined figure of the liquidity and correlation parts in the Arithmetic Brownian motion set-up. Since, in the Arithmetic Brownian motion set-up, the correlation part is on average negligible, this does not complicate the comparisons between the two set-ups.

In the case of Brazil, we see that the decomposition of the 2 year mid premia is very similar in both set-ups. Approximately 61.5% of the 2 year mid premium can be attributed to pure credit risk, while roughly 38.5% can be attributed to liquidity risk. We thus see that a substantial part of the CDS premium is not related to credit risk and this indeed confirms the conjecture that the CDS premium is not a pure measure of credit risk. In the case of Turkey, the decompositions in the different model set-ups differ slightly. This is not surprisingly, since we also found that the intensity processes in the different set-ups differed more than they did for Brazil. For Turkey, approximately 54% of the 2 year CDS mid premium can be attributed to credit risk and approximately 46% to liquidity risk. These results are in line with Badaoui, Cathcart and El-Jahel (2013) [5], who find that, on average, 56% of the sovereign CDS premia can be attributed to credit risk and 44% to liquidity risk. These results thus indeed confirm the conjecture that liquidity risk has a large impact on the CDS premia and that, therefore, these effects should
be filtered out if one wants to use CDS premia to extract implied default probabilities.

<table>
<thead>
<tr>
<th>Credit Part</th>
<th>Brazil</th>
<th>Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.6175</td>
<td>0.5783</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0249</td>
<td>0.0277</td>
</tr>
<tr>
<td>Max</td>
<td>0.6677</td>
<td>0.6398</td>
</tr>
<tr>
<td>Min</td>
<td>0.5494</td>
<td>0.5181</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liquidity Part</th>
<th>Brazil</th>
<th>Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.3824</td>
<td>0.4207</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0227</td>
<td>0.0272</td>
</tr>
<tr>
<td>Max</td>
<td>0.4417</td>
<td>0.4802</td>
</tr>
<tr>
<td>Min</td>
<td>0.3379</td>
<td>0.3602</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlation Part</th>
<th>Brazil</th>
<th>Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.0003</td>
<td>0.0008</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0047</td>
<td>0.0006</td>
</tr>
<tr>
<td>Max</td>
<td>0.0144</td>
<td>0.0026</td>
</tr>
<tr>
<td>Min</td>
<td>-0.0165</td>
<td>-0.0005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ornstein-Uhlenbeck set-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
</tr>
<tr>
<td>Credit Part</td>
</tr>
<tr>
<td>mean</td>
</tr>
<tr>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td>Min</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liquid Part</th>
<th>Brazil</th>
<th>Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.3844</td>
<td>0.5059</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0245</td>
<td>0.0416</td>
</tr>
<tr>
<td>Max</td>
<td>0.4480</td>
<td>0.5846</td>
</tr>
<tr>
<td>Min</td>
<td>0.3386</td>
<td>0.3933</td>
</tr>
</tbody>
</table>

Table 7.8: Decomposition of 2 year CDS premia. Entries are denoted as a fraction of the mid premium.

7.3 Chapter Summary

In this chapter, we showed the results of the calibration of our model on Turkish and Brazilian CDSs. We showed that the fit of our model-implied premia is sufficiently good for our purposes, with average relative pricing errors of roughly 2.2% for both countries in both model set-ups. The pricing errors can mainly be attributed to the fact that the grid-search algorithm does not allow us to search the whole process parameter space (only the discrete grid points).
7.3. Chapter Summary

Furthermore, since we made some simplifying assumptions on the pricing formulas, there is also some room for improvement by relaxing these assumptions. This, however, often comes at the cost of more (computationally) expensive formulas.

A remarkable result of the calibration is that all the correlation factors have very low values. This would suggest that the intensities are almost entirely driven by the corresponding pure intensities and that they are almost independent of each other. This is indeed what we expected for the default intensity, but, for the liquidity intensities, on the other hand, we expected that the default intensity would have a larger impact. A possible explanation for these results may be that we use a sequential calibration approach, in which the correlation factor parameters are calibrated separately from the intensities and process parameters. The reason we choose this sequential approach is that it allows us to calibrate the intensities at day $t$ to observed premia at day $t$ and the correlation factor parameters, which we assume to be constant over time, to all premia at once. We are, however, not completely satisfied with these results and, therefore, the calibration of the correlation factor parameters may be something that needs to be investigated further.

Regarding the pure intensities, we found that the results were very similar for Brazil in both model set-ups. For Turkey, however, there were some differences, especially in the liquidity intensities, in both model set-ups. In all cases, we found that the liquidity intensities were positive, which implies that the protection sellers get compensated for the liquidity risk in the market by means of higher premium payments.

In the last section of this chapter, we gave the results of the decomposition of the 2 year mid premia of Brazil and Turkey. For Brazil, we found that, on average, 61.5% of the 2 year mid premium can be attributed to credit risk. The other 38.5% can be attributed to liquidity risk. For Turkey we found that approximately 54% of the 2 year mid premium can be attributed to credit risk and the other 46% to liquidity risk. These results confirm that one needs to account for liquidity effects if one wants to use CDS implied default probabilities.
Maximum Likelihood Estimation of the Process Parameters

After calibrating the model to the data we obtained the time series of the pure $Q$-default and liquidity intensities. Furthermore, we obtained the parameters of the $Q$-dynamics of all three intensity processes. We would, however, like to determine the $P$-dynamics of the $Q$-default intensity, since, in combination with Rabobank’s model for the default event risk premium, these are sufficient to compute the real-world default probabilities (see formula (3.29)).

As explained in chapter 5, we can use the time series of the pure intensities to find their $P$-dynamics. Let us first recall that, if we assume that the density processes associated with the change of measure for the pure liquidity intensities and the pure $Q$-default intensity are of a specific affine form, then these intensities follow processes of the same form under $P$ as under $Q$, but with different drift parameters. Furthermore, the processes remain independent under $P$ as a consequence of (the multivariate) Lévy’s characterization theorem. Under these assumptions, we thus know what types of processes generate the historical time series for the pure intensities and we only have to find out what the values of the drift parameters are using these time series.

The estimation of (drift) parameters in continuous time diffusion-type models has received a great deal of attention in financial econometrics. In cases where transition densities are known, maximum likelihood (ML) methods may be preferred above other methods, since they have desirable asymptotic properties like consistency and asymptotic efficiency (Phillips & Yu, 2009) [51]. In the case of affine (multi)factor models, where the factors follow CIR, OU or Arithmetic Brownian motion processes, one knows the transition densities and, therefore, using the method of maximum likelihood for the estimation of the parameters is a logical choice (Aït-Sahalia & Kimmel, 2010) [2].
Chapter 8. Maximum Likelihood Estimation of the Process Parameters

An alternative to maximum likelihood estimation based on exact transition densities, is maximum likelihood estimation based on the Continuous Record likelihood. In this case the likelihood function is given by the Radon-Nikodym derivative associated with a change of measure between two absolutely continuous probability measures. Via Girsanov’s theorem, one can construct the Continuous Record likelihood explicitly.

The Continuous Record likelihood principle is based on a continuous observation of the process. Suppose we have a continuous diffusion process of the form

\[ dX_t = \mu(\theta, X_t)dt + \sigma(\theta, X_t)dW_t, \quad X_0 = x_0, \]  

(8.1)

where we assume for simplicity that \( \mu \) and \( \sigma \) are smooth functions and \( W \) is a Brownian motion under some measure \( P_0 \).

Suppose now that we have a probability measure \( P_1 \), such that \( P_0 \) is absolutely continuous with respect to \( P_1 \) (\( P_0 \ll P_1 \)). The Continuous likelihood process, \( L_T(\theta) \) associated with these two measures is given by

\[ L_t(\theta) = \frac{dP_0}{dP_1} \bigg| F_t, \quad t \geq 0, \]

where \( F_t \) is the canonical filtration of \( X_t \). The idea of Continuous Record maximum likelihood estimation is now to fix a measure \( P_1 \) such that the process \( X_t \) has no drift under \( P_1 \). Suppose now that we observe \( X \) on the interval \([0, T]\). Then, by Girsanov’s theorem, we know that the Continuous Record likelihood process at \( T \) is given by

\[ L_T(\theta) = \exp \left( \int_0^T \frac{\mu(\theta, X_t)}{\sigma^2(\theta, X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(\theta, X_t)} dt \right). \]

Equivalently, one can also evaluate the log-likelihood, which is given by

\[ \mathcal{L}_T(\theta) = \int_0^T \frac{\mu(\theta, X_t)}{\sigma^2(\theta, X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(\theta, X_t)} dt. \]  

(8.2)

Using the dynamics of \( X^{P_1} \), we can rewrite the previous equation as follows:

\[ \mathcal{L}_T(\theta) = \int_0^T \frac{\mu(\theta, X_t)}{\sigma^2(\theta, X_t)} \left( \sigma(\theta, X_t)dW_t^{P_1} \right) - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(\theta, X_t)} dt \]

\[ = \int_0^T \frac{\mu(\theta, X_t)}{\sigma(\theta, X_t)} dW_t^{P_1} - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(\theta, X_t)} dt. \]
Note that under the above measure change, we indeed have that \( dW_t^{\mathbb{P}_0} = dW_t^{\mathbb{P}_1} - \frac{\mu(\theta, X_t)}{\sigma(\theta, X_t)} \) \( dt \) and that therefore the process \( X_t \) has \( \mathbb{P}_0 \)-dynamics given by

\[
dX_t = \mu(\theta, X_t)dt + \sigma(\theta, X_t)dW_t^{\mathbb{P}_0}.
\]

We have to remark that in order to use the above method, the measure \( \mathbb{P}_1 \) should not depend on the unknown parameters, and, therefore, if the diffusion function depends on the unknown parameters \( \theta \), we cannot use the above method. Luckily, we can estimate the diffusion function separately by making use of the quadratic covariation of the process. The quadratic covariation of the process given in (8.1) is given by

\[
[X]_T = \int_0^T (dX_t)^2 = \int_0^T \sigma^2(\theta, X_t)dt.
\]

Since this function is continuously differentiable, we have that the quadratic variation of \( X \) provides a perfect estimate of the diffusion function and the parameters on which it depends. When a continuous path can be observed, we can therefore effectively assume that the diffusion function is known and does not depend on unknown parameters (Phillips & Yu, 2010) [51] (so we can write \( \sigma(\theta, X_t) = \sigma(X_t) \)). Note that, under the assumptions we made on the measure change in chapter 5, the diffusion function is indeed known, since the measure change induces only a drift change. We can, therefore, use the \( \mathbb{Q} \)-diffusion term that we obtain from the calibration procedure as the \( \mathbb{P} \)-diffusion term.

It can be shown that, in general, the Continuous Record likelihood method suffers from a finite sample bias. It can, however, also be shown that the method is consistent and asymptotically normal (Göing-Jaeschke, 1998) [30], provided we actually have a continuous observation period.

Often, one does not observe a diffusion process continuously, but only at discrete time points. In our situation, for example, we obtain time series of daily observations of the pure intensity processes. Therefore, we can only approximate the Continuous Record Likelihood function. In general, using a discrete-time approximation of a continuous-time estimation procedure introduces a bias. Florens-Zmirou (1989) [29] showed that the estimators of the discrete-time approximation of a Continuous Record log-likelihood with time-equidistant observations are even inconsistent.

We will, however, still investigate the performance of this type of maximum likelihood estimator in a discrete-time setting, since we want to see how this method compares to the method of exact maximum likelihood estimation. This is mainly done to gain intuition about the performance of the exact maximum likelihood estimators, which we will be using on our time series in order to extract the \( \mathbb{P} \)-dynamics of the intensities.
In this chapter, we will thus evaluate maximum likelihood estimators based on the exact transition densities and on the discrete-time approximation of the Continuous Record likelihood for the parameters of CIR, Ornstein-Uhlenbeck and Arithmetic Brownian motion processes. We will quickly evaluate their performance by means of a Monte Carlo study. We will end this chapter with applying the exact maximum likelihood estimators to the time series of the pure intensities, which were obtained in the calibration procedure, and we will compute the real-world default probabilities of Brazil and Turkey.

8.1 Maximum Likelihood Estimators for Arithmetic Brownian Motion

Let us consider the Arithmetic Brownian motion given by

\[ dy_t = \mu dt + \sigma dW_t. \]

We know that, conditional on \( F_s \), with \( s \leq t \), \( y_t \) is normally distributed with mean \( y_s + \mu (t - s) \) and variance \( \sigma^2 (t - s) \) (see Appendix D).

8.1.1 Exact Maximum Likelihood

Suppose that we have \( (N + 1) \) observations, \( y_i \), for \( i = 0, 1, \ldots, N \) with equidistant time-differences \( t_{i+1} - t_i = \Delta \) \( \forall i = 0, 1, \ldots, N \). We get that the conditional probability density function of an observation \( y_i \) given a previous observation \( y_{i-1} \) is given by

\[ f(y_i|y_{i-1}, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi \sigma^2 \Delta}} \exp \left( -\frac{(y_i - y_{i-1} - \mu \Delta)^2}{2\sigma^2 \Delta} \right). \]

The log-likelihood function of the complete sample is given by

\[ \mathcal{L}(\mu, \sigma^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2 \Delta) - \sum_{i=1}^{N} \frac{(y_i - y_{i-1} - \mu \Delta)^2}{2\sigma^2 \Delta}. \]

We find the maximum likelihood estimators by taking the partial derivatives and equating them to zero and then solving them for the parameters. We get the following maximum likelihood estimators (see Appendix D for the derivations):
8.2. Maximum Likelihood Estimators for Arithmetic Brownian Motion

\[ \hat{\mu} = \frac{1}{N\Delta} \sum_{i=1}^{N} (y_i - y_{i-1}) = \frac{y_N - y_0}{N\Delta}, \quad (8.3) \]

\[ \hat{\sigma}^2 = \frac{1}{N\Delta} \sum_{i=1}^{N} (y_i - y_{i-1})^2 - \frac{2}{N} \sum_{i=1}^{N} (y_i - y_{i-1}) + \hat{\mu}^2 \Delta. \quad (8.4) \]

8.1.2 Continuous Record Maximum Likelihood

The Continuous Record Likelihood method only works when the diffusion function is known and does not depend on the unknown parameters. Therefore, one can only estimate the drift parameters using this method. In our application this will not be a problem, since we obtain the diffusion parameters from the original model calibration.

So let us again consider the situation where we have \((N+1)\) observations, \(y_i\), for \(i = 0, 1, \ldots, N\) with equidistant time-differences \(t_{i+1} - t_i = \Delta \ \forall i = 0, 1, \ldots, N\). In the introduction of this section, we derived that the Continuous Record log-Likelihood function of a general continuous diffusion process (8.1) is given by

\[ \mathcal{L}_T(\theta) = \int_0^T \frac{\mu(\theta, X_t)}{\sigma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(X_t)} dt. \]

In the case of the Arithmetic Brownian motion, we have that \(\mu(\theta, y_t) = \mu\) and \(\sigma(y_t) = \sigma\), which is assumed to be known. The log-likelihood function therefore becomes

\[ \mathcal{L}_T(\mu) = \int_0^T \frac{\mu}{\sigma^2} dy_t - \frac{1}{2} \int_0^T \frac{\mu^2}{\sigma^2} dt. \]

It is shown in Appendix D that the discrete-time approximation of the Continuous Record maximum likelihood estimator of \(\mu\) is given by

\[ \hat{\mu} = \frac{1}{N\Delta} \sum_{i=1}^{N} (y_i - y_{i-1}) = \frac{y_N - y_0}{N\Delta}, \quad (8.5) \]

which is equal to the exact maximum likelihood estimator.
8.2 Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process

Consider the Ornstein-Uhlenbeck process given by

\[ dy_t = (\theta - \kappa y_t)dt + \sigma dW_t. \]

We know that, conditional on \( \mathcal{F}_s \), \( y_t \) is normally distributed with mean \( y_s e^{-\kappa(t-s)} + \theta \kappa (1 - e^{-\kappa(t-s)}) \) and variance \( \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-s)}\right) \) (see Appendix E).

8.2.1 Exact Maximum Likelihood

Suppose that we have \( (N+1) \) observations, \( y_i \), for \( i = 0, 1, \ldots, N \). We get that the conditional probability density function of an observation \( y_{i+1} \) given a previous observation \( y_i \) is given by

\[
f(y_{i+1}|y_i, \theta, \kappa, V) = \frac{1}{\sqrt{2\pi V^2}} \exp \left(-\frac{(y_{i+1} - y_i e^{-\kappa\Delta} - \frac{\theta}{\kappa} (1 - e^{-\kappa\Delta}))^2}{2V^2}\right),\]

where \( \Delta \) is the time difference between \( t_{i+1} \) and \( t_i \) (which we assume to be equal for all \( i \)), and \( V^2 = \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa\Delta}\right) \). Using the change of variables,

\[
\alpha = e^{-\kappa\Delta}, \\
\beta = \frac{\theta}{\kappa}, \\
V^2 = \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa\Delta}\right),
\]

we can rewrite the conditional probability density function as follows:

\[
f(y_{i+1}|y_i, \alpha, \beta, V) = \frac{1}{\sqrt{2\pi V^2}} \exp \left(-\frac{(y_{i+1} - y_i \alpha + \beta(1 - \alpha))^2}{2V^2}\right).
\]

The log-likelihood function is now given by
8.2. Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process

\[ L(\alpha, \beta, V) = \sum_{i=1}^{N} f(y_i | y_{i-1}, \alpha, \beta, V^2) \]

\[ = \frac{-N}{2} \log(2\pi) - N \log(V) - \frac{1}{2V^2} \sum_{i=1}^{N} (y_i - y_{i-1} - \beta(1 - \alpha))^2. \]

We find the maximum likelihood estimators by taking the partial derivatives and equating them to zero and then solving them for the parameters. We get the following maximum likelihood estimators for \( \alpha, \beta \) and \( V^2 \) (see Appendix E for the derivations):

\[ \hat{\beta} = \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_{i-1}}{N(1 - \hat{\alpha})}, \quad (8.6) \]

\[ \hat{\alpha} = \frac{N \sum_{i=1}^{N} y_i y_{i-1} - \sum_{i=1}^{N} y_i \sum_{i=1}^{N} y_{i-1}}{N \sum_{i=1}^{N} y_i^2 - \left( \sum_{i=1}^{N} y_{i-1} \right)^2}, \quad (8.7) \]

\[ \hat{V}^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{\alpha} y_{i-1} - \hat{\beta}(1 - \hat{\alpha}))^2. \quad (8.8) \]

The parameters \( \kappa, \theta \) and \( \sigma \) can be obtained by substituting back the above expressions.

8.2.2 Continuous Record Maximum Likelihood

In the case of the Ornstein-Uhlenbeck process, we have that, looking at the general Continuous Record log-likelihood function (8.2), \( \mu(\theta, y_t) = (\theta - \kappa y_t) \) and \( \sigma(y_t) = \sigma \), where we assume \( \sigma \) to be known. The Continuous Record log-likelihood function, therefore, becomes

\[ L_T(\theta, \kappa) = \int_{0}^{T} \frac{(\theta - \kappa y_t)}{\sigma^2} dy_t - \frac{1}{2} \int_{0}^{T} \frac{(\theta - \kappa y_t)^2}{\sigma^2} dt. \]

The ML estimates \( \hat{\theta} \) and \( \hat{\kappa} \) can be found by putting the partial derivatives w.r.t. \( \theta \) and \( \kappa \) equal to zero and solving them for \( \theta \) and \( \kappa \). Considering again the situation where we have \((N + 1)\) observations, \( y_i \), for \( i = 0, 1, \ldots, N \) with equidistant time-differences \( t_{i+1} - t_i = \Delta \forall i = 0, 1, \ldots, N \), we get the following discrete-time approximations of the Continuous Record maximum likelihood estimators of the drift parameters (see Appendix E for the derivations):
\begin{align*}
\hat{\theta} &= \frac{1}{N\Delta} \left( y_N - y_0 + \hat{\kappa} \Delta \sum_{i=1}^{N} y_{i-1} \right), \\
\hat{\kappa} &= \frac{1}{\Delta} \left( \frac{(y_N - y_0) \sum_{i=1}^{N} y_{i-1} - N \sum_{i=1}^{N} y_{i-1}(y_i - y_{i-1})}{N \sum_{i=1}^{N} y_{i-1}^2 - \left( \sum_{i=1}^{N} y_{i-1} \right)^2} \right). 
\end{align*}

\subsection{8.3 Maximum Likelihood Estimators for CIR Process}

In this subsection, we consider the CIR process given by

$$\text{d}x_t = (\alpha - \beta x_t)\text{d}t + \sigma \sqrt{x_t}\text{d}W_t.$$  

Let \( \theta = (\alpha, \beta, \sigma) \) denote the parameter vector. Cox et al. (1985) [16] observed that the above dynamics give rise to the following conditional transition densities:

$$p(x_{t+\Delta t} | x_t, \theta, \Delta t) = ce^{-u-v} \left( \frac{v}{u} \right)^\frac{\nu}{2} I_q(2\sqrt{uv}),$$  

where

\begin{align*}
c &= \frac{2\beta}{\sigma^2 \left( 1 - e^{-\beta \Delta t} \right)}, \\
u &= cx_te^{-\beta \Delta t}, \\
v &= cx_{t+\Delta t}, \\
q &= \frac{2\alpha}{\sigma^2} - 1,
\end{align*}

and \( I_q \) is the modified Bessel function of the first kind of order \( q \).

It can be shown (Cox et al, 1985) [16] that the distribution of \( x_{t+\Delta t} \) given \( x_t \) is given by

$$x_{t+\Delta t} | x_t \sim \eta \chi^2(\nu, \bar{\lambda}),$$

where \( \chi^2(\nu, \bar{\lambda}) \) is the distribution function of a non-central chi-square random variable with non-centrality parameter \( \bar{\lambda} \) and \( \nu \) degrees of freedom, where
8.3. Maximum Likelihood Estimators for CIR Process

\[ \bar{\lambda} = x_t \frac{4\beta e^{-\beta \Delta t}}{\sigma^2 (1 - e^{-\beta \Delta t})}, \]

\[ \nu = \frac{4\alpha}{\sigma^2}, \]

\[ \eta = \frac{\sigma^2 (1 - e^{-\beta \Delta t})}{4\beta}. \]

8.3.1 Exact Maximum Likelihood

Suppose that we have \( N \) observations \( x_{t_i}, i = 1, \ldots, N \), where we assume that the time steps are of equal length, so \( t_i - t_{i-1} = \Delta t \) for all \( i \).

The likelihood function of the \( N \) observations is given by

\[ L(\theta) = \prod_{i=1}^{N-1} p(x_{t_{i+1}}|x_{t_i}, \theta, \Delta t). \]

The log-likelihood function is given by

\[ \mathcal{L}(\theta) = \sum_{i=1}^{N-1} \ln(p(x_{t_{i+1}}|x_{t_i}, \theta, \Delta t)) . \]

Using the (conditional) transition density function of the CIR process as given by (8.11) gives

\[ \mathcal{L}(\theta) = (N - 1) \ln(c) + \sum_{i=1}^{N-1} \left( -u_i - v_{i+1} + \frac{q}{2} \ln(v_{i+1}) - \frac{q}{2} \ln(u_i) + \ln(I_q(2\sqrt{u_i v_{i+1}})) \right), \]

where \( u_i = cx_{t_i} e^{-\beta \Delta t} \) and \( v_{i+1} = cx_{t_{i+1}} \).

The maximum likelihood estimators \( \hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}) \) are now given by

\[ \hat{\theta} = \arg \max_\theta \mathcal{L}(\theta). \]

This optimization problem cannot be solved analytically, and therefore a numerical optimization procedure has to be used. We will use the MATLAB function \texttt{fminsearch} for this purpose. This function finds the (local) minimum of a scalar function of multiple variables and requires
as input the function to be optimized and initial values for the parameters to be optimized. The function to be optimized is in this case minus the log-likelihood function.

Since \texttt{fminsearch} is a search algorithm, it is important that the initial parameter values are chosen as good as possible. We will use an Ordinary Least Squares optimization procedure on the discretized version of the stochastic differential equation to find the initial parameter values. We get

\[
x_{t+\Delta t} = x_t + (\alpha - \beta x_t)\Delta t + \sigma \sqrt{x_t} \epsilon_t,
\]

where \(\epsilon_t \sim \mathcal{N}(0, \Delta t)\). Now we rewrite this to

\[
\frac{x_{t+\Delta t} - x_t}{\sqrt{x_t}} = \frac{\alpha \Delta t}{\sqrt{x_t}} - \beta \sqrt{x_t} \Delta t + \sigma \epsilon_t.
\]

Performing OLS on this equation gives the initial values for \(\alpha\) and \(\beta\). The initial value for \(\sigma^2\) will be given by taking the variance of the residual and dividing it by \(\Delta t\). The MATLAB implementation of the initial parameter value estimation is given in Listing 8.1.

\begin{verbatim}
%% initial parameter estimation using linear regression
s = x(1:end-1); % time series of CIR observations
dx = diff(x);
dx = transpose(dx./s.^0.5);
regressors = transpose([dt./s.^0.5; dt*s.^0.5]);
drift = inv(regressors'*regressors)*regressors'*dx; % OLS estimates of drift ... parameters (a,b)
residual = regressors*drift - dx;
initialA = drift(1);
initialB = -drift(2);
initialSigma = sqrt(var(residual,1)/dt);
initialParams = [initialA,initialB, initialSigma]; % vector of initial parameter
\end{verbatim}

The function that is minimized using \texttt{fminsearch} is minus the log-likelihood. In order to compute this function, the modified Bessel function of the first kind, \(I_q(2\sqrt{uv})\) must be calculated. This function approaches rapidly infinity and therefore it is not stable to use this function in an optimization procedure. MATLAB also has the command \texttt{besseli(q,2*sqrt(u*v),1)}, which is a scaled version of the modified Bessel function of the first kind. The value of \(\text{besseli}(q,2*\text{sqrt}(u*v),1)\) is \(I_q(2\sqrt{u}v)\exp(-2\sqrt{u}v)\) and is more stable (Kladívko, 2012) [42]. We adjust the log-likelihood for using this function and we get
8.3. Maximum Likelihood Estimators for CIR Process

\[L(\theta) = (N-1) \ln(c) + \sum_{i=1}^{N-1} \left(-u_i - v_{i+1} + \frac{q}{2} \ln(v_{i+1}) - \frac{q}{2} \ln(u_i) + \ln(I_0^1(2\sqrt{u_i v_{i+1}})) + 2\sqrt{u_i v_{i+1}}\right),\]

where \(I_0^1(2\sqrt{u_i v_{i+1}})\) is the scaled version of the modified Bessel function of the first kind. The MATLAB code for the objective function is given in Listing 8.2.

Listing 8.2: MATLAB code for the CIR objective function

```matlab
function logL = CIRobjective(params,dt, data)
    a = params(1);
    b = params(2);
    sigma = params(3);
    M = length(data);
    dataF = data(2:end);  % last (M-1) observations
    dataL = data(1:end-1);  % first (M-1) observations
    c = (2 * b) / (sigma^2 *(1-exp(-b*dt)));
    q = (2*a/sigma^2)-1;
    u = c * exp(-b*dt) * dataL;
    v = c * dataF;
    z = 2*sqrt(u.*v);
    bf = besseli(q,z,1);  % scaled modified bessel function of first kind
    logL = -(M-1)*log(c) + sum(u + v - 0.5*q*log(v./u) - log(bf) - z);
end
```

8.3.2 Continuous Record Maximum Likelihood

The Continuous Record log-likelihood function for the CIR process is given by

\[\mathcal{L}_T(\alpha, \beta) = \int_0^T \frac{(\alpha - \beta x_t)}{\sigma^2 x_t} dx_t - \frac{1}{2} \int_0^T \frac{(\alpha - \beta x_t)}{\sigma^2 x_t} dt,\]

where we assume that \(\sigma\) is known. Setting the partial derivatives equal to zero and solving them for the parameters gives the following discrete-time approximations of the Continuous Record maximum likelihood estimators (see Appendix F for the derivations):
\[ \hat{\alpha} = \frac{1}{\Delta \sum_{i=1}^{N} \frac{1}{x_{i-1}}} \left( \sum_{i=1}^{N} \frac{x_i}{x_{i-1}} - N + \hat{\beta} \Delta N \right), \quad (8.12) \]

\[ \hat{\beta} = \frac{1}{\Delta} \left( \frac{N \sum_{i=1}^{N} x_i - N^2 - (x_N - x_0) \sum_{i=1}^{N} \frac{1}{x_{i-1}}}{\sum_{i=1}^{N} x_{i-1} \sum_{i=1}^{N} \frac{1}{x_{i-1}} - N^2} \right). \quad (8.13) \]

### 8.4 Simulation Study

#### 8.4.1 Simulation of Arithmetic Brownian motion and CIR and OU Processes

In order to test the performance of the maximum likelihood estimators, we conduct a simulation experiment. We will simulate paths for the Arithmetic Brownian motion and CIR and the Ornstein-Uhlenbeck processes for some given parameter values and check whether the MLE’s are able to retrieve these parameter values.

Simulating diffusion processes can be done in different ways. One could use an Euler or higher-order approximation scheme to simulate the processes, but typically these methods induce some approximation errors. Since we know the exact (conditional) distributions of the Arithmetic Brownian motion, CIR process and OU process, we will use exact simulation procedures.

In the case of the Arithmetic Brownian motion, we know that \( y_t \) given \( y_s \) \((s < t)\), is normally distributed with mean \( y_s + \mu(t - s) \) and variance \( \sigma^2(t - s) \). Simulation is now straightforward by using a recursive algorithm. Figure 8.1 shows a typical simulated path. The MATLAB code used for simulating the Arithmetic Brownian motion is given by Listing 8.3.

Similarly, in the case of the Ornstein-Uhlenbeck process, we know that \( y_t \) given \( y_s \) \((s < t)\), is normally distributed with mean \( y_se^{-\kappa(t-s)} + \frac{\theta}{\kappa} \left(1 - e^{-\kappa(t-s)}\right) \) and variance \( \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-s)}\right) \). Again, using a recursive algorithm, simulation of a sample path is straightforward. Figure 8.2 shows a typical simulated path. The MATLAB code used for simulating the OU process is given by Listing 8.4.

In the case of the CIR process, we know that \( x_{t+\Delta} \) given \( x_t \), is distributed as follows:

\[ x_{t+\Delta t} | x_t \sim \eta \chi^2(\nu, \tilde{\lambda}), \]

where \( \chi^2(\nu, \tilde{\lambda}) \) is the distribution function of a non-central chi-square random variable with non-centrality parameter \( \tilde{\lambda} \) and \( \nu \) degrees of freedom, where
8.4. Simulation Study

\[ \hat{\lambda} = x_t \frac{4\beta e^{-\beta \Delta t}}{\sigma^2 (1 - e^{-\beta \Delta t})}, \]
\[ \nu = \frac{4\alpha}{\sigma^2}, \]
\[ \eta = \frac{\sigma^2 (1 - e^{-\beta \Delta t})}{4\beta}. \]

So if we are able to sample from a non-central chi-square distribution, we can simulate the CIR process.

Let us first consider the relation between central and non-central chi-square distributions. We have that the cumulative density function of a central chi-square random variable with \( \nu \) degrees of freedom:

\[ P(\chi^2(\nu, 0) \leq x) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} \int_0^x e^{-\frac{z}{2}} z^{\nu/2 - 1} dz. \]

The cumulative density function of a non-central chi-square distribution with \( \nu \) degrees of freedom and non-centrality parameter \( \hat{\lambda} \) is given by

\[ P(\chi^2(\nu, \hat{\lambda}) \leq x) = e^{-\frac{\hat{\lambda}^2}{2}} \sum_{j=0}^{\infty} \left( \frac{\hat{\lambda}^2}{4} \right)^j \frac{1}{j!} \int_0^x \frac{e^{-\frac{z}{2}}}{\Gamma(\frac{\nu}{2} + j)} \frac{z^{\nu/2 + j - 1}}{2^{\nu/2 + j}} dz. \]

Consider now a \( \chi^2(\nu + 2N, 0) \) random variable, where \( N \) has a Poisson distribution with rate \( \hat{\lambda}/2 \). We have that

\[ P(N = j) = e^{-\frac{\hat{\lambda}}{2}} \left( \frac{\hat{\lambda}/2}{j!} \right)^j, \quad j = 0, 1, \ldots, \]

and conditional on \( N = j \), we have

\[ P\left( \chi^2(\nu + 2N, 0) \leq x | N = j \right) = \frac{1}{2^{\nu/2 + j} \Gamma(\frac{\nu}{2} + j)} \int_0^x e^{-\frac{z}{2}} \frac{z^{\nu/2 + j - 1}}{2^{\nu/2 + j}} dz. \]

The unconditional distribution is now given by
which is exactly the cumulative density function of the non-central chi-square distribution with \(\nu\) degrees of freedom and non-centrality parameter \(\tilde{\lambda}\).

Now we know how to simulate \(x_{t+\Delta}|x_t\), and we can define an exact recursive simulation procedure for the CIR process:

1. Simulate a Poisson random variable \(N\) with intensity \(\tilde{\lambda}\), where \(\tilde{\lambda} = x_t \frac{4\beta e^{-\beta \Delta t}}{\sigma^2 (1 - e^{-\beta \Delta t})}\).
2. Given \(N\), simulate a central chi-square distributed random variable \(P\) with \(\nu+2N\) degrees of freedom.
3. Set \(x_{t+\Delta} = \eta P\) with \(\eta = \frac{\sigma^2 (1 - e^{-\beta \Delta t})}{4\beta}\).
4. Update \(\tilde{\lambda}\) for the next step (to obtain \(x_{t+2\Delta}\) the \(\tilde{\lambda}\) depends on \(x_{t+\Delta}\)) and go back to step 1.

Figure 8.3 shows a typical sample path of a CIR process. Listing 8.5 gives the MATLAB code for simulating a CIR process.
8.4. Simulation Study

Figure 8.1: Simulated path of Arithmetic Brownian motion with $\mu = \sigma = y_0 = 0.1$ and $\Delta = \frac{1}{250}$.
Listing 8.3: MATLAB code for exact simulation of an Arithmetic Brownian motion

```
% Exact simulation of Arithmetic Brownian motion

N = 1; % number of simulations

for i=1:N;
    %%% Initialize parameters
    mu = 0.1; % mean-reversion speed
    sigma = 0.1; % volatility
    T = 6; % time-horizon in years
    M = 1500; % number of time-steps
    dt = T / M; % time-step size
    y0 = 0.1; % starting value

    %%% Simulate the whole time-grid using exact distribution
    y = [y0];
    for j=1:M;
        mean1 = y(end)+mu*dt;
        variance = sigma^2 * dt;
        y(end+1) = normrnd(mean1,sqrt(variance));
    end
end
```
Figure 8.2: Simulated path of OU process with \( \theta = \kappa = \sigma = y_0 = 0.1 \) and \( \Delta = \frac{1}{250} \).
Listing 8.4: MATLAB code for exact simulation of a OU process

```matlab
% Exact simulation of Ornstein-Uhlenbeck process

N = 500; % number of simulation trials

for i=1:N;
    % Initialize parameters
    theta = 0.1; % mean-reversion level
    kappa = 0.1; % mean reversion speed
    sigma = 0.1; % volatility
    
    T = 6; % time-horizon in years
    M = 1500; % number of time-steps
    dt = T / M; % time-step size
    y0 = 0.1; % starting value

    % Simulate the whole time-grid using exact distribution
    y = [y0];
    for j=1:M;
        mean1 = y(end)*exp(-kappa*dt)+(theta/kappa)*(1-exp(-kappa*dt));
        variance = (sigma^2 / (2*kappa)) * (1-exp(-2*kappa*dt));
        y(end+1) = normrnd(mean1,sqrt(variance));
    end
end
```

Chapter 8. Maximum Likelihood Estimation of the Process Parameters
8.4. Simulation Study

Figure 8.3: Simulated path of CIR process with $\alpha = \beta = \sigma = x_0 = 0.1$ and $\Delta = \frac{1}{250}$.
Chapter 8. Maximum Likelihood Estimation of the Process Parameters

Listing 8.5: MATLAB code for exact simulation of a CIR process

```
1 %==========================================================================
2 % Exact simulation of CIR process
3 %==========================================================================
4
5 N = 20; % number of simulation trials
6
7 for i=1:N;
8
9     %% Initialize parameters
10     a = 0.1; % mean-reversion level
11     b = 0.1; % mean reversion speed
12     sigma = 0.1; % volatility
13
14     T = 6; % time-horizon in years
15     M = 1500; % number of time-steps
16     dt = T / M; % time-step size (dt = 1/250)
17     x0 = 0.1; % starting value
18
19     d = 4*a / sigmaˆ2; % degrees of freedom
20     c = (4 * b) / (sigmaˆ2 * (1-exp(-b*dt))); % weight
21     lambda = c * x0 * exp(-b*dt); % initial non-centrality parameter
22
23     %% Simulate the whole time-grid
24
25     x = [x0]; % initial value
26     for j=1:M;
27         P = poissrnd(lambda/2); % generate poisson R.V. with parameter lambda/2
28         R = chi2rnd(d+2*P); % chi-square R.V. with d+2P degrees of freedom
29         x(end+1) = (1/c)*R; % append values to list of observations
30         lambda = c * x(end) * exp(-b*dt); % update lambda value
31     end
32 end
```
8.4. Simulation Study

8.4.2 Results Simulation Study

In order to investigate the performance of the different estimators, we conducted a simulation study. For each test, we simulated 500 sample paths of the relevant process and calculated the mean value of the estimates, their standard errors and the resulting root mean squared errors (RMSE). Each sample path consists of 6 years of daily observations (250 business days per year), which is roughly equal to the amount of observations we will obtain from our calibration. We will first give the results for the Arithmetic Brownian motion and after that the results for Ornstein-Uhlenbeck process and the CIR process.

Results for the Arithmetic Brownian Motion

To test the maximum likelihood estimators for the Arithmetic Brownian motion, we simulated the process

\[ dy_t = \mu dt + \sigma dW_t, \]

with the following parameter values:

\[ \mu = 0.1, \quad \sigma = 0.1, \quad y_0 = 0.1, \]

using the exact simulation procedure as given in Listing 8.3.

Since the exact maximum likelihood estimator of \( \mu \) is equal to the Continuous Record maximum likelihood estimator of \( \mu \), we only display the results for the exact maximum likelihood estimators. Table 8.1 shows the results of the simulation study. The (exact) maximum likelihood seems to perform well as the average estimates are very close to the real values and the standard errors are not that high.

<table>
<thead>
<tr>
<th>Maximum Likelihood Results Arithmetic Brownian Motion</th>
<th>True value ( \mu = 0.1 )</th>
<th>True value ( \sigma = 0.1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Exact</td>
<td>Exact</td>
</tr>
<tr>
<td>Mean</td>
<td>0.1013</td>
<td>0.0999</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0391</td>
<td>0.0020</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0359</td>
<td>0.0101</td>
</tr>
</tbody>
</table>

Table 8.1: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of an Arithmetic Brownian motion with parameters \( \mu = \sigma = y_0 = 0.1 \) are simulated.
Chapter 8. Maximum Likelihood Estimation of the Process Parameters

Results for the Ornstein-Uhlenbeck Process

To test the maximum likelihood estimators for the OU process, we simulated the process

$$d y_t = (\theta - \kappa y_t)dt + \sigma dW_t,$$

with the following parameter values:

$$\theta = 0.1, \quad \kappa = 0.1, \quad \sigma = 0.1, \quad y_0 = 0.1^1,$$

using the exact simulation procedure as given in Listing 8.4.

Table 8.2 shows the results for the maximum likelihood estimators based on the exact maximum likelihood (columns Exact) and the discrete-time approximations of the Continuous Record maximum likelihood (columns CR). It is clear from the table that the estimators behave poorly, with large biases and relatively large standard errors and root mean squared errors. For the discrete-time approximation of the Continuous Record likelihood method this is not very surprising, since we know that the discretization introduces a bias on top of the bias the method possesses in general. The exact maximum likelihood estimator suffers from a finite sample bias.

<table>
<thead>
<tr>
<th>Method</th>
<th>True value $\theta = 0.1$</th>
<th>True value $\kappa = 0.1$</th>
<th>True value $\sigma = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact CR</td>
<td>Exact CR</td>
<td>Exact CR</td>
</tr>
<tr>
<td>Mean</td>
<td>0.2263 0.2259</td>
<td>0.5072 0.5061</td>
<td>0.1 0.1</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1436 0.1431</td>
<td>0.5288 0.5266</td>
<td>0.0018 -</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.3553 0.3548</td>
<td>0.6381 0.6373</td>
<td>0.0032 -</td>
</tr>
</tbody>
</table>

Table 8.2: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of an OU process with parameters $\theta = \kappa = \sigma = y_0 = 0.1$ are simulated.

To make the maximum likelihood estimators useful, we have to find a way to reduce the biases. Phillips and Yu (2005) [50] propose to use the bias reduction technique called jackknife’s estimation.

The jackknife estimator of a parameter $\theta$ uses estimates from both the whole sample as well as subsamples and combines them in order to reduce the bias. Suppose that we have $N$

---

$^1$The performance of the maximum likelihood estimators is sensitive to the choice of these parameters. In general, we observed that the estimates were more precise for larger parameter values (not close to zero) and when $\kappa$ is larger than $\theta$. 

128
observations in a sample. Now we decompose these \( N \) observations in \( m \) consecutive subsamples of equal length \( l \), that is \( N = m \times l \). The jackknife estimator is then given by

\[
\hat{\theta}_{\text{jack}} = \frac{m}{m-1} \hat{\theta}_N - \frac{1}{m^2 - m} \sum_{i=1}^{m} \hat{\theta}_i,
\]

where \( \hat{\theta}_N \) is the estimated value of \( \theta \) using the whole sample and \( \hat{\theta}_i \) is the estimated value using subsample \( i \). The estimates of the whole sample and the subsamples can be based on exact maximum likelihood estimation and approximated Continuous Record maximum likelihood estimation. Tables 8.3 and 8.4 show results for the jackknife estimation procedure using \( m = 2 \) and \( m = 3 \) subsamples respectively.

### Table 8.3: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of an OU process with parameters \( \theta = \kappa = \sigma = y_0 = 0.1 \) are simulated. Results are given for a jackknife estimation procedure with \( m = 2 \) subsamples.

<table>
<thead>
<tr>
<th>Method</th>
<th>True value ( \theta = 0.1 )</th>
<th>True value ( \kappa = 0.1 )</th>
<th>True value ( \sigma = 0.1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact ((m = 2))</td>
<td>CR ((m = 2))</td>
<td>Exact ((m = 2))</td>
</tr>
<tr>
<td>Mean</td>
<td>0.1038</td>
<td>0.1020</td>
<td>0.1038</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0233</td>
<td>0.0224</td>
<td>0.0230</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0590</td>
<td>0.0415</td>
<td>0.0616</td>
</tr>
</tbody>
</table>

### Table 8.4: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of an OU process with parameters \( \theta = \kappa = \sigma = y_0 = 0.1 \) are simulated. Results are given for a jackknife estimation procedure with \( m = 3 \) subsamples.

<table>
<thead>
<tr>
<th>Method</th>
<th>True value ( \theta = 0.1 )</th>
<th>True value ( \kappa = 0.1 )</th>
<th>True value ( \sigma = 0.1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact ((m = 3))</td>
<td>CR ((m = 3))</td>
<td>Exact ((m = 3))</td>
</tr>
<tr>
<td>Mean</td>
<td>0.1020</td>
<td>0.1003</td>
<td>0.1019</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0216</td>
<td>0.0209</td>
<td>0.0213</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0450</td>
<td>0.0182</td>
<td>0.0434</td>
</tr>
</tbody>
</table>

Clearly, the jackknife estimation procedure improves the performance of both types of maximum likelihood estimators. For each parameter the biases are reduced and the estimators have...
smaller standard errors and root mean squared errors. The approximated Continuous Record maximum likelihood estimator seems to perform slightly better than the exact maximum likelihood estimator, but this might caused by the fact that for the approximated Continuous Record maximum likelihood estimator we use the true value of $\sigma$, since this method assumes the diffusion function to be known.

**Results for the CIR Process**

To test the maximum likelihood estimators for the CIR process, we simulated the process

\[ dx_t = (\alpha - \beta x_t)dt + \sigma \sqrt{x_t}dW_t, \]

with the following parameter values:

\[ \alpha = 0.1, \quad \beta = 0.1, \quad \sigma = 0.1, \quad x_0 = 0.1^2, \]

using the exact simulation procedure as given in Listing 8.5.

Table 8.5 gives the results for the exact maximum likelihood method and the approximated Continuous Record maximum likelihood method. Again, the estimators suffer from finite sample biases, although the results seem to be better than in the case of the Ornstein-Uhlenbeck process. Especially the estimate of $\beta$ suffers from a large bias and therefore we want to improve the performance of the estimators by using the jackknife estimation procedure.

<table>
<thead>
<tr>
<th>Method</th>
<th>True value $\alpha = 0.1$</th>
<th>True value $\beta = 0.1$</th>
<th>True value $\sigma = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Exact 0.1275  CR 0.1274</td>
<td>Exact 0.1888  CR 0.1887</td>
<td>Exact 0.0998  CR 0.1</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0582  CR 0.0581</td>
<td>0.2087  CR 0.2085</td>
<td>0.0018  -</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1658  CR 0.1656</td>
<td>0.2979  CR 0.2978</td>
<td>0  -</td>
</tr>
</tbody>
</table>

Table 8.5: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of a CIR process with parameters $\alpha = \beta = \sigma = x_0 = 0.1$ are simulated.

\(^2\)Again, the performance of the estimators depends on the choice of these parameters. We observed that, in general, the performance of the estimators is better when $\beta$ is relatively large compared to $\alpha$. We have chosen to display the results for the current parameter values in order to give a warning for the use of bias reduction techniques.
Tables 8.6 and 8.7 give the results for the jackknife estimation procedure using \( m = 2 \) and \( m = 10 \) subsamples respectively. We see that for \( m = 2 \) the performance of both estimators is even worse than without using the jackknife procedure. For \( m = 10 \) on the other hand, the exact maximum likelihood estimator performs much better using the jackknife estimator. The approximated Continuous Record maximum likelihood estimator, on the other hand, also performs bad in this case. Also for other values of \( m \) (which we do not report here), we see that the exact maximum likelihood estimator outperforms the approximated Continuous Record maximum likelihood estimator.

<table>
<thead>
<tr>
<th>Method</th>
<th>True value ( \alpha = 0.1 )</th>
<th>True value ( \beta = 0.1 )</th>
<th>True value ( \sigma = 0.1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Exact ((m = 2))</td>
<td>Exact ((m = 2))</td>
<td>Exact ((m = 2))</td>
</tr>
<tr>
<td></td>
<td>0.0292</td>
<td>-0.0559</td>
<td>-0.0918</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.2212</td>
<td>0.2032</td>
<td>0.5327</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.2661</td>
<td>0.3949</td>
<td>0.4380</td>
</tr>
</tbody>
</table>

Table 8.6: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of a CIR process with parameters \( \alpha = \beta = \sigma = x_0 = 0.1 \) are simulated. Results are given for a jackknife estimation procedure with \( m = 2 \) subsamples.

<table>
<thead>
<tr>
<th>Method</th>
<th>True value ( \alpha = 0.1 )</th>
<th>True value ( \beta = 0.1 )</th>
<th>True value ( \sigma = 0.1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Exact ((m = 10))</td>
<td>Exact ((m = 10))</td>
<td>Exact ((m = 10))</td>
</tr>
<tr>
<td></td>
<td>0.0939</td>
<td>0.1136</td>
<td>0.1097</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0749</td>
<td>0.0875</td>
<td>0.2408</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.07</td>
<td>0.475</td>
<td>0.116</td>
</tr>
</tbody>
</table>

Table 8.7: MLE results of a simulation study of 500 sample runs, where in each sample run 6 years of daily observations of a CIR process with parameters \( \alpha = \beta = \sigma = x_0 = 0.1 \) are simulated. Results are given for a jackknife estimation procedure with \( m = 10 \) subsamples.

### 8.4.3 Conclusions Simulation Study

The exact maximum likelihood method works very well for the Arithmetic Brownian motion. The parameter estimates are very close to the real parameter values and the standard errors
are low. We observed this for many different parameter values. For this process, there is, therefore, no need to use bias reduction techniques.

For both the Ornstein-Uhlenbeck process as the CIR process the exact maximum likelihood estimators and the approximated Continuous Record maximum likelihood estimators suffer from finite sample bias and have relatively high standard errors. Without bias reduction techniques, performance of both estimators is roughly the same for both the Ornstein-Uhlenbeck as the CIR processes. Using the jackknife estimation procedure, the performance of the maximum likelihood estimators mostly became better (provided the right number of subsamples were chosen), except for the approximated Continuous Record maximum likelihood estimator of the CIR process.

In case of the CIR process we, therefore, conclude that the exact maximum likelihood estimator is to be preferred over the approximated Continuous Record maximum likelihood estimator. A note of caution is that we see from Tables 8.6 and 8.7 that the jackknife estimation procedure can increase performance of an estimator, but can also decrease the performance. In the case of the CIR process, using 2 subsamples made the performance of both the exact maximum likelihood and the approximated Continuous Record maximum likelihood worse compared to the pure estimators. Therefore, it is a priori not clear whether the use of a jackknife estimation procedure improves or worsen the estimation errors and by how much.

In the case of the OU process, the exact maximum likelihood estimator and the approximated Continuous Record maximum likelihood estimator have similar performance. The jackknife bias reduction technique seems to work very well on this process and, therefore, we would always try to use this when estimating parameter values of an OU process. Because of its theoretical advantages (efficiency, consistency), we again prefer the exact maximum likelihood estimators over the approximated Continuous Record estimators.

### 8.5 Maximum Likelihood Results on Intensity Time Series

Having derived and evaluated the different maximum likelihood estimators, we are now able to use them in our model. Let us first consider the set-up where we model the pure liquidity intensities as driftless Arithmetic Brownian motions. We have the following $\mathbb{Q}$-dynamics:

\[
\begin{align*}
\text{d}x^Q(t) & = (\alpha - \beta x^Q(t))dt + \sigma \sqrt{x^Q(t)}dW^Q_x(t), \\
\text{d}y^{\text{bid}}(t) & = \sigma^{\text{bid}}dW^Q_{y^{\text{bid}}}(t), \\
\text{d}y^{\text{ask}}(t) & = \sigma^{\text{ask}}dW^Q_{y^{\text{ask}}}(t).
\end{align*}
\]
8.5. Maximum Likelihood Results on Intensity Time Series

The values of the above parameters can be obtained by the calibration procedure that was proposed in chapter 6 and the results for the Brazilian and Turkish CDSs were given in table 7.6. Under certain assumptions on the change of measure (see section 5.3.2), we have that the pure intensity processes are of the same type under $\mathbb{P}$ as under $\mathbb{Q}$, but with different drift parameters. Making these assumptions, we thus get the following $\mathbb{P}$-dynamics of the pure intensities:

$$
\begin{align*}
\text{d}x^Q(t) &= (\tilde{\alpha} - \tilde{\beta}x^Q(t))dt + \sigma \sqrt{x^Q(t)} dW^P_x(t), \\
\text{d}y^{\text{bid}}(t) &= \mu^{\text{bid}}dt + \sigma^{\text{bid}} dW^P_{y^{\text{bid}}}(t), \\
\text{d}y^{\text{ask}}(t) &= \mu^{\text{ask}}dt + \sigma^{\text{ask}} dW^P_{y^{\text{ask}}}(t).
\end{align*}
$$

Now, $\tilde{\alpha}$, $\tilde{\beta}$, $\mu^{\text{bid}}$ and $\mu^{\text{ask}}$ are unknown and need to be estimated. The volatility parameters $\sigma$, $\sigma^{\text{bid}}$ and $\sigma^{\text{ask}}$ are still the same as under $\mathbb{Q}$ and were already obtained by our calibration procedure. As a consequence of (the multivariate) Lévy’s characterization theorem, these processes remain independent under $\mathbb{P}$ and, therefore, we can analyze these processes separately.

In the light of this chapter, we can use the (discrete) time series of the pure intensities, that were obtained in our calibration procedure, to estimate the unknown drift parameters by using maximum likelihood estimators. These (historical) time series give us information on the $\mathbb{P}$-dynamics of these processes (see for example Brigo and Mercurio (2006) [9] for an explanation of this in the case of short-rate interest rate models) and, therefore, using the maximum likelihood estimators as discussed above, will give us exactly the estimates of the unknown $\mathbb{P}$-drift parameters.

We will thus estimate $\mu^{\text{bid}}$ and $\mu^{\text{ask}}$ by using formula (8.3) and the time series of $y^{\text{bid}}$ and $y^{\text{ask}}$, respectively. Furthermore, we will estimate $\tilde{\alpha}$ and $\tilde{\beta}$ by using the time series of $x^Q$ and the optimization procedure corresponding to the exact maximum likelihood estimators as explained in the section above. We will not use bias reduction techniques in the estimation procedure of the $\mathbb{P}$-dynamics of the pure $\mathbb{Q}$-default intensity process, since, as we have shown above, the use of such techniques on a CIR process can have an adverse impact on the performance of the estimators.

In a similar way, we will consider the following $\mathbb{P}$-dynamics of the pure intensities in the Ornstein-Uhlenbeck set-up:

$$
\begin{align*}
\text{d}x^Q(t) &= (\alpha - \beta x^Q(t))dt + \sigma \sqrt{x^Q(t)} dW^P_x(t) \\
\text{d}y^{\text{bid}}(t) &= (\theta^{\text{bid}} - \kappa^{\text{bid}} y^{\text{bid}}(t))dt + \sigma^{\text{bid}} dW^P_{y^{\text{bid}}}(t), \\
\text{d}y^{\text{ask}}(t) &= (\theta^{\text{ask}} - \kappa^{\text{ask}} y^{\text{ask}}(t))dt + \sigma^{\text{ask}} dW^P_{y^{\text{ask}}}(t),
\end{align*}
$$

133
where $\theta^{bid}, \theta^{ask}, \kappa^{bid}, \kappa^{ask}, \bar{\alpha}$ and $\bar{\beta}$ are the unknown parameters. The example that we worked out in our simulation study suggested that the Ornstein-Uhlenbeck process responds well to the jackknife bias reduction technique. In estimating the liquidity intensity drift parameters, however, we obtained very different results for different numbers of subsamples and, therefore, we think it is better not to use this technique. Just as in the CIR case, the wrong number of subsamples may worsen the outcomes and since there is no way of telling what number of subsamples is correct, we think it is better to use the full sample estimates for $\theta^{bid}, \theta^{ask}, \kappa^{bid}, \kappa^{ask}, \tilde{\alpha}$ and $\tilde{\beta}$.

Table 8.8 shows the estimation results for both model set-ups for both Brazil and Turkey.

<table>
<thead>
<tr>
<th></th>
<th>Brazil</th>
<th></th>
<th>Turkey</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arithmetic BM</td>
<td>OU</td>
<td>Arithmetic BM</td>
<td>OU</td>
</tr>
<tr>
<td>$\bar{\alpha}$</td>
<td>0.0258</td>
<td>0.0256</td>
<td>$\bar{\alpha}$</td>
<td>0.0255</td>
</tr>
<tr>
<td>$\bar{\beta}$</td>
<td>4.0099</td>
<td>3.9722</td>
<td>$\bar{\beta}$</td>
<td>2.5701</td>
</tr>
<tr>
<td>$\mu^{bid}$</td>
<td>0.0211</td>
<td>0.4871</td>
<td>$\mu^{bid}$</td>
<td>0.0194</td>
</tr>
<tr>
<td>$\mu^{ask}$</td>
<td>0.0198</td>
<td>1.1541</td>
<td>$\mu^{ask}$</td>
<td>0.01644</td>
</tr>
<tr>
<td>$\kappa^{bid}$</td>
<td>0.4170</td>
<td>0.8972</td>
<td>$\kappa^{bid}$</td>
<td>0.3244</td>
</tr>
<tr>
<td>$\kappa^{ask}$</td>
<td>0.8972</td>
<td>1.1541</td>
<td>$\kappa^{ask}$</td>
<td>3.2110</td>
</tr>
</tbody>
</table>

Table 8.8: Estimated values of drift parameters of the $\mathbb{P}$-dynamics of the pure intensities.

In order to get some feeling for the correctness of these estimates, let us consider the stationary distributions of an Ornstein-Uhlenbeck process and a CIR process. We know that the stationary mean of an Ornstein-Uhlenbeck process is given by $\theta^{bid}\kappa^{bid}$. In the case of Brazil, we thus see that the stationary mean of the OU process with parameter values as described in table 8.8, is equal to $0.0256 \times 0.4871 \approx 0.422$ for the bid liquidity intensity. We note that this is close to the actual observed mean of the bid liquidity intensity of 0.4033 (see table 7.7). Similarly, for the ask liquidity intensity, the stationary mean of the estimated process is given by $0.4170 \times 0.8972 \approx 0.46$, which is again close to the actual observed mean of 0.4428. For the CIR process, the stationary mean is given by $\frac{\bar{\alpha}}{\bar{\beta}}$, which, in the case of the OU set-up for Brazil, is equal to $\frac{0.0256}{3.9722} \approx 0.006445$, and is again close to the observed mean of 0.0065 (see table 7.7). Similar results hold for Turkey. Of course, this is no hard evidence that the estimates are correct, but it indicates that they do make sense.

Using the above estimates, we can now compute the default probabilities of Turkey and Brazil. Let us recall that the probability to survive until time $T$, given that no default has occurred until time $t$, is given by

$$
\mathbb{P}\left\{ \tau > T | \mathcal{F}_t \right\} = \mathbb{E}\mathbb{P}\left[ e^{- \int_t^T A^\mathbb{P}(s)ds} | \mathcal{F}_t \right] = \mathbb{E}\mathbb{P}\left[ e^{- \int_t^T K \Lambda^\mathbb{Q}(s)ds} | \mathcal{F}_t \right],
$$

(8.14)
Maximum Likelihood Results on Intensity Time Series

where $K$ is a country-specific constant that represents the risk premium on the default event. In our model set-up, we can rewrite the last expression as follows:

$$
E_P \left[ e^{-\int_t^T K\lambda_Q(s)ds} \bigg| \mathcal{F}_t \right] = \mathbb{E}^P \left[ e^{-\int_t^T K(x_Q(s)+g_{bid}y_{bid}(s)+g_{ask}y_{ask}(s))ds} \bigg| \mathcal{F}_t \right]
$$

where the last step follows from lemma 5.1 and the fact that $x_Q$, $y_{bid}$ and $y_{ask}$ are driven by independent Brownian motions under $P$. We again recognize the product of three bond price formulas and, since the processes $x_Q$, $y_{bid}$ and $y_{ask}$ are still affine (and of the same type) under $P$, we can compute these (and thus the default probability) explicitly.

In the confidential version, the remainder of this section consists of a discussion on how our PD estimates compare to Rabobank's ratings. Due to the confidential nature of Rabobank's internal estimates, we cannot publish the complete discussion. The discussion of some results is, therefore, kept superficial.

Using the estimated values of $K$, the estimated parameters of the $P$-dynamics and the pure intensity values on the days that correspond to the days on which Rabobank estimated the default probabilities gives the following results:

<table>
<thead>
<tr>
<th>Estimated One-Year Default Probabilities Brazil and Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
</tr>
<tr>
<td>Arithmetic BM</td>
</tr>
<tr>
<td>26.212</td>
</tr>
</tbody>
</table>

Table 8.9: Estimated one-year default probabilities Brazil and Turkey (in basis points).

Table 8.9 shows that our estimates of the one-year default probability of Turkey are similar to the estimates of Rabobank.

In our model, we have a default intensity that changes daily and therefore the default probability estimate changes daily. Of course, the other parameters in the model are estimated using historical data and therefore our model is a combination of a so-called through-the-cycle and point-in-time model. The first type refers to models that use a history of data to calibrate parameters and that are, therefore, less sensitive to daily changes. The second type refers to models that look at time-specific points. Rabobank has a slightly different rating philosophy.

To compensate for the difference between our hybrid model and Rabobank’ model, we look at the average rating over the sample period that Rabobank gives and we compare this with
the default probability estimate that we obtain by taking the average intensities instead of the intensities of a specific date. The average rating is obtained by assigning a value to each rating and by weighting it with the time period that this rating holds during the sample period. The model-implied default probabilities using the average intensities during this period are given in table 8.10.

<table>
<thead>
<tr>
<th>Arithemetic BM</th>
<th>OU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>27.84</td>
</tr>
<tr>
<td>Turkey</td>
<td>56.44</td>
</tr>
</tbody>
</table>

Table 8.10: Estimated one-year default probabilities Brazil and Turkey (in basis points) using average intensities.

We can see that, by using the average intensities, the estimates are similar to the average rating Rabobank gave to these sovereigns. We, furthermore, see from table 8.10 that we find a higher default probability for Turkey than for Brazil. This is in line with what we would expect from a quick look at the levels of the CDS premia. As can be seen from table 6.2, the average mid premia of Turkish CDSs are higher than those of Brazil, and, ignoring liquidity effects, a higher CDS premium should, in principle, reflect a higher credit risk. This is again no proof that our model is correct, but it is an indication that our model gives plausible results, since it satisfies the ranking property.

All-in-all, we conclude that our estimates can be viewed as a benchmark in line with Rabobank’s estimates (and vice versa). A small note that we make to the estimates that we obtained, is that, as we have shown in this chapter, the maximum likelihood estimates of our parameters may be biased and that, therefore, our estimates of the default probabilities may be biased. As argued above, the use of the jackknife bias reduction technique does not always prevents us from obtaining biased parameter estimates, since the wrong number of subsamples may actually decrease the performance of the maximum likelihood estimators. We did, however, argue that our estimation results make sense, since the estimated stationary means coincided with the actual observed means of the processes. This, however, is no concrete proof of the correctness of the estimates and, therefore, the use of suitable bias reduction techniques for maximum likelihood estimators for discrete-time observed processes may be something that needs to be investigated further.

### 8.6 Chapter Summary

In this chapter, we discussed the topic of maximum likelihood estimation of the drift parameters of a diffusion process based on discrete-time observations. We specifically focused on the
8.6. Chapter Summary

parameter estimation of an Arithmetic Brownian motion, an Ornstein-Uhlenbeck process and a CIR process. For all these processes, we derived maximum likelihood estimators based on their exact (conditional) distributions and on a discrete-time approximation of the Continuous Record log-likelihood function.

We evaluated the performance of the different estimators for the different processes by means of a Monte Carlo simulation study. We saw that in the Arithmetic Brownian motion case, both the exact and the (discrete-time approximation of the) Continuous Record maximum likelihood estimators performed very well. The estimates were very close to the parameter values that were used to simulate the sample paths. For the OU and CIR processes, however, the performance of the estimators depended on the choice of parameters with which we simulated the data. We gave some results for parameter choices for which the estimators did not perform very well. In order to increase the performance of the estimators, we introduced the jackknife bias reduction technique. At first sight, the maximum likelihood estimators of the OU process seemed to improve using this technique. The performance of the estimators of the CIR parameters, however, did not always increase by the jackknife technique and sometimes even decreased. We, therefore, concluded that the use of the jackknife technique cannot guarantee improvement and that we should be careful when using it. But, given the possible existence of bias in the parameter estimates, we do think that the topic of bias reduction techniques should be investigated further.

We ended this chapter with applying the maximum likelihood estimators on the time series of the intensity processes that we obtained from our calibration in the previous chapter, since, under certain assumptions on the change of measure, we know that these time series are generated by either CIR, OU or Arithmetic Brownian motion processes. By using these historic time series of the \( Q \)-intensities, we are able to extract the \( P \)-dynamics of the \( Q \)-default intensity, which, combined with Rabobank’s model of the default event risk premium, allows us to compute the real-world default probabilities.

Using the estimated parameters of the \( P \)-dynamics of the \( Q \)-default intensity, we found the one-year default probabilities of Turkey and Brazil to be approximately 0.4% and 0.25%, respectively. These estimates can be viewed as a benchmark in line with Rabobank’s estimates of the default probabilities. We, furthermore, concluded that our model satisfies the ‘ranking’ property of assigning higher default probabilities to countries that are also perceived more risky.
9

Conclusions

In this thesis, we investigated the possibility of extracting default probabilities from sovereign CDS premia in order to benchmark the estimates of the sovereign default probabilities that Rabobank currently uses. We argued that the large observed bid-ask spreads in the sovereign CDS market indicate that CDS premia are probably not pure measures of credit risk, but also of liquidity risk. The goal of our research was, therefore, to quantify the effects of liquidity risk on the sovereign CDS premia and to obtain ‘uncontaminated’ estimates of CDS-implied default probabilities.

We introduced an intensity-based model that incorporates liquidity effects by means of extra liquidity discount factors. We were able to derive closed-form formulas for the CDS bid and ask premia, which allow for a natural decomposition of the corresponding mid premium into a credit and a liquidity part. We tested our model on Brazil and Turkey by simultaneously calibrating the formulas to data on their 2, 3, 5 and 10 year CDSs. We found that, on average, 61.5% of the 2 year CDS mid premium of Brazil can be attributed to credit risk and 38.5% to liquidity risk. For Turkey, we found that approximately 54% of the 2 year CDS mid premium can be attributed to credit risk and 46% to liquidity risk. The model-implied one-year default probabilities were found to be 0.25% and 0.4% for Brazil and Turkey, respectively. Furthermore, we were able to deduce that the compensation for liquidity risk in the sovereign CDS market can be attributed to the sell-side, since they receive higher premium payments than they would have received in a perfectly liquid market.

These results indeed confirm that liquidity risk is heavily priced into sovereign CDS premia and that, therefore, this ‘distorting’ component should be accounted for if one wants to use CDS premia to extract default probabilities. The sizes of the different components of the CDS premia are in line with the results of Badaoui, Cathcart and El-Jahel (2013) [5], who find that, on average, 56% of sovereign CDS premia can be attributed to credit risk and 44% to
liquidity risk. Concerning our estimates of the default probabilities, we observe that they are in line with Rabobank’s internal estimates. Of course, this is still no conclusive answer on what the actual default probabilities of these sovereigns are, but the fact that two completely independent methods give more or less the same results strengthens the belief in them.

One of the major attractions to our model is that it allows for a country-specific analysis, which is exactly what we need to compute country-specific default probabilities. Another attraction to the model is that, from a theoretical perspective, it offers much freedom. For example, our model, unlike many other models, allows for a dependence structure between default and liquidity factors (and even interest rate factors if desired). Furthermore, there is, in principle, also great flexibility in modeling the intensities, since the general model structure is invariant to the stochastic processes we assume for the intensities. Of course, using the analytically tractable affine processes, as we did in our analysis, has the advantage that the formulas work out nicely and are more easily calibrated, but, in principle, other processes could be taken too. One could also easily extend the current model to include the bond market, which is done in Bühler and Trapp (2008) [11] and Badaoui, Cathcart and El-Jahel (2013) [5].

Unfortunately, we are currently not able to fully benefit from the model’s flexibility, since the calibration is computationally expensive. Within each grid point several least squares optimization routines have to be passed through and, therefore, we are not able to run very dense grids, since the total calibration time would become too long. The danger of running coarse grids, however, is that, even though we go deeper into the grid, we are more prone to end up in a local optimum. Furthermore, since the number of grid points that need to be evaluated grows exponentially with the number of process parameters, we were now not able to take the most general formulations of some of the processes we used. Apart from the computational speed, we are also not completely satisfied with the calibration results of the correlation factor parameters. Our current results suggest that there are hardly any dependencies between the liquidity and default intensities. We would, however, expect that the default intensity has a significant impact on the liquidity intensities. We think that these results are the consequence of our calibration procedure, since we now completely separate the calibration of the correlation factor parameters from the calibration of the process parameters and intensities.

In the light of the above mentioned comments, we thus think that our model could be improved by speeding up the calibration, since this would allow for more general processes and more dense grids. One possibility could be to specify a very dense grid and to use a simulated annealing approach on this grid. In this way, one does not have to consider all the grid points, since the annealing algorithm chooses the points in a more clever way. Another improvement could thus be made by estimating the correlation factor parameters differently. The challenge here is that, if one looks at the formulas, it appears that the correlation between the intensities are time-dependent. Lastly, we showed in a simulation study that maximum likelihood estimators of process parameters based on discrete-time observations may be biased, and we were not yet able
to find a method that could successfully reduce this. We, therefore, recommend investigating this problem further. A promising solution could be the use of an indirect inference estimation procedure, which is explained by Phillips and Yu (2009)[51]. Unfortunately, we were not able to investigate this any further ourselves.

**Recommendation**

Since estimating sovereign default probabilities is no exact science, we recommend to use our model as a benchmark next to Rabobank’s internal estimates. Our model estimates the market’s perception of the default probability, while Rabobank’s estimates are based on other input. Because of this, we think that our model nicely complements Rabobank’s model and that, therefore, using the models together to benchmark each other is to be preferred over choosing the one over the other. By having two independent estimates of the default probability, Rabobank can better decide on policy issues and better detect risks. We would not yet recommend to let our model replace Rabobank’s model, since there are still some improvements that can be made. For benchmarking purposes we feel, however, that the current model is already sufficient, since the results we obtained for Turkey and Brazil are plausible.
Appendix A

Hazard Process of a Random Time

Within the framework of the intensity-based models, the default time \( \tau \) is modeled as a stopping time that is not predictable with respect to the relevant filtration. The relevant filtration for intensity-based models is often assumed to be given by \( \mathcal{G} = \mathbb{H} \lor \mathcal{F} \), where \( \mathbb{H} \) is the natural filtration generated by the jump process \( H_t = \mathbb{1}_{\{\tau \leq t\}} \) and \( \mathcal{F} \) is some reference filtration, which in applications follows naturally as the filtration generated by a certain stochastic process. For example, Schönbucher (2003) [52], introduces a \( d \)-dimensional background driving process \( X_t \) and states that \( \mathcal{F} \) is the filtration generated by \( X_t \). All default-free processes, such as default-free interest rates, are assumed to be adapted to this filtration \( (\mathcal{F}_t)_{t \geq 0} \). He states that it is not essential that the background filtration is generated by such a stochastic process \( X_t \), but that it convenient to think of it that way.

The main modeling tool in the intensity-based approach is the specification of the conditional probability of default, given that no default has yet occurred. In most cases this is done by modeling the so-called hazard rate or intensity of default (hence the name intensity-based model).

In this Appendix, we will discuss some fundamental results on the hazard process of a random time \( \tau \) with a filtration structure as described above. Some proofs of results stated in chapter 3 are given, but also some other, more fundamental, results will be discussed (e.g. martingale representation theorems). The results of this Appendix are mainly based on chapter 5 of Bielecki & Rutkowski (2002) [6], unless stated otherwise.
A.1 The $\mathcal{F}$-Hazard Process

We consider the set-up where we have a non-negative random time $\tau : \Omega \rightarrow \mathbb{R}_+$ on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$, such that $\mathbb{P}\{\tau = 0\} = 0$ and $\mathbb{P}\{\tau > t\} > t$ for every $t \in \mathbb{R}_+$. Unless stated otherwise, we will assume that the following assumptions are satisfied:

**Condition 4.** The filtration $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$ on $(\Omega, \mathcal{G}, \mathbb{P})$ represents the total information available at time $t$ and is such that we can write $\mathcal{G}_t = \mathcal{H}_t \vee \mathcal{F}_t$ (i.e. $\mathcal{H}_t = \sigma(H_u : u \leq t)$). We also assume, for simplicity, that $\mathcal{F}_0$ is trivial and therefore $\mathcal{G}_0$ is trivial as well. Furthermore, all filtrations are assumed to satisfy the usual conditions of right-continuity and completeness.

We should note that, for given filtrations $\mathcal{H} \subseteq \mathcal{G}$, we do not have a unique filtration $\mathcal{F}$ such that $\mathcal{G}_t = \mathcal{H}_t \vee \mathcal{F}_t$ holds. If, for example, we would have $\mathcal{G}_t = \mathcal{H}_t$, then we could take both $\mathcal{F} = \mathcal{H}$ or $\mathcal{F} = \mathcal{F}_0$ for the assumptions of Condition 4 to hold. Under the above assumptions, we have that $\tau$ is $\mathcal{G}$-stopping time, but not necessarily an $\mathcal{F}$-stopping time. In fact, the case where we have that $\tau$ is an $\mathcal{F}$-stopping time, so $\mathcal{H} \subseteq \mathcal{F}$ (and therefore $\mathcal{F} = \mathcal{G}$), is a special case for which some of the results below do not hold. This is because the $\mathcal{F}$-hazard process, which we will defined below, is not well-defined in this case. Therefore, we will assume throughout that we are not in the case that $\mathcal{F} = \mathcal{G}$. As mentioned before, we usually have that $\mathcal{F}$ follows as the filtration of some pre-defined stochastic process to which all non-defaultable stochastic processes are adapted (Schönbucher, 2003) [52].

In this Appendix, we will write $F_t = \mathbb{P}\{\tau \leq t|\mathcal{F}_t\}$ for every $t \in \mathbb{R}_+$ and we denote the $\mathcal{F}$-survival process of $\tau$ with respect to the filtration $\mathcal{F}$ by $G$, where we can write

$$G_t := 1 - F_t = \mathbb{P}\{\tau > t|\mathcal{F}_t\}, \quad \forall t \in \mathbb{R}_+.$$

**Lemma A.1.** Under the present assumptions we may deal with the càdlàg modifications of $F$ and $G$.

**Proof.** First we note that $F$ is bounded, nonnegative $\mathcal{F}$-submartingale, since for $t \leq s$ we have

$$\mathbb{E}^\mathcal{F}[F_s|\mathcal{F}_t] = \mathbb{E}^\mathcal{F}[\mathbb{E}^\mathcal{F}[\mathbb{P}\{\tau \leq s|\mathcal{F}_s}\]|\mathcal{F}_t] = \mathbb{P}\{\tau \leq s|\mathcal{F}_t\} \geq \mathbb{P}\{\tau \leq t|\mathcal{F}_t\} = F_t.$$

The inequality follows from the fact that, for any $0 \leq t \leq s$, we have $\{\tau \leq t\} \subseteq \{\tau \leq s\}$. We also have that $t \mapsto F_t = \mathbb{E}^\mathcal{F}[\mathbb{P}\{\tau \leq t|\mathcal{F}_t\}] = \mathbb{P}\{\tau \leq t\}$ is right-continuous, since $F_t$ is the cumulative distribution function of $\tau$, which is by definition right-continuous. By Doob’s regularity theorem, it follows that there exists a càdlàg modification of $F$ on $\mathcal{F}$ (see for example Spreij (2012) [53], theorem 2.3). The result for $G$ follows in a similar way. □
Now we have all the notation to define the $\mathcal{F}$-hazard process of $\tau$:

**Definition A.2.** Assume that $F_t < 1$ for $t \in \mathbb{R}_+$. We denote by $\Gamma$ the $\mathcal{F}$-hazard process of $\tau$ under $\mathbb{P}$. Here we have that $\Gamma_t = -\ln G_t = -\ln(1 - F_t)$.

In the case where $\tau$ is an $\mathcal{F}$-stopping time, the above definition of the $\mathcal{F}$-hazard process is not well-defined, since in this case the conditioning on $\mathcal{F}_t$ does not make sense. We will, therefore, explicitly assume that this situation does not occur in the following.

### A.2 Conditional Expectations

In this section, we will give some results concerning conditional expectations and we will investigate how conditioning on $\mathcal{G}_t$ is related to conditioning on $\mathcal{F}_t$. We will start with a result that proves to be useful when dealing with conditional expectations.

**Lemma A.3.** Assume that the filtration $\mathcal{G} \subseteq \mathcal{H} \vee \mathcal{F}$, that is, $\mathcal{G}_t \subseteq \mathcal{H}_t \vee \mathcal{F}_t$ for every $t \in \mathbb{R}_+$. Then $\mathcal{G} \subseteq \mathcal{G}^*$, where $\mathcal{G}^* = (\mathcal{G}^*_t)_{t \geq 0}$ with

$$\mathcal{G}^*_t := \{ A \in \mathcal{G} : \exists B \in \mathcal{F}_t, A \cap \{ \tau > t \} = B \cap \{ \tau > t \} \}.$$ 

**Proof.** First, we will show that $\mathcal{G}^*_t$ is a $\sigma$-algebra.

1. Obviously $\Omega \in \mathcal{G}^*_t$, since we can take $B = \Omega$ and $\Omega \in \mathcal{F}_t$.
2. Let $A \in \mathcal{G}^*_t$, then we want $A^c = \Omega \setminus A \in \mathcal{G}^*_t$. So we need a $\tilde{B} \in \mathcal{F}_t$ such that $A^c \cap \{ \tau > t \} = \tilde{B} \cap \{ \tau > t \}$. We define the following 3 cases:
   (i) Suppose $A \in \mathcal{F}_t$, then we can take $\tilde{B} = A^c$.
   (ii) Suppose $A \in \mathcal{H}_t$, then we can take $\tilde{B} = \emptyset$ or $\tilde{B} = \Omega$, depending on whether $A^c \cap \{ \tau > t \} = \emptyset$ or $A^c \cap \{ \tau > t \} = \{ \tau > t \}$ (and exactly 1 of these cases must occur).
   (iii) Suppose $A \in \mathcal{H}_t \vee \mathcal{F}_t$, so for example $A = M \cap N$ or $A = M \cup N$ for some $M \in \mathcal{H}_t$ and $N \in \mathcal{F}_t$. Then there are of course again different possibilities, but we will just show for 1 of these how we can find $\tilde{B}$, since the other cases are similar. So suppose that $A = M \cap N$. Now $A^c = M^c \cup N^c$ by DeMorgan’s law. Since $M \in \mathcal{H}_t$, we have either $A \cap \{ \tau > t \} = M \cap N \cap \{ \tau > t \} = N \cap \{ \tau > t \}$ if $M \cap \{ \tau > t \} = \{ \tau > t \}$ or $A \cap \{ \tau > t \} = \emptyset$ if $M \cap \{ \tau > t \} = \emptyset$ (and these are the only options possible). In the latter case we have that $A^c \cap \{ \tau > t \} = (M^c \cup N^c) \cap \{ \tau > t \} = \{ \tau > t \} \cup (N^c \cap \{ \tau > t \}) = N^c \cap \{ \tau > t \}$. So we can take $\tilde{B} = N^c \in \mathcal{F}_t$. In the former case we have $A^c \cap \{ \tau > t \} = (M^c \cup N^c) \cap \{ \tau > t \} = N^c \cap \{ \tau > t \}$, so again we can take $\tilde{B} = N^c \in \mathcal{F}_t$. The case where $A = M \cup N$ follows in a similar fashion.
3. Let $A_1, A_2 \in G_t^*$, then $\exists B_1, B_2 \in \mathcal{F}_t$ such that $A_1 \cap \{\tau > t\} = B_1 \cap \{\tau > t\}$ and $A_2 \cap \{\tau > t\} = B_2 \cap \{\tau > t\}$. Now we get $(A_1 \cup A_2) \cap \{\tau > t\} = (B_1 \cup B_2) \cap \{\tau > t\}$, and we know $B_1 \cup B_2 \in \mathcal{F}_t$ since $\mathcal{F}_t$ is a $\sigma$-algebra. In a similar way we have that for $A_n \in G_t^*$ for $n = 1, 2, \ldots$, that also $\bigcup_{n=1}^{\infty} A_n \in G_t^*$, since we can put $\tilde{B} = \bigcup_{n=1}^{\infty} B_n \in \mathcal{F}_t$.

We have thus showed that $G_t^*$ is a $\sigma$-algebra. Now we will show that $H_t \subseteq G_t^*$ and $\mathcal{F}_t \subseteq G_t^*$. To see the latter statement, we simply take $B = A$ for every $A \in \mathcal{F}_t$. For the former statement we note that for any $A \in H_t$, we have $A \cap \{\tau > t\} = \emptyset$ or $A \cap \{\tau > t\} = \{\tau > t\}$. So we can take $B = \emptyset$ or $B = \Omega$ respectively. We have thus showed that $G_t^*$ is a $\sigma$-algebra containing $\mathcal{F}_t$ and $H_t$. By assumption, we know that $G_t \subseteq H_t \lor C \mathcal{F}_t$, and, therefore, we have showed that $G_t \subseteq G_t^*$ and, hence, $G \subseteq G^*$.

We will use Lemma A.3 in the proof of the Lemma A.4 below. Brigo and Mercurio (2006) [9] call the result of Lemma A.4 the "filtration switching formula".

**Lemma A.4.** Let $Y$ be a $\mathcal{G}$-measurable random variable.

(i) Assume we have $\mathcal{F}_t \subseteq \mathcal{G}_t \subseteq H_t \lor \mathcal{F}_t$ for every $t \in \mathbb{R}_+$, then we have

$$
\mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{G}_t \right] = \mathbb{P}\{\tau > t \mid \mathcal{G}_t \} \mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{F}_t \right]. 
$$

(A.1)

(ii) If, in addition, $H_t \subseteq \mathcal{G}_t$ (now all assumptions stated in Condition 4 hold), then

$$
\mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{F}_t \right]. 
$$

(A.2)

In particular, for any $t \leq s$, we have

$$
\mathbb{P}\{t < \tau \leq s \mid \mathcal{G}_t \} = 1_{\{\tau > t\}} \frac{\mathbb{P}\{t < \tau \leq s \mid \mathcal{F}_t \}}{\mathbb{P}\{\tau > t \mid \mathcal{F}_t \}}.
$$

(A.3)

**Proof.** In order to prove (A.1), we will show that

$$
\mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{G}_t \right] \mathbb{P}\{\tau > t \mid \mathcal{F}_t \} = \mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mathbb{P}\{\tau > t \mid \mathcal{F}_t \} \mid \mathcal{G}_t \right] \quad \text{(since $\mathcal{F}_t \subseteq \mathcal{G}_t$)}
$$

is equal to

$$
\mathbb{P}\{\tau > t \mid \mathcal{G}_t \} \mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{F}_t \right] = \mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} \mathbb{E}^\mathbb{P}\left[1_{\{\tau > t\}} Y \mid \mathcal{F}_t \right] \mid \mathcal{G}_t \right].
$$

So we will check for every $A \in \mathcal{G}_t$ that the following holds:
A.2. Conditional Expectations

\[ \int_A 1_{\{\tau > t\}} Y \mathbb{P}\{\tau > t|\mathcal{F}_t\} \, d\mathbb{P} = \int_A 1_{\{\tau > t\}} \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y \bigg| \mathcal{F}_t \right] \, d\mathbb{P}. \]

By Lemma A.3, we have that for any \( A \in \mathcal{G}_t \) there exists a \( B \in \mathcal{F}_t \) such that \( A \cap \{\tau > t\} = B \cap \{\tau > t\} \). We get

\[ \int_A 1_{\{\tau > t\}} Y \mathbb{P}\{\tau > t|\mathcal{F}_t\} \, d\mathbb{P} = \int_{A \cap \{\tau > t\}} Y \mathbb{P}\{\tau > t|\mathcal{F}_t\} \, d\mathbb{P} = \int_{B \cap \{\tau > t\}} Y \mathbb{P}\{\tau > t|\mathcal{F}_t\} \, d\mathbb{P} = \int_B \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y \bigg| \mathcal{F}_t \right] \mathbb{P}\{\tau > t|\mathcal{F}_t\} \, d\mathbb{P} = \int_B \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y \bigg| \mathcal{F}_t \right] \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y \bigg| \mathcal{F}_t \right] \, d\mathbb{P} = \int_{B \cap \{\tau > t\}} \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y \bigg| \mathcal{F}_t \right] \, d\mathbb{P} = \int_A \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y \bigg| \mathcal{F}_t \right] \, d\mathbb{P}. \]

Formula (A.1) follows. To prove (A.2), we note that we get \( \mathbb{P}\{\tau > t|\mathcal{G}_t\} = \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} | \mathcal{G}_t \right] = 1_{\{\tau > t\}} \), since \( 1_{\{\tau > t\}} \) is \( \mathcal{H}_t \)-measurable (and therefore also \( \mathcal{G}_t \)-measurable). The rest follows as in the proof of (A.1). In order to prove equation (A.3), we use (A.2) and we take \( Y = 1_{\{\tau \leq s\}} \).

The following result gives some modifications of equations (A.1)-(A.3).

**Corollary A.5.** Let \( Y \) be \( \mathcal{G} \)-measurable random variable and let \( t \leq s \).

(i) If for every \( t \in \mathbb{R}_+ \) we have \( \mathcal{F}_t \subseteq \mathcal{G}_t \subseteq \mathcal{H}_t \lor \mathcal{F}_t \), then

\[ \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > s\}} Y | \mathcal{G}_t \right] = \mathbb{P}\{\tau > t|\mathcal{G}_t\} \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > s\}} e^{r_t} Y \bigg| \mathcal{F}_t \right]. \]  

(A.4)

(ii) If the assumptions stated in Condition 4 hold, then

\[ \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > s\}} Y | \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > s\}} e^{r_t} Y \bigg| \mathcal{F}_t \right], \]

(A.5)

147
and

\[ \mathbb{E}^P \left[ 1_{\{t < \tau \leq s\}} Y \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^P \left[ 1_{\{t < \tau \leq s\}} e^{\Gamma_s} Y \mid \mathcal{F}_t \right]. \tag{A.6} \]

(iii) If in addition \( Y \) is \( \mathcal{F}_s \)-measurable, then

\[ \mathbb{E}^P \left[ 1_{\{\tau > s\}} Y \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^P \left[ e^{\Gamma_t - \Gamma_s} Y \mid \mathcal{F}_t \right] \tag{A.7} \]

and

\[ \mathbb{E}^P \left[ 1_{\{t < \tau \leq s\}} Y \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^P \left[ 1_{\{\tau > t\}} (1 - e^{-\Gamma_s}) e^{\Gamma_t} Y \mid \mathcal{F}_t \right]. \]

Proof. See Bielecki and Rutkowski (2002) [6], Corollary 5.1.1.

Until now, we have focused on conditional expectations on the set \( \{\tau > t\} \). Now we want to look at expectations of the form \( \mathbb{E}^P \left[ 1_{\{\tau \leq t\}} Y \mid \mathcal{G}_t \right] \).

Lemma A.6. Under the assumptions stated in Condition 4, we have, for any \( \mathcal{G} \)-measurable random variable \( Y \), that

\[ \mathbb{E}^P \left[ 1_{\{\tau \leq t\}} Y \mid \mathcal{G}_t \right] = 1_{\{\tau \leq t\}} \mathbb{E}^P \left[ Y \mid \mathcal{H}_\infty \vee \mathcal{F}_t \right]. \tag{A.8} \]

Proof. First note that, under the present conditions, we have, for any \( t \in \mathbb{R}_+ \), and any event \( A \in \mathcal{H}_\infty \vee \mathcal{F}_t \), that \( A \cap \{\tau \leq t\} \in \mathcal{G}_t \). We get for any \( A \in \mathcal{H}_\infty \vee \mathcal{F}_t \) the following:

\[
\int_A \mathbb{E}^P \left[ 1_{\{\tau \leq t\}} Y \mid \mathcal{H}_\infty \vee \mathcal{F}_t \right] d\mathbb{P} = \int_A 1_{\{\tau \leq t\}} Y d\mathbb{P} = \int_{A \cap \{\tau \leq t\}} Y d\mathbb{P} = \mathbb{E}^P \left[ Y \mid \mathcal{G}_t \right] d\mathbb{P} = \int_A 1_{\{\tau \leq t\}} \mathbb{E}^P \left[ Y \mid \mathcal{G}_t \right] d\mathbb{P}
\]

We have that the random variable \( 1_{\{\tau \leq t\}} \mathbb{E}^P \left[ Y \mid \mathcal{G}_t \right] \) is \( \mathcal{H}_t \vee \mathcal{G}_t = \mathcal{H}_t \vee \mathcal{F}_t \)-measurable and, therefore, also \( \mathcal{H}_\infty \vee \mathcal{F}_t \)-measurable. This concludes the proof.

We are now able to deal with the full conditional expectation \( \mathbb{E} \left[ Y \mid \mathcal{G}_t \right] \):
Corollary A.7. For any $G$-measurable random variable $Y$ we have

$$
\mathbb{E}^P [Y | G_t] = \mathbb{E}^P \left[ 1_{\{\tau \leq t\}} Y \bigg| G_t \right] + \mathbb{E}^P \left[ 1_{\{\tau > t\}} e^{\Gamma_t} Y \bigg| F_t \right].
$$

Proof. First we write $\mathbb{E}^P [Y | G_t] = \mathbb{E}^P \left[ 1_{\{\tau \leq t\}} Y \bigg| G_t \right] + \mathbb{E}^P \left[ 1_{\{\tau > t\}} Y | G_t \right]$. Now we use for the first part Lemma A.6 and for the second part Corollary A.7. We note that we can write $\mathbb{E}^P \left[ 1_{\{\tau > t\}} | F_t \right] = e^{\Gamma_t}$, which is $F_t$-measurable. The result follows. $\square$

The following result deals with the conditional expectations of a bounded and continuous function of $\tau$ and of a bounded, $F$-predictable, process.

Proposition A.8. (i) Let $h : \mathbb{R}_+ \to \mathbb{R}$ be a bounded, continuous function. Then for any $t < s \leq \infty$

$$
\mathbb{E}^P \left[ 1_{\{t < \tau \leq s\}} h(\tau) \bigg| G_t \right] = \mathbb{E}^P \left[ 1_{\{\tau > t\}} e^{\Gamma_t} \mathbb{E}^P \left[ \int_{[t,s]} h(u) dF_u \bigg| F_t \right] \right]. \tag{A.9}
$$

(ii) Let $Z$ be a bounded $F$-predictable process. Then for any $t < s \leq \infty$

$$
\mathbb{E}^P \left[ 1_{\{t < \tau \leq s\}} Z_{\tau} \bigg| G_t \right] = \mathbb{E}^P \left[ 1_{\{\tau > t\}} e^{\Gamma_t} \mathbb{E}^P \left[ \int_{[t,s]} Z_u dF_u \bigg| F_t \right] \right]. \tag{A.10}
$$

Proof. The proof of (A.9) follows by a similar argument as the proof of (A.10) and, therefore, we will focus only on the proof of the latter.

Let us assume that $Z$ is a stepwise $F$-predictable process of the form: $Z_u = \sum_{i=0}^{n} Z_{t_i} 1_{\{t_i < u \leq t_{i+1}\}}$ for $t < u \leq s$ and $t_0 = t < \cdots < t_{n+1} = s$ and where $Z_{t_i}$ is a $F_{t_i}$-measurable random variable for $i = 0, \ldots, n$. We get the following:

$$
\mathbb{E}^P \left[ 1_{\{t < \tau \leq s\}} Z_{\tau} \bigg| G_t \right] = \mathbb{E}^P \left[ \sum_{i=0}^{n} \mathbb{E}^P \left[ 1_{\{t_i < \tau \leq t_{i+1}\}} Z_{t_i} \bigg| F_{t_{i+1}} \right] \bigg| F_t \right]
$$

$$
= \mathbb{E}^P \left[ \sum_{i=0}^{n} 1_{\{t_i < \tau \leq t_{i+1}\}} Z_{t_i} \bigg| F_t \right]
$$

$$
= \mathbb{E}^P \left[ \sum_{i=0}^{n} Z_{t_i} \mathbb{E}^P \left[ 1_{\{t_i < \tau \leq t_{i+1}\}} \bigg| F_{t_i} \right] \bigg| F_t \right]
$$

$$
= \mathbb{E}^P \left[ \sum_{i=0}^{n} Z_{t_i} (F_{t_{i+1}} - F_{t_i}) \bigg| F_t \right]
$$
If we now have a bounded and \( \mathbb{F} \)-predictable process \( Z \), then we can approximate it by a sequence of bounded, stepwise, \( \mathbb{F} \)-predictable processes. By the dominated convergence theorem, we conclude that the result holds and that, for any bounded, \( \mathbb{F} \)-predictable process \( Z \), we have

\[
\mathbb{E}^\mathbb{P}\left[ \mathbb{1}_{\{t<\tau\leq s\}} Z_\tau \mid \mathcal{G}_t \right] = \mathbb{E}^\mathbb{P}\left[ \int_{[t,s]} Z_u d\mathcal{F}_u \bigg| \mathcal{F}_t \right]. \tag{A.11}
\]

\[\square\]

A.3 Martingale Results

In this section, we will give several results on martingales associated with the hazard process \( \Gamma \) and some martingale representation theorems that follow from these results. Some of the martingale representation theorems are used in section 3.4, where the change of probability measure is discussed.

A.3.1 Martingales Associated with \( \Gamma \)

The following two results will be used in some martingale representation theorems below.

**Lemma A.9.** The process \( L \), given by

\[
L_t := \mathbb{1}_{\{\tau>t\}} e^{\Gamma_t} = (1 - H_t) e^{\Gamma_t} = \frac{1 - H_t}{1 - F_t}, \tag{A.12}
\]

follows a \( \mathcal{G} \)-martingale. Furthermore, for any bounded \( \mathbb{F} \)-martingale \( m \), the product \( Lm \) is \( \mathcal{G} \)-martingale. If \( m \) also follows a \( \mathcal{G} \)-martingale then the quadratic covariation \([L,m]\), given by

\[
[L,m]_t := L_t m_t - L_0 M_0 - \int_{[0,t]} L_s dm_s - \int_{[0,t]} m_s dL_s, \tag{A.13}
\]

follows a \( \mathcal{G} \)-martingale.

**Proof.** By formula (A.5), we have, for any \( t \leq s \), that

\[
\mathbb{E}^\mathbb{P}\left[ \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mid \mathcal{G}_t \right] = \mathbb{1}_{\{\tau>t\}} e^{\Gamma_t} \mathbb{E}^\mathbb{P}\left[ \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mid \mathcal{F}_t \right].
\]

We can rewrite the last expression, however, as follows.
A.3. Martingale Results

\[ 1_{\{\tau > t\}} e^{\Gamma_t} \mathbb{E}^P \left[ 1_{\{\tau > s\}} e^{\Gamma_s} \mid \mathcal{F}_t \right] = 1_{\{\tau > t\}} e^{\Gamma_t} \mathbb{E}^P \left[ e^{\Gamma_s} \mathbb{E}^P \left[ 1_{\{\tau > s\}} \mid \mathcal{F}_s \right] \mid \mathcal{F}_t \right] = 1_{\{\tau > t\}} e^{\Gamma_t} e^{-\Gamma_s} \mathbb{E}^P \left[ 1_{\{\tau > s\}} \mid \mathcal{F}_s \right] \mid \mathcal{F}_t \]  

Hence, we have showed that

\[ \mathbb{E}^P \left[ 1_{\{\tau > s\}} e^{\Gamma_s} \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} e^{\Gamma_t}, \]

and, therefore, we conclude that \( L \) follows a \( \mathbb{G} \)-martingale.

For the second statement, we have, for \( t \leq s \), that

\[ \mathbb{E}^P \left[ L_s m_s \mid \mathcal{G}_t \right] = \mathbb{E}^P \left[ 1_{\{\tau > s\}} e^{\Gamma_s} m_s \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} e^{\Gamma_t} e^{\Gamma_s} m_s \mathbb{E}^P \left[ 1_{\{\tau > s\}} \mid \mathcal{F}_s \right] \mid \mathcal{F}_t \]

Hence, the product \( Lm \) follows a \( \mathbb{G} \)-martingale. Now it is easy to see that the last statement follows directly. \( \square \)

**Proposition A.10.** Assume that the \( \mathbb{F} \)-hazard process \( \Gamma \) of \( \tau \) is a continuous and increasing process. Then the following statements are valid:

(i) The process \( \hat{M}_t = H_t - \Gamma_{t \wedge \tau} \) follows a \( \mathbb{G} \)-martingale, specifically,

\[ \hat{M}_t = -\int_{[0,t]} e^{-\Gamma_u} dL_u. \tag{A.14} \]

Furthermore, \( L_t = \mathcal{E}_t(-\hat{M}) \), so \( L \) solves the following equation:

\[ L_t = 1 - \int_{[0,t]} L_u - d\hat{M}_u. \tag{A.15} \]

(ii) If a bounded \( \mathbb{F} \)-martingale \( m \) is also a \( \mathbb{G} \)-martingale, then the product \( \hat{M}m \) is a \( \mathbb{G} \)-martingale.
(iii) If $m$ is a bounded, predictable, $\mathbb{F}$-martingale, then the product $\hat{M}\tilde{m}$ is a $\mathbb{G}$-martingale, where $\tilde{m} = m_{t\wedge \tau}$.

Proof. See Bielecki and Rutkowski (2002) [6], Proposition 5.1.3. \hfill \Box

A.3.2 Martingale Representation Theorems

In this section, we will give some martingale representation theorems for $\mathbb{G}$-martingales. We will assume that the process $F_t = \mathbb{P}\{\tau \leq t|\mathcal{F}_t\}$ (and as a consequence $\Gamma$) is continuous and increasing. We denote $\hat{M}_t = H_t - \Gamma_{t\wedge \tau}$ and we know by Proposition A.10 that it is a $\mathbb{G}$-martingale.

Proposition A.11. Assume that the $\mathbb{F}$-hazard process $\Gamma$ of $\tau$ follows an increasing continuous process. Let $Z$ be an $\mathbb{F}$-predictable process such that the random variable $Z_\tau$ is integrable. Then the $\mathbb{G}$-martingale $M^Z_t := \mathbb{E}^\mathbb{P}[Z_\tau|\mathcal{G}_t]$ admits the following decomposition

$$M^Z_t = m_0 + \int_{[0,t]} e^{\Gamma_u}d\tilde{m}_u + \int_{[0,t]} (Z_u - D_u)d\hat{M}_u,$$

where $\tilde{m}_t = m_{t\wedge \tau}$, and $m$ is an $\mathbb{F}$-martingale, namely,

$$m_t = \mathbb{E}^\mathbb{P}\left[\int_0^\infty Z_u e^{-\Gamma_u}d\Gamma_u \bigg| \mathcal{F}_t\right] = \mathbb{E}^\mathbb{P}\left[\int_0^\infty Z_u dF_u \bigg| \mathcal{F}_t\right], \tag{A.17}$$

hence, in particular, $m_0 = M^Z_0$. Moreover,

$$D_t = \mathbb{E}^\mathbb{P}\left[\int_0^\infty Z_u e^{\Gamma_t - \Gamma_u}d\Gamma_u \bigg| \mathcal{F}_t\right] = e^{\Gamma_t}\mathbb{E}^\mathbb{P}\left[\int_t^\infty Z_u dF_u \bigg| \mathcal{F}_t\right].$$

Proof. We have

$$M^Z_t = \mathbb{E}^\mathbb{P}[Z_\tau|\mathcal{G}_t] = \mathbb{1}_{\{\tau \leq t\}}Z_\tau + \mathbb{1}_{\{\tau > t\}}\mathbb{E}^\mathbb{P}\left[\int_t^\infty Z_u e^{\Gamma_t - \Gamma_u}d\Gamma_u \bigg| \mathcal{F}_t\right]
= \mathbb{1}_{\{\tau \leq t\}}Z_\tau + \mathbb{1}_{\{\tau > t\}}D_t.$$

Here the first equality follows from (A.10) and the fact that we have $dF_u = e^{-\Gamma_u}d\Gamma_u$ as $\Gamma$ is continuous. We now have $M^Z_t = M^{Z^\tau}_{t\wedge \tau}$ for every $t \in \mathbb{R}_+$. We can write
A.3. Martingale Results

\[ D_t = e^{\Gamma_t} m_t - e^{\Gamma_t} \int_0^t Z_u e^{-\Gamma_u} d\Gamma_u. \]

We have that \( e^{\Gamma_t} \) is a continuous, increasing process, since \( \Gamma_t \) is increasing and continuous. We, therefore, have that the quadratic covariations \( [e^{\Gamma}, m] \) and \( [e^{\Gamma}, \int_0^t Z_u e^{-\Gamma_u} d\Gamma_u] \) are zero. We get by the (Itô) integration by parts formula the following:

\[
e^{\Gamma_t} m_t = e^{\Gamma_0} m_0 + \int_0^t e^{\Gamma_u} dm_u + \int_0^t m_u e^{\Gamma_u} d\Gamma_u = m_0 + \int_0^t e^{\Gamma_u} dm_u + \int_0^t m_u e^{\Gamma_u} d\Gamma_u,
\]

and

\[
e^{\Gamma_t} \int_0^t Z_u e^{-\Gamma_u} d\Gamma_u = 0 + \int_0^t e^{\Gamma_u} Z_u e^{-\Gamma_u} d\Gamma_u + \int_0^t \left( \int_0^u Z_v e^{-\Gamma_v} d\Gamma_v \right) e^{\Gamma_u} d\Gamma_u = \int_0^t Z_u d\Gamma_u + \int_0^t e^{\Gamma_u} \int_0^u Z_v e^{-\Gamma_v} d\Gamma_v d\Gamma_u.
\]

Combining these expressions gives

\[
D_t = m_0 + \int_0^t e^{\Gamma_u} dm_u + \int_0^t m_u e^{\Gamma_u} d\Gamma_u - \int_0^t Z_u d\Gamma_u - \int_0^t e^{\Gamma_u} \int_0^u Z_v e^{-\Gamma_v} d\Gamma_v d\Gamma_u = m_0 + \int_0^t e^{\Gamma_u} dm_u - \int_0^t Z_u d\Gamma_u + \int_0^t \left( m_u e^{\Gamma_u} - \int_0^u Z_v e^{-\Gamma_v} d\Gamma_v \right) d\Gamma_u = m_0 + \int_0^t e^{\Gamma_u} dm_u + \int_0^t (D_u - Z_u) d\Gamma_u.
\]

Furthermore, we have \( D_t = m_0 + \int_0^t dD_u \) and, therefore, we get

\[
1_{\{\tau > t\}} D_t = m_0 + \int_0^t dD_u - 1_{\{\tau \leq t\}} D_\tau = m_0 + \int_0^{t \wedge \tau} dD_u - 1_{\{\tau \leq t\}} D_\tau = m_0 + \int_0^{t \wedge \tau} e^{\Gamma_u} dm_u + \int_0^{t \wedge \tau} (D_u - Z_u) d\Gamma_u - 1_{\{\tau \leq t\}} D_\tau.
\]
We conclude that

\[ M^Z_t = m_0 + \int_0^{t \land \tau} e^{\Gamma_u} \, dm_u + \int_0^{t \land \tau} (D_u - Z_u) \, d\Gamma_u + 1_{\{\tau \leq t\}}(Z_\tau - D_\tau), \]

and we observe that

\[ 1_{\{\tau \leq t\}}(Z_\tau - D_\tau) = \int_0^t (Z_t - D_t) \, dH_t. \]

When we combine this with \( \int_0^{t \land \tau} (D_u - Z_u) \, d\Gamma_u \), we get

\[ \int_0^{t \land \tau} (D_u - Z_u) \, d\Gamma_u + 1_{\{\tau \leq t\}}(Z_\tau - D_\tau) = \int_0^t (Z_u - D_u) \, d\hat{M}_u. \]

This concludes the proof of formula (A.16).

\[ \square \]

The following result is specifically relevant in the situation where \( \mathbb{F} \) is generated by a Brownian motion.

**Corollary A.12.** Suppose all conditions of Proposition A.11 hold, and if, in addition, \( \Delta D_\tau = 0 \) (or equivalently \( \Delta m_\tau = 0 \)), then

\[ M^Z_t = m_0 + \int_{[0,t]} e^{\Gamma_u} \, d\tilde{m}_u + \int_{[0,t]} (Z_u - M^Z_u) \, d\hat{M}_u, \quad (A.18) \]

so that

\[ M^Z_t = m_0 + \tilde{M}^Z_t + \hat{M}^Z_t, \]

where the two \( \mathbb{G} \)-martingales \( \tilde{M}^Z_t \) and \( \hat{M}^Z_t \) are mutually orthogonal (so the product \( \tilde{M}^Z_t \hat{M}^Z_t \) is a \( \mathbb{G} \)-martingale).

**Proof.** We saw in the proof of Proposition A.11 that we could write \( M^Z_t = 1_{\{\tau \leq t\}}Z_\tau + 1_{\{\tau > t\}}D_t \). This means that for \( u < \tau \) we have \( D_u = M^Z_u \). Since, by assumption, \( \Delta D_\tau = 0 \), we get that \( D_\tau = D_{\tau -} = M^Z_{\tau -} \). Consequently, we get

\[ 1_{\{\tau > t\}}D_t = m_0 + \int_0^t dD_u - 1_{\{\tau \leq t\}}M^Z_{\tau -}. \]

This combined with the continuity of \( \Gamma \) proves (A.18). For the second part of the Corollary, we note that the process \( m \), given by (A.17), is a bounded and \( \mathbb{F} \)-predictable \( \mathbb{F} \)-martingale, since
A.3. Martingale Results

it is integrable and since $Z$ is an $\mathcal{F}$-predictable process and $\Gamma$ is continuous and increasing. By Proposition A.10 part (iii), we get that the quadratic covariation $[\hat{M}, \tilde{m}]$ is a $\mathcal{G}$-martingale and, therefore,

$$[\hat{M}^Z, \hat{M}^Z]_t = \int_{[0,t]} e^{\Gamma_u} (Z_u - M_u^Z) d[\hat{M}, \tilde{m}]_u$$

follows a $\mathcal{G}$-martingale. We have that the above equality follows from Proposition 6.9 in Spreij (2012) [53]. By Itô’s integration by parts formula, we get that

$$\tilde{M}^Z \hat{M}^Z = 0 + \int_0^t \tilde{M}^Z_u d\hat{M}^Z_u + \int_0^t \hat{M}^Z_u d\tilde{M}^Z_u + [\tilde{M}^Z, \hat{M}^Z]_t$$

is an $\mathcal{G}$-martingale and, therefore, $\tilde{M}^Z \hat{M}^Z$ is a $\mathcal{G}$-martingale. \hfill \Box

In the case that the filtration $\mathcal{F}$ is the natural filtration of some Brownian motion $W$, we have that $\mathcal{F}$ only supports continuous martingales. Therefore, the process $m$ given by formula (A.17) is continuous and we have $\Delta m_t = 0$ (so the conditions of Corollary A.12 are satisfied). If we, furthermore, postulate that the Brownian motion $W$ remains a Brownian motion with respect to the filtration $\mathcal{G}$, then for a fixed $T > 0$ we get the following proposition:

**Proposition A.13.** Let $L$ be as defined in formula (A.12). For a $\mathcal{G}$-measurable and integrable random variable $X$, we define the $\mathcal{G}$-martingale $M^X_t = \mathbb{E}^\mathcal{G}[X|G_t]$ for $t \in [0,T]$. Then $M^X$ admits the following representation

$$M^X_t = M^X_0 + \int_0^t \xi^X_u dW_u + \int_{[0,t]} \tilde{\xi}^X_u dL_u = M^X_t + \tilde{L}^X_t + \hat{L}^X_t,$$

where $\xi^X$ and $\tilde{\xi}^X$ are $\mathcal{G}$-predictable stochastic processes. Moreover, the $\mathcal{G}$-martingales $\tilde{L}^X_t$ and $\hat{L}^X_t$ are mutually orthogonal.

**Proof.** See Bielecki and Rutkowski (2002) [6], Proposition 5.2.2. \hfill \Box

We will end this section with the following result:

**Corollary A.14.** Suppose that, next to the assumptions of Proposition A.13, the hazard process $\Gamma$ is continuous. Then for any $\mathcal{G}$-martingale $N$ we have

$$N = N_0 + \int_0^t \xi^N_u dW_u + \int_{[0,t]} \zeta^N_u d\hat{M}_u = N_0 + \tilde{M}^N_t + \hat{M}^N_t,$$
where $\xi^N$ and $\zeta^N$ are $\mathbb{G}$-predictable stochastic processes. The $\mathbb{G}$-martingales $\tilde{M}_t^N$ and $\hat{M}_t^N$ are mutually orthogonal.

Proof. Note that for every $\mathbb{G}$-martingale we can write $N_t = \mathbb{E}^P[X|\mathcal{G}_t]$ for a $\mathcal{G}$-measurable and integrable random variable $X$, since we can simply take $X = N_s$ for $s > t$. Therefore, the result of Proposition A.13 holds and we can write

$$N_t = N_0 + \int_0^t \xi^X_u dW_u + \int_{[0,t]} \tilde{\zeta}^X_u dL_u,$$

where $\xi^X$ and $\tilde{\zeta}^X$ are $\mathbb{G}$-predictable stochastic processes. By part (i) in Proposition A.10, we get that

$$\int_{[0,t]} \tilde{\zeta}^X_u dL_u = -\int_{[0,t]} \tilde{\eta}^X_u L_u - d\hat{M}_u.$$

We have that the result follows by taking $\zeta^N_t = -\tilde{\zeta}^X_t L_t -$ and $\xi^X_t = \xi^N_t$. The mutual orthogonality follows from the fact that $\hat{M}$ is of finite variation, and therefore the quadratic covariation $[W, \hat{M}]$ is zero and we get

$$[\tilde{M}_t^N, \hat{M}_t^N]_t = \int_0^t \xi^N_u \tilde{\zeta}^N_u d[W, \hat{M}]_u.$$

This equality follows from proposition 6.9 of Spreij (2012) [53]. Now by Itô we get

$$\tilde{M}_t^N \hat{M}_t^N = \int_0^t \tilde{M}_u^N \hat{M}_u^N + \int_0^t \hat{M}_u^N d\tilde{M}_u^N,$$

which is obviously a $\mathbb{G}$-martingale, and hence, $\hat{M}^N$ and $\tilde{M}^N$ are mutually orthogonal. \qed
Appendix B

Martingale Hazard Process of a Random Time

In this Appendix, we will discuss some properties of the so-called martingale hazard process. As the name already suggests, this process is closely related to the hazard process and in most practical applications these notions even coincide. The martingale hazard process is a valuable tool in the situation where we are considering measure changes. In the case of hazard processes, it is not always obvious what the hazard process of a stopping time is under an equivalent measure, whereas for martingale hazard processes more general results hold. The results stated in this Appendix are mainly taken from Bielecki & Rutkowski (2002) [6].

B.1 The $(\mathcal{F}, \mathcal{G})$-Martingale Hazard Process

In this section, we will give the definition of the martingale hazard process and we will give a characterization of it.

We again consider a non-negative random time $\tau$ on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}\{\tau = 0\} = 0$ and $\mathbb{P}\{\tau > t\} > t$ for every $t \in \mathbb{R}_+$. Furthermore, we assume that the assumptions as stated in Condition 4 hold, and that we, therefore, have $\mathcal{G} = \mathcal{H} \vee \mathcal{F}$. We get the following definition of the $(\mathcal{F}, \mathcal{G})$-martingale hazard process:

Definition B.1. An $\mathcal{F}$-predictable, right-continuous, increasing process $\Lambda$, with $\Lambda_0 = 0$, is called a $(\mathcal{F}, \mathcal{G})$-martingale hazard process if and only if the process $\tilde{M}_t := H_t - \Lambda_{t \wedge \tau}$ follows a $\mathcal{G}$-martingale. If $\Lambda_t = \int_0^t \lambda_u du$, then the $\mathcal{F}$-progressively measurable non-negative process $\lambda$ is referred to as the $(\mathcal{F}, \mathcal{G})$-martingale intensity process.
Chapter B. Martingale Hazard Process of a Random Time

Note that, since $G = H \vee F$, the filtration $F$ uniquely specifies $G$, and, therefore, we will refer to the $F$-martingale hazard process instead of the $(F, G)$-martingale hazard process. From now on, we will make the following, additional assumption:

**Condition 5.** For all $t \in \mathbb{R}^+$ and every $u \leq t$, we have

$$\mathbb{P}\{\tau > u|\mathcal{F}_\infty\} = \mathbb{P}\{\tau > t|\mathcal{F}_t\}.$$ 

The following Lemma, taken from Filipović (2009) [27], puts this assumption to context:

**Lemma B.2.** The following statements are equivalent:

(i) Condition 5 holds.

(ii) We have $\mathbb{E}^P[X|G_t] = \mathbb{E}^P[X|F_t]$ for every bounded, $\mathcal{F}_\infty$-measurable random variable $X$ and for any $t \in \mathbb{R}_+$.

(iii) Every $F$-martingale also follows a $G$-martingale (martingale invariance property of $F$ w.r.t. $G$).

**Proof.** We will first show $(i) \iff (ii)$. Let $A \in \mathcal{F}_t$, $u \leq t$ and let $X$ be a bounded, $\mathcal{F}_\infty$-measurable random variable. We define

$$I_1 = \int_{A \cap \{\tau > u\}} X d\mathbb{P} = \int_A X \mathbb{E}^P[1_{\{\tau > u\}}|\mathcal{F}_\infty]d\mathbb{P},$$

and

$$I_2 = \int_{A \cap \{\tau > u\}} \mathbb{E}^P[X|\mathcal{F}_t]d\mathbb{P} = \int_A X \mathbb{E}^P[1_{\{\tau > u\}}|\mathcal{F}_t]d\mathbb{P}.$$ 

Since $G_t$ is generated by sets of the form $A \cap \{\tau > u\}$ (since $G_t = H_t \vee \mathcal{F}_t$), we get that $(i) \Rightarrow (ii)$ implies $I_1 = I_2$ and also $(ii) \Rightarrow (i)$ implies $I_1 = I_2$ and therefore we conclude $(i) \iff (ii)$ holds.

Now we consider the proof of $(ii) \iff (iii)$. Suppose that $(ii)$ holds. Let $M$ be an arbitrary $F$-martingale, then we get for any $t \leq s$

$$\mathbb{E}^P[M_s|G_t] = \mathbb{E}^P[M_s|F_t] = M_t.$$ 

Hence, $M$ is also a $G$-martingale. Now suppose $(iii)$ holds, then for any fixed $t \leq s$ and arbitrary set $A \in \mathcal{G}_\infty$, we consider the $F$-martingale $M_t = \mathbb{P}[A|\mathcal{F}_t], t \in \mathbb{R}_+$. Now we get

$$\mathbb{P}\{A|\mathcal{F}_t\} = M_t = \mathbb{P}\{A|\mathcal{G}_t\},$$

158
B.1. The \((\mathbb{F}, \mathbb{G})\)-Martingale Hazard Process

since \(M\) is also a \(\mathbb{G}\)-martingale. Now it is easy to see that (ii) holds.

Under the assumptions of Condition 4 and Condition 5, we are able to specify the \(\mathbb{F}\)-martingale hazard process.

**Proposition B.3.** Assume that \(F\) is an increasing, \(\mathbb{F}\)-predictable process. Then the process \(\Lambda\) given by

\[
\Lambda_t = \int_{[0,t]} \frac{dF_u}{1 - F_u} 
\]  

(B.1)

is the \(\mathbb{F}\)-martingale hazard process of \(\tau\).

**Proof.** We need to check that \(H_t - \Lambda_{t\wedge \tau}\) is a \(\mathbb{G}\)-martingale. We get

\[
\mathbb{E}^\mathbb{P}\left[ H_s - H_t | \mathcal{G}_t \right] = \mathbb{P}\left\{ t < \tau \leq s | \mathcal{G}_t \right\} 
= \mathbb{P}\left\{ t < \tau \leq s | \mathcal{F}_t \right\} \mathbb{P}\left\{ \tau > t | \mathcal{F}_t \right\} \quad \text{(by formula (A.3))}
= \mathbb{P}\left\{ \tau > t \right\} \left( \mathbb{E}^\mathbb{P}\left[ 1_{\{t \leq s\}} - 1_{\{t \leq \tau\}} | \mathcal{F}_t \right] \right) / 1 - F_t
= \mathbb{P}\left\{ \tau > t \right\} \left( \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}\left[ 1_{\{t \leq s\}} | \mathcal{F}_s \right] | \mathcal{F}_t \right] - \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}\left[ 1_{\{t \leq \tau\}} | \mathcal{F}_t \right] | \mathcal{F}_t \right] \right) / 1 - F_t
= \mathbb{P}\left\{ \tau > t \right\} \left( \mathbb{E}^\mathbb{P}\left[ F_s | \mathcal{F}_t \right] - F_t \right) / 1 - F_t.
\]

We, furthermore, get

\[
\mathbb{E}^\mathbb{P}\left[ \Lambda_{s\wedge \tau} - \Lambda_{t\wedge \tau} | \mathcal{G}_t \right] = \mathbb{E}^\mathbb{P}\left[ \int_{[t \wedge \tau, s\wedge \tau]} \frac{dF_u}{1 - F_u} | \mathcal{G}_t \right] = \mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y | \mathcal{G}_t \right],
\]

where \(Y = \int_{[t \wedge \tau, s\wedge \tau]} \frac{dF_u}{1 - F_u} = 1_{\{\tau > t\}}\). By formula (A.5) we get

\[
\mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y | \mathcal{G}_t \right] = \frac{\mathbb{E}^\mathbb{P}\left[ 1_{\{\tau > t\}} Y | \mathcal{F}_t \right]}{\mathbb{P}\left\{ \tau > t | \mathcal{F}_t \right\}}.
\]
Chapter B. Martingale Hazard Process of a Random Time

If we can now show that \( \mathbb{E}^\mathbb{P} \left[ 1_{\{\tau > t\}} Y \mid \mathcal{F}_t \right] = \mathbb{E}^\mathbb{P} \left[ F_s - F_t \mid \mathcal{F}_t \right] \), then we are done. We get

\[
\mathbb{E}^\mathbb{P} \left[ 1_{\{\tau > t\}} Y \mid \mathcal{F}_t \right] = \mathbb{E}^\mathbb{P} \left[ \int_{[t,s]} \frac{dF_u}{1 - F_u^-} \mid \mathcal{F}_t \right]
\]

\[
= \mathbb{E}^\mathbb{P} \left[ 1_{\{\tau > s\}} \int_{[t,s]} \frac{dF_u}{1 - F_u^-} + 1_{\{t < \tau \leq s\}} \int_{[t,s]} \frac{dF_u}{1 - F_u^-} \mid \mathcal{F}_t \right]
\]

\[
= \mathbb{E}^\mathbb{P} \left[ \mathbb{E}^\mathbb{P} \left[ 1_{\{\tau > s\}} \int_{[t,s]} \frac{dF_u}{1 - F_u^-} \mid \mathcal{F}_s \right] + 1_{\{t < \tau \leq s\}} \int_{[t,u]} \frac{dF_u}{1 - F_u^-} \mid \mathcal{F}_t \right]
\]

\[
= \mathbb{E}^\mathbb{P} \left[ (1 - F_s) \int_{[t,s]} \frac{dF_u}{1 - F_u^-} + \int_{[t,u]} \frac{dF_u}{1 - F_u^-} \mid \mathcal{F}_t \right]
\]

\[
= \mathbb{E}^\mathbb{P} \left[ (1 - F_s)(\Lambda_s - \Lambda_t) + \int_{[t,s]} (\Lambda_u - \Lambda_t) dF_u \mid \mathcal{F}_t \right],
\]

where the fourth equality follows from (A.11), since we can take \( Z_s = \int_{[t,s]} \frac{dF_u}{1 - F_u^-} \), which is \( \mathbb{F} \)-predictable, since \( F \) is \( \mathbb{F} \)-predictable. Since \( F \) is also increasing, we get that \( \tilde{F} \) and \( \Lambda \) are processes of finite variation and therefore we get by the product rule

\[
\int_{[t,s]} \Lambda_u dF_u = \Lambda_s F_s - \Lambda_t F_t - \int_{[t,s]} F_u^- d\Lambda_u.
\]

And we also get that \( \int_{[t,s]} F_u^- d\Lambda_u = \Lambda_s - \Lambda_t - F_s + F_t \), since

\[
\Lambda_s - \Lambda_t - F_s + F_t = \Lambda_s - \Lambda_t - \int_{[t,s]} dF_u
\]

\[
= \int_{[t,s]} \frac{1}{1 - F_u^-} dF_u - \int_{[t,s]} dF_u
\]

\[
= \int_{[t,s]} \left( \frac{1}{1 - F_u^-} - 1 \right) dF_u
\]

\[
= \int_{[t,s]} \frac{F_u^-}{1 - F_u^-} dF_u
\]

\[
= \int_{[t,s]} F_u^- d\Lambda_u.
\]

Therefore, we get
\[ E^F \left[ \mathbb{1}_{\{\tau > t\}} Y \right| F_t] = E^F \left[ (1 - F_s)(\Lambda_s - \Lambda_t) + \int_{[t,s]} \Lambda_u dF_u - \int_{[t,s]} \Lambda_t dF_u \left| F_t \right. \right] \\
= E^F \left[ (1 - F_s)(\Lambda_s - \Lambda_t) + \int_{[t,s]} \Lambda_u dF_u - \Lambda_t(F_s - F_t) \left| F_t \right. \right] \\
= E^F [F_s - F_t|F_t]. \]

This concludes the proof. \(\square\)

It can be shown that the \(F\)-martingale hazard process \(\Lambda\) of \(\tau\) is unique up to time \(\tau\) (Bielecki & Rutkowski, 2002 [6]). This result follows from the fact that the process \(H\) (or equivalently \(\tau\)) has a unique (by the Doob-Meyer decomposition) \(G\)-compensator, which is defined as follows:

**Definition B.4.** A process \(A\) is a \(G\)-compensator of \(\tau\) if and only if \(A\) is a \(G\)-predictable, right-continuous, increasing process with \(A_0 = 0\) and if the process \(H - A\) is a \(G\)-martingale.

It can be shown that the notions of a \(G\)-compensator \(A\) and an \(F\)-martingale hazard process \(\Lambda\) of \(\tau\) satisfy the relation \(A_t = \Lambda_{t\wedge \tau}\), and, therefore, by the uniqueness of the \(G\)-compensator, we also have a uniqueness result for the \(F\)-martingale hazard process.

### B.2 Relationships Between \(\Gamma\) and \(\Lambda\)

The following results show the relationship between the hazard process and the martingale hazard process.

**Proposition B.5.** Let the assumptions of Conditions 4 and 5 hold.

(i) If the increasing process \(F\) is \(F\)-predictable, but \(F\) is not continuous, then the \(F\)-martingale hazard process \(\Lambda\) is also a discontinuous process and we have

\[ e^{-\Gamma_t} = e^{-\Lambda^c_t} \prod_{0 < u \leq t} (1 - \Delta \Lambda_u), \]

where \(\Lambda^c\) is the continuous component of \(\Lambda\).

(ii) If the increasing process \(F\) is continuous, then the \(F\)-martingale hazard process \(\Lambda\) is also continuous an \(\Gamma_t = \Lambda_t = -\ln(1 - F_t), \quad \forall \ t \in \mathbb{R}^+. \)
If, in addition, the process $\Lambda = \Gamma$ is absolutely continuous then for an integrable $\mathcal{F}_s$-measurable random variable $Y$ we get

$$
\mathbb{E}^P\left[1_{\{\tau>s\}}Y \bigg| \mathcal{G}_t\right] = 1_{\{\tau>t\}}\mathbb{E}^P\left[Ye^{-\int_t^s \lambda_u \, du} \bigg| \mathcal{F}_t\right].
$$

Proof. All conditions of Proposition B.3 are satisfied and therefore we know that the $\mathbb{F}$-martingale hazard process $\Lambda$ is given by

$$
\Lambda_t = \int_{[0,t]} \frac{dF_u}{1 - F_u}. 
$$

Now we note that we can write

$$
G_t = 1 - F_t = 1 - \int_{[0,t]} dF_u 
= 1 - \int_{[0,t]} \frac{1 - F_u}{1 - F_u} dF_u 
= 1 - \int_{[0,t]} G_u - d\Lambda_u.
$$

We recognize the SDE for which the solution is given by the Doléans exponential (see, e.g., Cont & Tankov (2004) [15]), and, therefore, we conclude that

$$
G_t = e^{-\Gamma_t} = e^{-\Lambda_t^t} \prod_{0<u\leq t} (1 - \Delta\Lambda_u). 
$$

Note that $\Gamma$ is discontinuous if and only if $F$ is discontinuous, and, therefore, the first result is proved. If, in addition, $F$ is continuous, then it follows from $(i)$, that $\Lambda$ is also continuous. For the last expression of part $(ii)$, we refer to formula (A.7).

The following result is a consequence of Propositions A.8 and B.5.

Corollary B.6. Let the assumptions of Conditions 4 and 5 hold and assume furthermore that $F$ is a continuous, increasing process. Let $Y = h(\tau)$ for some bounded, continuous function $h : \mathbb{R}_+ \to \mathbb{R}$. Then

$$
\mathbb{E}^P\left[Y \bigg| \mathcal{G}_t\right] = 1_{\{\tau\leq t\}}h(\tau) + 1_{\{\tau>t\}}\mathbb{E}^P\left[\int_t^\infty h(u)e^{\Lambda_u \Delta\lambda u} \, du \bigg| \mathcal{F}_t\right].
$$
Let $Z$ be a bounded, $\mathbb{F}$-predictable process. Then for any $t \leq s$

$$
\mathbb{E}^\mathbb{F} \left[ Z_t 1_{\{t<\tau \leq s\}} \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^\mathbb{F} \left[ \int_t^s Z_u e^{\Lambda_t - \Lambda_u} d\Lambda_u \mid \mathcal{F}_t \right].
$$

Proof. We note that, since $F$ is a continuous and increasing process, we have by Proposition B.5 part (ii) that $\lambda = \Gamma$. Furthermore, since $F$ is continuous, we get $dF_t = e^{-\Gamma_t} d\Gamma_t$. Using the results of Proposition A.8, we have the results.

We saw in Proposition B.5 that continuity of $\Gamma$ implies $\Gamma = \Lambda$. We could also ask whether continuity of $\Lambda$ implies $\Lambda = \Gamma$, but it seems that we cannot give a full result on this question. The following, partial, answer can be proved (see Bielecki and Rutkowski (2002) [6]):

**Proposition B.7.** Let the assumptions of Conditions 4 and 5 hold and assume furthermore that the filtration $\mathbb{F}$ supports only continuous martingales. If the $\mathbb{F}$-martingale hazard process $\Lambda$ is continuous, then the hazard process $\Gamma$ is also continuous and $\Lambda = \Gamma$.

The following result gives some useful expressions in terms of the martingale hazard process.

**Proposition B.8.** Let the assumptions of Conditions 4 and 5 hold and assume furthermore that $\mathbb{F}$-martingale hazard process $\Lambda$ of the random time $\tau$ is continuous. For a fixed $s > 0$, let $Y$ be an $\mathcal{F}_s$-measurable, integrable random variable.

(i) If the (right-continuous) process $V$, given by

$$
V_t = \mathbb{E}^\mathbb{F} \left[ Ye^{\Lambda_t - \Lambda_s} \mid \mathcal{F}_t \right], \quad \forall \ t \in [0,s],
$$

is continuous at $\tau$ (so $\Delta V_{s\wedge \tau} = 0$), then for any $t < s$ we have

$$
\mathbb{E}^\mathbb{F} \left[ 1_{\{\tau > s\}} Y \mid \mathcal{G}_t \right] = 1_{\{\tau > t\}} \mathbb{E}^\mathbb{F} \left[ Ye^{\Lambda_t - \Lambda_s} \mid \mathcal{F}_t \right].
$$

(ii) If the process $V$, given by

$$
V_t = \mathbb{E}^\mathbb{F} \left[ e^{\Lambda_t - \Lambda_s} \mid \mathcal{F}_t \right], \quad \forall \ t \in [0,s],
$$

is continuous at $\tau$, then for any $t \leq s$ we have

$$
\mathbb{P} \{ \tau > s \mid \mathcal{G}_t \} = 1_{\{\tau > t\}} \mathbb{E}^\mathbb{F} \left[ e^{\Lambda_t - \Lambda_s} \mid \mathcal{F}_t \right].
$$
Appendix C

Derivations of the Discount Factors

In this Appendix, we will give the explicit formulas for the discount factors and expectation terms of formulas (5.6) and (5.7). In these formulas, we still have to compute the following expressions:

\[
\begin{align*}
\mathbb{E}\left[\bar{P}(t, T_i) \mid \mathcal{F}_t\right] &= \mathbb{E}\left[e^{-\int_t^{T_i} \lambda_s ds} \mid \mathcal{F}_t\right], \\
\mathbb{E}\left[\bar{P}(t, T_i) \bar{L}(t, T_i) \mid \mathcal{F}_t\right] &= \mathbb{E}\left[e^{-\int_t^{T_i} \lambda_s ds} e^{-\int_t^{T_i} \gamma_l ds} \mid \mathcal{F}_t\right],
\end{align*}
\]

where \(l \in \{\text{bid, ask}\}\) and \(\mathcal{F}_t\) is the filtration generated by the Brownian motions \(W = (W_x, W_{y^{\text{bid}}}, W_{y^{\text{ask}}})\).

As stated in chapter 5, we consider two different types of processes for modeling the pure liquidity intensities. We either model the pure liquidity intensities as Gaussian (or Arithmetic Brownian motions without drift) processes, so

\[
dy^l(t) = \sigma^l dW_{y^l}(t), \quad l \in \{\text{bid, ask}\},
\]

or we model them as Ornstein-Uhlenbeck processes with the same mean-reversion speed parameter and zero mean-reversion level parameter:

\[
dy^l(t) = -\eta y^l(t)dt + \sigma^l dW_{y^l}(t), \quad l \in \{\text{bid, ask}\}.
\]

We will discuss both models separately, starting with the Gaussian liquidity intensity set-up.
Chapter C. Derivations of the Discount Factors

C.1 Model with Gaussian Processes for Liquidity Intensities

In this model set-up, we have the following dependence structure:

\[
\begin{pmatrix}
  d\lambda(t) \\
  d\gamma^{bid}(t) \\
  d\gamma^{ask}(t)
\end{pmatrix} =
\begin{pmatrix}
  1 & g_{bid} & g_{ask} \\
  f_{bid} & 1 & \omega_{ask,bid} \\
  f_{ask} & \omega_{bid,ask} & 1
\end{pmatrix}
\begin{pmatrix}
  dx(t) \\
  dy^{bid}(t) \\
  dy^{ask}(t)
\end{pmatrix},
\]

where

\[
\begin{align*}
  dx(t) &= (\alpha - \beta x(t))dt + \sigma \sqrt{x(t)}dW_x(t), \\
  dy^{bid}(t) &= \sigma^{bid}dW_y^{bid}(t), \\
  dy^{ask}(t) &= \sigma^{ask}dW_y^{ask}(t),
\end{align*}
\]

with \( W_x, W_y^{bid/ask} \) independent \( \mathbb{Q} \)-Brownian motions. By Lemma 5.1, we get for (C.1) that

\[
\mathbb{E}\left[e^{-\int_t^{T_i} \lambda_s \, ds} \, \bigg| \mathcal{F}_t \right] =
\mathbb{E}\left[e^{-\int_t^{T_i} (x(s) + g_{bid}y^{bid}(s) + g_{ask}y^{ask}(s)) \, ds} \, \bigg| \mathcal{F}_t \right]
\mathbb{E}\left[e^{-\int_t^{T_i} g_{bid}y^{bid}(s) \, ds} \, \bigg| \mathcal{F}_t \right]
\mathbb{E}\left[e^{-\int_t^{T_i} g_{ask}y^{ask}(s) \, ds} \, \bigg| \mathcal{F}_t \right].
\]

Each of these expectations can be computed analytically, since we recognize the so-called bond price formulas for known analytically tractable affine processes (see Proposition 4.5). The explicit calculations of these bond price formulas are given in Appendices D and F. Here, we will just state the results. We get for the first expectation

\[
\mathbb{E}\left[e^{-\int_t^{T_i} x(s) \, ds} \, \bigg| \mathcal{F}_t \right] = e^{-A(t,T_i) - B(t,T_i)x(t)},
\]

with

\[
\begin{align*}
  A(t,T_i) &= -\frac{2\gamma}{\sigma^2} \ln \left( \frac{2\gamma e^{(\gamma+\beta)(T_i-t)}}{2\gamma + (\gamma + \beta) (e^{\gamma(T_i-t)} - 1)} \right), \\
  B(t,T_i) &= \frac{2}{\beta + \gamma} \left( e^{\gamma(T_i-t)} - 1 \right), \\
  \gamma &= \sqrt{\beta^2 + 2\sigma^2}.
\end{align*}
\]
C.1. Model with Gaussian Processes for Liquidity Intensities

The second and third expectations are given by

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} g(t') y(t') dt'} \mid \mathcal{F}_t \right] = e^{-A(t, T_i) - B(t, T_i) g(t)}.
\]

where

\[
A(t, T_i) = \frac{(g(t) \sigma)^2}{6} (T_i - t)^3,
\]

\[
B(t, T_i) = (T_i - t),
\]

with \( l \in \{\text{bid, ask}\} \).

The conditional expectation given by (C.2) can, again by Lemma 5.1, be written as the product of independent conditional expectations:

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} \lambda_s dt} e^{-\int_t^{T_i} \gamma_s^l dt} \mid \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^{T_i} (1+f_l) x(s) ds} \mid \mathcal{F}_t \right] \cdot \mathbb{E} \left[ e^{-\int_t^{T_i} (1+g_l) y^l(s) ds} \mid \mathcal{F}_t \right] \cdot \mathbb{E} \left[ e^{-\int_t^{T_i} (g_{k \neq l} + \omega_{k \neq l, l}) y^l(s) ds} \mid \mathcal{F}_t \right],
\]

with \( l, k \in \{\text{bid, ask}\} \). The first expectation can be written as

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} (1+f_l) x(s) ds} \mid \mathcal{F}_t \right] = e^{-A(t, T_i) - B(t, T_i) (1+f_l) x(t)},
\]

where

\[
A(t, T_i) = -\frac{2\alpha}{\sigma^2} \ln \left( \frac{2 e^{(\gamma + \beta)(T_i - t)}}{2 \gamma + (\gamma + \beta) (e^{(T_i - t)} - 1)} \right),
\]

\[
B(t, T_i) = \frac{2 (e^{(T_i - t)} - 1)}{(\beta + \gamma) (e^{(T_i - t)} - 1) + 2 \gamma},
\]

\[
\gamma = \sqrt{\beta^2 + 2(1 + f_l) \sigma^2}.
\]

The second expectation is given by
\[ \mathbb{E} \left[ e^{-\int_t^{T_i} (1+g_l) y_l(s) ds} \right] = e^{-A(t,T_i) - B(t,T_i)(1+g_l)y_l(t)}, \]

where

\[
A(t, T_i) = -\frac{(1 + g_l)^2 (\sigma_l^2) (T_i - t)^3}{6}, \\
B(t, T_i) = (T_i - t). 
\]

Similarly, the third expectation can be written as

\[ \mathbb{E} \left[ e^{-\int_t^{T_i} (g_{k \neq l} + \omega_{k \neq l}) y_{k \neq l}(s) ds} \right] = e^{-A(t,T_i) - B(t,T_i)(g_{k \neq l} + \omega_{k \neq l}) y_{k \neq l}(t)}, \]

where

\[
A(t, T_i) = -\frac{(g_{k \neq l} + \omega_{k \neq l})^2 (\sigma_{k \neq l}^2) (T_i - t)^3}{6}, \\
B(t, T_i) = (T_i - t). 
\]

### C.2 Model with Ornstein-Uhlenbeck Processes for Liquidity Intensities

The structure of the model with Ornstein-Uhlenbeck processes for the liquidity intensities is completely similar to the structure of the model with Gaussian processes for the liquidity intensities. The only difference is that expectations containing \( y_l \) are different. In this model set-up, we have the following dependence structure:

\[
\begin{pmatrix}
\frac{d\lambda(t)}{dt} \\
\frac{d\gamma^{bid}(t)}{dt} \\
\frac{d\gamma^{ask}(t)}{dt}
\end{pmatrix} =
\begin{pmatrix}
1 & g_{bid} & g_{ask} \\
1 & \omega_{bid,bid} & \omega_{ask,bid} \\
\omega_{bid,ask} & 1 & 1
\end{pmatrix}
\begin{pmatrix}
\frac{dx(t)}{dt} \\
\frac{dy^{bid}(t)}{dt} \\
\frac{dy^{ask}(t)}{dt}
\end{pmatrix},
\]

where

168
C.2. Model with Ornstein-Uhlenbeck Processes for Liquidity Intensities

\[
\begin{align*}
\text{d}x(t) &= (\alpha - \beta x(t)) \text{d}t + \sigma \sqrt{x(t)} \text{d}W_x(t), \\
\text{d}y^{\text{bid}}(t) &= -\eta y^{\text{bid}}(t) \text{d}t + \sigma^{\text{bid}} \text{d}W_y^{\text{bid}}(t), \\
\text{d}y^{\text{ask}}(t) &= -\eta y^{\text{ask}}(t) \text{d}t + \sigma^{\text{ask}} \text{d}W_y^{\text{ask}}(t),
\end{align*}
\]

with \(W_x, W_{y^{\text{bid/ask}}}\) independent \(Q\)-Brownian motions. We have to calculate the same expectations in this model set-up, so we will just state the results here. We get for (C.1):

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} \lambda_x(s) \text{d}s} \mid \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^{T_i} x(s) \text{d}s} \mid \mathcal{F}_t \right] \mathbb{E} \left[ e^{-\int_t^{T_i} g_{\text{bid}} y^{\text{bid}}(s) \text{d}s} \mid \mathcal{F}_t \right] \mathbb{E} \left[ e^{-\int_t^{T_i} g_{\text{ask}} y^{\text{ask}}(s) \text{d}s} \mid \mathcal{F}_t \right],
\]

where the first equation is the same as in the first model set-up, and

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} g_i y'(s) \text{d}s} \mid \mathcal{F}_t \right] = e^{-A(t,T_i) - B(t,T_i) g_i y'(t)},
\]

where

\[
A(t, T_i) = \frac{(g_i \sigma_i)^2}{4 (g_i \eta)^2} \left( e^{-2g_i \eta(T_i-t)} - 4 e^{-g_i \eta(T_i-t)} - 2g_i \eta(T_i-t) + 3 \right),
\]

\[
B(t, T_i) = \frac{1}{g_i \eta} \left( 1 - e^{-g_i \eta(T_i-t)} \right),
\]

with \(l \in \{\text{bid, ask}\}\). Similar to the first model set-up, we can write for (C.2):

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} \chi_{l,k} \text{d}s} e^{-\int_t^{T_i} \gamma_{l,k}' \text{d}s} \mid \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^{T_i} (1+f_l) x(s) \text{d}s} \mid \mathcal{F}_t \right] \cdot \mathbb{E} \left[ e^{-\int_t^{T_i} (1+g_l) y'(s) \text{d}s} \mid \mathcal{F}_t \right] \cdot \mathbb{E} \left[ e^{-\int_t^{T_i} (g_{k,l} + \omega_{k,l}) y^{k\neq l}(s) \text{d}s} \mid \mathcal{F}_t \right],
\]

with \(l, k \in \{\text{bid, ask}\}\). The first of these expectations is the same as in the first model set-up. The second expectation is given by

\[
\mathbb{E} \left[ e^{-\int_t^{T_i} (1+g_l) y'(s) \text{d}s} \mid \mathcal{F}_t \right] = e^{-A(t,T_i) - B(t,T_i) (1+g_l) y'(t)},
\]

169
Chapter C. Derivations of the Discount Factors

where

\[ A(t, T_i) = \frac{((1 + g_l)\sigma_l)^2}{4((1 + g_l)\eta)^3} \left( e^{-2(1+g_l)\eta(T_i-t)} - 4e^{-(1+g_l)\eta(T_i-t)} - 2(1 + g_l)\eta(T_i - t) + 3 \right), \]

\[ B(t, T_i) = \frac{1}{(1 + g_l)\eta} \left( 1 - e^{-(1+g_l)\eta(T_i-t)} \right), \]

with \( l \in \{\text{bid, ask}\} \). The last expectation is given by

\[ \mathbb{E} \left[ e^{-\int_t^{T_i}(g_{k\neq l} + \omega_{k\neq l})y_{k\neq l}(s)ds} \mid \mathcal{F}_t \right] = e^{-A(t, T_i) - B(t, T_i)(g_{k\neq l} + \omega_{k\neq l})y_{k\neq l}(t)} , \]

where

\[ A(t, T_i) = \frac{((g_{k\neq l} + \omega_{k\neq l})\sigma_l)^2}{4((g_{k\neq l} + \omega_{k\neq l})\eta)^3} \left( e^{-2(g_{k\neq l} + \omega_{k\neq l})\eta(T_i-t)} - 4e^{-(g_{k\neq l} + \omega_{k\neq l})\eta(T_i-t)} - 2(g_{k\neq l} + \omega_{k\neq l})\eta(T_i - t) + 3 \right), \]

\[ B(t, T_i) = \frac{1}{(g_{k\neq l} + \omega_{k\neq l})\eta} \left( 1 - e^{-(g_{k\neq l} + \omega_{k\neq l})\eta(T_i-t)} \right), \]

with \( l, k \in \{\text{bid, ask}\} \).
Appendix D

Arithmetic Brownian Motion

In this Appendix, we will give some results on the Arithmetic Brownian motion, which is given by

\[ dy_t = \mu dt + \sigma dW_t. \]  \hspace{1cm} (D.1)

We will also refer to this process as the Gaussian process (with drift). In the first model option of the credit-liquidity model, which is presented in chapter 5, we assume that the pure liquidity intensities, \( y_{bid} \) and \( y_{ask} \), follow Arithmetic Brownian motions without drifts under \( Q \). That is,

\[ dy_{bid}(t) = \sigma_{bid} dW_{y_{bid}}(t), \]
\[ dy_{ask}(t) = \sigma_{ask} dW_{y_{ask}}(t). \]

In order to compute the model-implied bid and ask premia, we are interested in expressions of the form

\[ \mathbb{E}^Q \left[ e^{-\int_t^T c y_{bid/ask}(s) ds} \bigg| \mathcal{F}_t \right], \]  \hspace{1cm} (D.2)

where \( c \) is some constant. If we would interpret the processes \( cy \) as the short-rate interest rate process, the above expression would be exactly the formula of a zero-coupon bond price. Therefore, we will refer to expression (D.2) as the bond price under an Arithmetic Brownian
motion. We note that an Arithmetic Brownian motion multiplied by a constant $c$ is again an Arithmetic Brownian motion, since

$$
d(ce^t) = cde^t = c\mu dt + c\sigma dW_t = \tilde{\mu} dt + \tilde{\sigma} dW_t.
$$

In the next section, we will derive an explicit formula for the bond price under an Arithmetic Brownian motion. After that, we will solve the SDE (D.1), since we can deduce (conditional) distributional properties from this, which will be needed for the derivation of maximum likelihood estimators of the process parameters. The computations of these maximum likelihood estimators will be the topic of the last section in this Appendix. The results in this Appendix will be given for the general Arithmetic Brownian motion with drift, since the results for the arithmetic Brownian motion without drift follow as special cases from these results.

## D.1 Bond Price under Arithmetic Brownian Motion

Consider the Arithmetic Brownian motion given in (D.1). Since this process falls in the class of affine processes, we have by Proposition 4.5 that

$$
\mathbb{E}^Q\left[e^{-\int_t^T c\mu ds} \bigg| \mathcal{F}_t\right] = e^{-A(t,T)-B(t,T)c^t},
$$

where $A(t,T)$ and $B(t,T)$ satisfy the following system of ordinary differential equations:

$$
\begin{align*}
\partial_t A(t,T) &= \frac{1}{2}(c\sigma)^2 B^2(t,T) - c\mu B(t,T), \quad A(T,T) = 0, \\
\partial_t B(t,T) &= -1, \quad B(T,T) = 0.
\end{align*}
$$

Clearly, a solution for $B(t,T)$ is given by

$$
B(t,T) = (T - t).
$$

Plugging this into the differential equation of $A(t,T)$ gives
D.2. Explicit Solution of the Arithmetic Brownian Motion SDE

\[ \partial_t A(t, T) = \frac{1}{2}(c\sigma)^2 (T - t)^2 - c\mu (T - t), \quad A(T, T) = 0. \]

Integrating both sides gives

\[ A(t, T) = \frac{(c\sigma)^2}{6} (T - t)^3 + \frac{c\mu}{2} (T - t)^2. \]

We thus conclude that

\[ \mathbb{E}^Q \left[ e^{-\int_t^T c\mu \, ds} \middle| \mathcal{F}_t \right] = e^{-A(t, T) - B(t, T)c\mu}, \]

with

\[ A(t, T) = \frac{(c\sigma)^2}{6} (T - t)^3 + \frac{c\mu}{2} (T - t)^2, \]
\[ B(t, T) = (T - t). \]

**D.2 Explicit Solution of the Arithmetic Brownian Motion SDE**

The explicit solution of (D.1) can easily be found by integrating both sides:

\[ y_t = y_0 + \int_0^t \mu \, ds + \int_0^t \sigma \, dW_s. \]  \hspace{1cm} (D.3)

We see that \( y_t \) is normally distributed with mean \( y_0 + \mu t \) and variance \( \sigma^2 t \). Equivalently, when we consider the distribution of \( y_t \), conditional on \( \mathcal{F}_s \) (\( s < t \)), we get that \( y_t \) is normally distributed with mean \( y_s + \mu (t - s) \) and variance \( \sigma^2 (t - s) \). This follows from the following Lemma:

**Lemma D.1.** Let \( \sigma(t) \) be a given deterministic function of time and define the process \( X \) by

\[ X_t = \int_0^t \sigma(s) \, dW_s. \]

Under suitable integrability conditions, \( X_t \) has a normal distribution with zero mean and variance given by \( \text{Var}(X_t) = \int_0^t \sigma^2(s) \, ds \).
Proof. Consider the the following (twice continuously differentiable function) \( f(x) = e^{iu x} \). Then we have \( f_x = iue^{iu x} \) and \( f_{xx} = -u^2e^{iu x} \) (where \( f_x \) and \( f_{xx} \) denote the first and second derivatives of \( f \) w.r.t. \( x \) respectively). By applying Itô’s rule we get

\[
df(X_t) = iue^{iuX_t}dX_t - \frac{u^2}{2}e^{iuX_t}d\langle X \rangle_t
\]

Integrating both sides gives

\[
e^{iuX_t} = 1 + iu \int_0^t f(X_s)\sigma(s)dW_s - \frac{u^2}{2} \int_0^t f(X_s)\sigma^2(s)ds.
\]

Taking expectations on both sides will make the part with the stochastic integral disappear, since \( \mathbb{E} \left[ \int_0^t f(X_s)\sigma(s)dW_s \right] = 0 \). Furthermore, taking the expectation inside the \( ds \)-integral gives the following:

\[
\mathbb{E}[e^{iuX_t}] = 1 - \frac{u^2}{2} \int_0^t \mathbb{E}[e^{iuX_s}]\sigma^2(s)ds
\]

\[
M(t) = 1 - \frac{u^2}{2} \int_0^t M(s)\sigma^2(s)ds.
\]

Differentiating both sides gives us the following ordinary differential equation:

\[
\dot{M}(t) = -\frac{u^2}{2} M(t)\sigma^2(t).
\]

Clearly, the solution to this ODE is given by \( M(t) = e^{-\frac{u^2}{2} \int_0^t \sigma^2(s)ds} \). From this we conclude that the characteristic function of \( X_t \) is exactly the characteristic function of a normally distributed random variable with mean zero and variance \( \int_0^t \sigma^2(s)ds \).

D.3 Maximum Likelihood Estimators for Arithmetic Brownian Motion

Let us again consider the Arithmetic Brownian motion

\[174\]
\[ dy_t = \mu dt + \sigma dW_t. \]

By the discussion above, we know that, conditional on \( F_s \), with \( s \leq t \), \( y_t \) is normally distributed with mean \( y_s + \mu (t-s) \) and variance \( \sigma^2 (t-s) \). We are now able to derive the maximum likelihood estimators of the process parameters using this distributional property. We will refer to this as the exact maximum likelihood estimators, as they are based on the exact distribution of the process.

D.3.1 Exact Maximum Likelihood

Suppose that we have \((N + 1)\) observations, \( y_i \), for \( i = 0, 1, \ldots, N \) with equidistant time-differences \( t_{i+1} - t_i = \Delta \forall i = 0, 1, \ldots, N \). We get that the conditional probability density function of an observation \( y_i \) given a previous observation \( y_{i-1} \) is given by

\[
 f(y_i | y_{i-1}, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp\left(-\frac{(y_i - y_{i-1} - \mu\Delta)^2}{2\sigma^2\Delta}\right).
\]

The log-likelihood function of the complete sample is given by

\[
 L(\mu, \sigma^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2\Delta) - \sum_{i=1}^{N} \frac{(y_i - y_{i-1} - \mu\Delta)^2}{2\sigma^2\Delta}.
\]

We find the maximum likelihood estimators by taking the partial derivatives and equating them to zero and then solving them for the parameters. We get the following partial derivatives:

\[
 \frac{\partial L(\mu, \sigma)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{N} (y_i - y_{i-1} - \mu\Delta), \tag{D.4}
\]

\[
 \frac{\partial L(\mu, \sigma)}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \sum_{i=1}^{N} \frac{(y_i - y_{i-1} - \mu\Delta)^2}{2\sigma^4\Delta}. \tag{D.5}
\]

The partial derivatives evaluated in the maximum likelihood estimators \( \hat{\mu} \) and \( \hat{\sigma} \) make the above equations equal to zero. Setting equation (D.4) equal to zero by evaluating in the maximum likelihood estimators \( \hat{\mu} \) and \( \hat{\sigma} \) gives

\[
 \frac{\partial L(\hat{\mu}, \hat{\sigma})}{\partial \mu} = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^{N} (y_i - y_{i-1}) - \frac{N\hat{\mu}\Delta}{\hat{\sigma}^2} = 0.
\]
Chapter D. Arithmetic Brownian Motion

Solving this for $\hat{\mu}$ gives

$$\hat{\mu} = \frac{1}{N\Delta} \sum_{i=1}^{N} (y_i - y_{i-1}).$$ (D.6)

Similarly, setting equation (D.5) equal to zero by evaluating in the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ gives

$$\frac{\partial L(\mu, \sigma)}{\partial \sigma^2} = -\frac{N}{2\hat{\sigma}^2} + \sum_{i=1}^{N} \frac{(y_i - y_{i-1} - \hat{\mu} \Delta)^2}{2\hat{\sigma}^4 \Delta} = 0.$$

Solving for $\hat{\sigma}^2$ gives

$$\hat{\sigma}^2 = \frac{1}{N\Delta} \sum_{i=1}^{N} (y_i - y_{i-1})^2 - \frac{2\hat{\mu}}{N} \sum_{i=1}^{N} (y_i - y_{i-1}) + \hat{\mu}^2 \Delta.$$

D.3.2 Continuous Record Maximum Likelihood

The Continuous Record Likelihood method is introduced in chapter 8. This method only works if the diffusion function is known and does not depend on the unknown parameters. Therefore, one can only estimate the drift parameters using this method.

In chapter 8, it was derived that the Continuous Record log-Likelihood function of a general continuous diffusion process

$$dX_t = \mu(\theta, X_t)dt + \sigma(\theta, X_t)dW_t, \quad X_0 = x_0,$$ (D.7)

is given by

$$L_T(\theta) = \int_0^T \frac{\mu(\theta, X_t)}{\sigma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(X_t)} dt.$$

In the case of the Arithmetic Brownian motion, we have that $\mu(\theta, y_t) = \mu$ and $\sigma(y_t) = \sigma$, and we thus assume $\sigma$ to be known. The log-likelihood function, therefore, becomes
D.3. Maximum Likelihood Estimators for Arithmetic Brownian Motion

\[ L_T(\mu) = \int_0^T \frac{\mu}{\sigma^2} dy_t - \frac{1}{2} \int_0^T \frac{\mu^2}{\sigma^2} dt. \]

Setting the derivative with respect to \( \mu \) equal to zero and solving for \( \mu \) gives the maximum likelihood estimator of \( \mu \). We get

\[ \frac{\partial L_T(\mu)}{\partial \mu} = \frac{1}{\sigma^2} \int_0^T dy_t - \frac{1}{\sigma^2} \int_0^T \mu dt. \]

We replace the Lebesgue and Itô integrals by Riemann-Itô sums, and we get:

\[ \frac{\partial L_T(\mu)}{\partial \mu} \approx \frac{1}{\sigma^2} \sum_{i=1}^N (y_i \Delta - y_{(i-1)} \Delta) - \frac{1}{\sigma^2} \sum_{i=1}^N \mu \Delta. \]

Equating this expression to zero and solving for \( \mu \) gives the Continuous Record maximum likelihood estimator \( \hat{\mu} \):

\[ \hat{\mu} = \frac{1}{N \Delta} \sum_{i=1}^N (y_i \Delta - y_{(i-1)} \Delta). \]  \hspace{1cm} (D.8)

We see that the Continuous Record maximum likelihood estimator of \( \mu \) is equal to the exact maximum likelihood estimator.
Appendix E

Ornstein-Uhlenbeck Process

In this Appendix, we will discuss some properties of the Ornstein-Uhlenbeck process, which is given by

\[ dy_t = (\theta - \kappa y_t)dt + \sigma dW_t. \]  \hspace{1cm} (E.1)

Here, \( \theta \) is the level of mean-reversion, \( \kappa \) the speed of mean-reversion and \( \sigma \) the volatility parameter.

In the second model variant of the credit-liquidity model of chapter 5, we model the pure liquidity intensities \( y^{\text{bid}} \) and \( y^{\text{ask}} \) as Ornstein-Uhlenbeck processes with mean-reversion level of zero, and mean-reversion speed parameter \( \eta \). In order to compute the model-implied bid and ask premia, we are interested in expressions of the form

\[ \mathbb{E}_Q \left[ e^{-\int_t^T c y_s ds} \bigg| \mathcal{F}_t \right], \]  \hspace{1cm} (E.2)

where \( c \) is some constant. Again, if we would interpret \( cy \) as the short-rate, expression (E.2) would be the price of a zero-coupon bond. We will, therefore, refer to (E.2) as the bond price under the Ornstein-Uhlenbeck process. We note that \( cy_t \) is again an Ornstein-Uhlenbeck process, since
Chapter E. Ornstein-Uhlenbeck Process

\[
\begin{align*}
\text{d}(cy_t) &= cd_{y_t} \\
&= c(\theta - \kappa y_t)dt + c\sigma dW_t \\
&= (c\theta - c\kappa y_t)dt + c\sigma dW_t \\
&= (\tilde{\theta} - \kappa c y_t)dt + \tilde{\sigma} dW_t.
\end{align*}
\]

In the next section, we will derive an explicit solution for the bond price under the Ornstein-Uhlenbeck process. After that, we will derive the explicit solution of the Ornstein-Uhlenbeck SDE, which gives the (conditional) distributional properties of the process. In the last section, we will derive maximum likelihood estimators of the process parameters using these distributional properties (exact maximum likelihood) and we will also derive approximate maximum likelihood estimators using the Continuous Record maximum likelihood method, which was introduced in chapter 8.

E.1 Bond Price under Ornstein-Uhlenbeck Process

Since the Ornstein-Uhlenbeck process (E.1) is an affine process, we get by Proposition 4.5 that

\[
\mathbb{E}^Q \left[ e^{-\int_t^T cy_s ds} \mid \mathcal{F}_t \right] = e^{-A(t,T) - B(t,T)c y_t},
\]

where \( A(t, T) \) and \( B(t, T) \) satisfy the following system of ordinary differential equations:

\[
\begin{align*}
\partial_t A(t, T) &= \frac{1}{2} (c\sigma)^2 B^2(t, T) - c\theta B(t, T), \quad A(T, T) = 0, \\
\partial_t B(t, T) &= c\kappa B(t, T) - 1, \quad B(T, T) = 0.
\end{align*}
\]

Clearly, the solution of \( B(t, T) \) is given by

\[
B(t, T) = \frac{1}{c\kappa} \left( 1 - e^{-c\kappa(T-t)} \right).
\]

Now, \( A(t, T) \) can be computed as an ordinary integral. We get
E.1. Bond Price under Ornstein-Uhlenbeck Process

\[ A(t, T) = A(T, T) - \int_t^T \partial_s A(s, T) ds \]

\[ = -\frac{(c\sigma)^2}{2} \int_t^T B^2(s, T) ds + c\theta \int_t^T B(s, T) ds \]

\[ = -\frac{(c\sigma)^2}{2} \int_t^T \frac{1}{(c\kappa)^2} \left(1 - e^{-c\kappa(T-s)}\right)^2 ds + c\theta \int_t^T \frac{1}{c\kappa} \left(1 - e^{-c\kappa(T-s)}\right) ds \]

\[ = -\frac{(c\sigma)^2}{2(c\kappa)^2} \left(\int_t^T ds - 2 \int_t^T e^{-c\kappa(T-s)} ds + \int_t^T e^{-2c\kappa(T-s)} ds\right) \]

\[ + \frac{c\theta}{c\kappa} \left(\int_t^T ds - \int_t^T e^{-c\kappa(T-s)} ds\right) \]

\[ = -\frac{(c\sigma)^2}{2(c\kappa)^2} \left(T - t - \frac{2}{c\kappa} + \frac{2}{c\kappa} e^{-c\kappa(T-t)} + \frac{1}{2c\kappa} - \frac{1}{2c\kappa} e^{-2c\kappa(T-t)}\right) \]

\[ + \frac{c\theta}{c\kappa} \left(\left(T - t - \frac{1}{c\kappa} + \frac{1}{c\kappa} e^{-c\kappa(T-t)}\right) \right) \]

\[ = -\frac{(c\sigma)^2}{4(c\kappa)^3} \left(2c\kappa(T - t) + 4e^{-c\kappa(T-t)} - e^{-2c\kappa(T-t)} - 3\right) \]

\[ + \frac{c\theta}{(c\kappa)^2} \left(c\kappa(T - t) - 1 + e^{-c\kappa(T-t)}\right) \]

\[ = \frac{(c\sigma)^2}{4(c\kappa)^3} \left(e^{-2c\kappa(T-t)} - 4e^{-c\kappa(T-t)} - 2c\kappa(T - t) + 3\right) + \frac{c\theta}{(c\kappa)^2} \left(c\kappa(T - t) - 1 + e^{-c\kappa(T-t)}\right). \]

We thus conclude that

\[ \mathbb{E}^Q \left[ e^{-\int_t^T cy\, ds} \bigg| \mathcal{F}_t \right] = e^{-A(t,T)-B(t,T)c\kappa}, \]

with

\[ A(t, T) = \frac{(c\sigma)^2}{4(c\kappa)^3} \left(e^{-2c\kappa(T-t)} - 4e^{-c\kappa(T-t)} - 2c\kappa(T - t) + 3\right) + \frac{c\theta}{(c\kappa)^2} \left(c\kappa(T - t) - 1 + e^{-c\kappa(T-t)}\right), \]

\[ B(t, T) = \frac{1}{c\kappa} \left(1 - e^{-c\kappa(T-t)}\right). \]
E.1.1 Bond Price under Specific Ornstein-Uhlenbeck process

In the credit-liquidity model, we do not use the full, general, Ornstein-Uhlenbeck process, but we take

\[ dy = \eta y(t) dt + \sigma y(t) dW(t), \quad l \in \{bid, ask\}. \]

Although, the computations are very similar to those in the general case, we will again explicitly derive the bond price under the above version of the Ornstein-Uhlenbeck process. We again get by Proposition 4.5 that

\[ \mathbb{E}_Q \left[ e^{-\int_t^T c y(s)ds | \mathcal{F}_t} \right] = e^{\alpha(t,T) - B(t,T)c y(t)}, \]

where \( A(t, T) \) and \( B(t, T) \) satisfy the following differential equations (we drop the superscript that denote the bid and ask parameters for notational convenience):

\[
\begin{align*}
\partial_t A(t, T) &= \frac{1}{2} (c \sigma)^2 B^2(t, T), \quad A(T, T) = 0, \\
\partial_t B(t, T) &= c \eta B(t, T) - 1, \quad B(T, T) = 0.
\end{align*}
\]

We get as solution for \( B(t, T) \)

\[ B(t, T) = \frac{1}{c \eta} \left( 1 - e^{-c \eta (T-t)} \right). \]

Now \( A(t, T) \) can be computed as an ordinary integral
E.2 Explicit Solution of the Ornstein-Uhlenbeck SDE

E.2.1 General Ornstein-Uhlenbeck Process

Consider the Ornstein-Uhlenbeck process given by (E.1). In order to find the solution to this SDE, we first define the function \( g(y_t, t) = ye^{nt} \). Applying Itô’s lemma to this function gives

\[
A(t, T) = A(T, T) - \int_t^T \partial_s A(s, T) ds
\]

\[
= -\frac{(c\sigma)^2}{2} \int_t^T B^2(s, T) ds
\]

\[
= -\frac{(c\sigma)^2}{2} \int_t^T \frac{1}{(c\eta)^2} \left(1 - e^{-c\eta(T-s)}\right)^2 ds
\]

\[
= -\frac{(c\sigma)^2}{2(c\eta)^2} \left( \int_t^T ds - 2 \int_t^T e^{-c\eta(T-s)} ds + \int_t^T e^{-2c\eta(T-s)} ds \right)
\]

\[
= -\frac{(c\sigma)^2}{2(c\eta)^2} \left( (T - t) - \frac{2}{c\eta} + \frac{2}{c\eta} e^{-c\eta(T-t)} + \frac{1}{2c\eta} - \frac{1}{2c\eta} e^{-2c\eta(T-t)} \right)
\]

\[
= -\frac{(c\sigma)^2}{4(c\eta)^3} \left( 2c\eta(T - t) + 4e^{-c\eta(T-t)} - e^{-2c\eta(T-t)} - 3 \right)
\]

\[
= \frac{(c\sigma)^2}{4(c\eta)^3} \left( e^{-2c\eta(T-t)} - 4e^{-c\eta(T-t)} - 2c\eta(T - t) + 3 \right).
\]

We conclude that

\[
E_Q \left[ e^{-\int_t^T c\eta^{-1}(s) ds} \mid F_t \right] = e^{-A(t, T) - B(t, T)c\eta^{-1}(t)}, \tag{E.3}
\]

with

\[
A(t, T) = \frac{(c\sigma)^2}{4(c\eta)^3} \left( e^{-2c\eta(T-t)} - 4e^{-c\eta(T-t)} - 2c\eta(T - t) + 3 \right), \tag{E.4}
\]

\[
B(t, T) = \frac{1}{c\eta} \left( 1 - e^{-c\eta(T-t)} \right). \tag{E.5}
\]

E.2 Explicit Solution of the Ornstein-Uhlenbeck SDE

E.2.1 General Ornstein-Uhlenbeck Process

Consider the Ornstein-Uhlenbeck process given by (E.1). In order to find the solution to this SDE, we first define the function \( g(y_t, t) = ye^{nt} \). Applying Itô’s lemma to this function gives

\[
A(t, T) = A(T, T) - \int_t^T \partial_s A(s, T) ds
\]

\[
= -\frac{(c\sigma)^2}{2} \int_t^T B^2(s, T) ds
\]

\[
= -\frac{(c\sigma)^2}{2} \int_t^T \frac{1}{(c\eta)^2} \left(1 - e^{-c\eta(T-s)}\right)^2 ds
\]

\[
= -\frac{(c\sigma)^2}{2(c\eta)^2} \left( \int_t^T ds - 2 \int_t^T e^{-c\eta(T-s)} ds + \int_t^T e^{-2c\eta(T-s)} ds \right)
\]

\[
= -\frac{(c\sigma)^2}{2(c\eta)^2} \left( (T - t) - \frac{2}{c\eta} + \frac{2}{c\eta} e^{-c\eta(T-t)} + \frac{1}{2c\eta} - \frac{1}{2c\eta} e^{-2c\eta(T-t)} \right)
\]

\[
= -\frac{(c\sigma)^2}{4(c\eta)^3} \left( 2c\eta(T - t) + 4e^{-c\eta(T-t)} - e^{-2c\eta(T-t)} - 3 \right)
\]

\[
= \frac{(c\sigma)^2}{4(c\eta)^3} \left( e^{-2c\eta(T-t)} - 4e^{-c\eta(T-t)} - 2c\eta(T - t) + 3 \right).
\]

We conclude that

\[
E_Q \left[ e^{-\int_t^T c\eta^{-1}(s) ds} \mid F_t \right] = e^{-A(t, T) - B(t, T)c\eta^{-1}(t)}, \tag{E.3}
\]

with

\[
A(t, T) = \frac{(c\sigma)^2}{4(c\eta)^3} \left( e^{-2c\eta(T-t)} - 4e^{-c\eta(T-t)} - 2c\eta(T - t) + 3 \right), \tag{E.4}
\]

\[
B(t, T) = \frac{1}{c\eta} \left( 1 - e^{-c\eta(T-t)} \right). \tag{E.5}
\]
\[\begin{align*}
\text{Chapter E. Ornstein-Uhlenbeck Process}\\
\text{d} g(y_t, t) &= \kappa y_t e^{\kappa t} dt + e^{\kappa t} dy_t \\
&= \left[ \kappa y_t e^{\kappa t} + e^{\kappa t} \left( \theta - \kappa y_t \right) \right] dt + \sigma e^{\kappa t} dW_t \\
&= \theta e^{\kappa t} dt + \sigma e^{\kappa t} dW_t
\end{align*}\]

Now, for each \( s \leq t \), integrating both sides gives

\[y_t e^{\kappa t} - y_s e^{\kappa s} = \int_s^t \theta e^{\kappa u} du + \int_s^t \sigma e^{\kappa u} dW_u.\]

Rewriting gives

\[\begin{align*}
y_t &= y_s e^{-\kappa (t-s)} + e^{-\kappa t} \int_s^t \theta e^{\kappa u} du + \sigma e^{-\kappa t} \int_s^t e^{\kappa u} dW_u \\
&= y_s e^{-\kappa (t-s)} + e^{-\kappa t} \left. \frac{\theta}{\kappa} e^{\kappa u} \right|_{u=s} + \sigma e^{-\kappa t} \int_s^t e^{\kappa u} dW_u \\
&= y_s e^{-\kappa (t-s)} + \frac{\theta}{\kappa} \left( 1 - e^{-\kappa (t-s)} \right) + \sigma e^{-\kappa t} \int_s^t e^{\kappa u} dW_u.
\end{align*}\]

We thus get that a solution to the Ornstein-Uhlenbeck SDE is given by

\[y_t = y_s e^{-\kappa (t-s)} + \theta \kappa \left( 1 - e^{-\kappa (t-s)} \right) + \sigma e^{-\kappa t} \int_s^t e^{\kappa u} dW_u. \tag{E.6}\]

Now we will show that this is also the unique solution to the Ornstein-Uhlenbeck stochastic differential equation. Let us first write the Ornstein-Uhlenbeck SDE as follows:

\[dy_t = z(t, y_t) dt + \sigma(t, y_t) dW_t \text{ with } z(t, y_t) = \theta - \kappa y_t \text{ and } \sigma(t, y_t) = \sigma.\]

If we consider the functions \( z(t, x) \) and \( \sigma(t, x) \), we have that Theorem 4.4 of Filipović (2009) [27] tells us that there exists a unique solution to the SDE with initial time-space point \((0, y_0)\), if there exists a finite constant \( K \) such that the following two conditions hold for all \( t \geq 0 \) and \( x, y \in \mathbb{R} \):

- \( ||z(t, x) - z(t, y)|| + ||\sigma(t, x) - \sigma(t, y)|| \leq K ||x - y|| \),
E.2. Explicit Solution of the Ornstein-Uhlenbeck SDE

- \(|z(t, x)|^2 + |\sigma(t, x)|^2 \leq K^2(1 + |x|^2)|.

Note that in the Ornstein-Uhlenbeck case, the norms can be replaced by absolute values, since \(z(t, x), \sigma(t, x), x\) and \(t\) are all real-valued. It easily follows that there exists such a finite \(K\) for the OU SDE, since we have for the first inequality that

\[
||z(t, x) - z(t, y)|| + ||\sigma(t, x) - \sigma(t, y)|| = ||(\theta - \kappa x) - (\theta - \kappa y)|| + ||\sigma - \sigma|| \\
= ||\kappa(y - x)|| + 0 \\
\leq |\kappa||y - x||.
\]

So for the first inequality, \(K \geq |\kappa|\) would suffice. For the second inequality we get:

\[
||z(t, x)||^2 + ||\sigma(t, x)||^2 = ||\theta - \kappa x||^2 + |\sigma|^2 \\
\leq |\theta|^2 + 2|\theta\kappa x| + |\kappa|^2|x|^2 + |\sigma|^2
\]

Now it is obvious that we can always find a finite \(K\) such that \(|\theta|^2 + 2|\theta\kappa x| + |\kappa|^2|\sigma|^2 \leq K^2(1 + |x|^2)|, since for large enough \(K\) we have that \(K^2 \) dominates \(|\theta|^2 + |\sigma|^2\) and \(K^2|x|^2\) dominates \(2|\theta\kappa x| + |\kappa|^2|x|^2\). This last statement follows since the dominant factor in the \((2|\theta\kappa x| + |\kappa|^2|x|^2)\)-term is, for large values of \(|x|\), the \((|\kappa|^2|x|^2)\)-term and therefore, for large \(K\) we get that \(K^2|x|^2\) dominates the entire \((2|\theta\kappa x| + |\kappa|^2|x|^2)\)-term. We conclude that the above solution is the unique solution to the Ornstein-Uhlenbeck SDE.

We get from equation (E.6) that, conditional on \(\mathcal{F}_s\), the variable \(y_t\) is normally distributed with mean \(y_s e^{-\kappa(t-s)} + \frac{\theta}{\kappa} \left(1 - e^{-\kappa(t-s)}\right)\) and variance \(\frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-s)}\right)\). This follows again from Lemma D.1.

E.2.2 Ornstein-Uhlenbeck Process with Mean-Reversion Level of Zero

In a completely similar way, we can show that the unique solution of the following stochastic differential equation:

\[
dy_t = -\eta y_t dt + \sigma dW_t,
\]

is given by

\[
y_t = y_s e^{-\eta(t-s)} + \sigma e^{-\eta t} \int_s^t e^{\eta u} dW_u. \tag{E.7}
\]
We have that, conditional on $\mathcal{F}_s$, the variable $y_t$ is normally distributed with mean $y_se^{-\eta(t-s)}$ and variance $\frac{\sigma^2}{2\eta} \left( 1 - e^{-2\eta(t-s)} \right)$.

### E.3 Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process

Consider the Ornstein-Uhlenbeck process given by

$$dy_t = (\theta - \kappa y_t)dt + \sigma dW_t.$$ 

We know, by the previous section, that, conditional on $\mathcal{F}_s$, $y_t$ is normally distributed with mean $y_se^{-\kappa(t-s)} + \frac{\theta}{\kappa} \left( 1 - e^{-\kappa(t-s)} \right)$ and variance $\frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa(t-s)} \right)$.

### E.3.1 Exact Maximum Likelihood

Suppose that we have $(N+1)$ observations, $y_i$, for $i = 0, 1, \ldots, N$. We get that the conditional probability density function of an observation $y_{i+1}$, given a previous observation $y_i$, is given by

$$f(y_{i+1}|y_i, \theta, \kappa, V) = \frac{1}{\sqrt{2\pi V^2}} \exp \left( - \frac{(y_{i+1} - y_i e^{-\kappa \Delta} - \frac{\theta}{\kappa} \left( 1 - e^{-\kappa \Delta} \right))^2}{2V^2} \right),$$

where $\Delta$ is the time difference between $t_{i+1}$ and $t_i$ (which we assume to be equal for all $i$), and $V^2 = \frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa \Delta} \right)$. Using the change of variables,

$$\alpha = e^{-\kappa \Delta},$$
$$\beta = \frac{\theta}{\kappa},$$
$$V^2 = \frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa \Delta} \right),$$

we can rewrite the conditional probability density function as follows:

$$f(y_{i+1}|y_i, \alpha, \beta, V) = \frac{1}{\sqrt{2\pi V^2}} \exp \left( - \frac{(y_{i+1} - (y_i \alpha + \beta(1 - \alpha)))^2}{2V^2} \right).$$
E.3. Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process

The log-likelihood function is now given by

$$L(\alpha, \beta, V) = \sum_{i=1}^{N} f(y_i|y_{i-1}, \alpha, \beta, V^2)$$

$$= -\frac{N}{2} \log(2\pi) - N \log(V) - \frac{1}{2V^2} \sum_{i=1}^{N} (y_i - y_{i-1}\alpha - \beta(1-\alpha))^2.$$

We find the maximum likelihood estimators by taking the partial derivatives and equating them to zero and then solving them for the parameters. We get the following partial derivatives:

$$\frac{\partial L(\alpha, \beta, V)}{\partial \alpha} = -\frac{1}{V^2} \sum_{i=1}^{N} (y_i - (y_{i-1}\alpha + \beta(1-\alpha))) (\beta - y_{i-1}), \quad (E.8)$$

$$\frac{\partial L(\alpha, \beta, V)}{\partial \beta} = -\frac{1}{V^2} \sum_{i=1}^{N} (y_i - (y_{i-1}\alpha + \beta(1-\alpha))) (-1) \alpha), \quad (E.9)$$

$$\frac{\partial L(\alpha, \beta, V)}{\partial V} = -\frac{N}{V} + \frac{1}{V^3} \sum_{i=1}^{N} (y_i - y_{i-1}\alpha - \beta(1-\alpha))^2. \quad (E.10)$$

The partial derivatives evaluated in the maximum likelihood estimators $\hat{\alpha}, \hat{\beta}$ and $\hat{V}$ make the above equations equal to zero. Solving equation (E.9) for $\hat{\beta}$ gives

$$\frac{\partial L(\hat{\alpha}, \hat{\beta}, \hat{V})}{\partial \beta} = \sum_{i=1}^{N} (y_i - (y_{i-1}\hat{\alpha} + \hat{\beta}(1-\hat{\alpha}))\) \ (-1 - \hat{\alpha})$$

$$= \sum_{i=1}^{N} y_i(-1 - \hat{\alpha}) - \sum_{i=1}^{N} y_{i-1}\hat{\alpha}(-1 - \hat{\alpha}) - \sum_{i=1}^{N} \hat{\beta}(1 - \hat{\alpha})(-1 - \hat{\alpha})$$

$$= \sum_{i=1}^{N} y_i\hat{\alpha} - \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_{i-1}\hat{\alpha}^2 + \sum_{i=1}^{N} y_{i-1}\hat{\alpha} + \sum_{i=1}^{N} \hat{\beta}(1 - \hat{\alpha})^2$$

$$= \sum_{i=1}^{N} y_i\hat{\alpha} - \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_{i-1}\hat{\alpha}^2 + \sum_{i=1}^{N} y_{i-1}\hat{\alpha} + N(1 - \hat{\alpha})^2 \hat{\beta}$$

$$= 0.$$

This gives
\[ \hat{\beta} = \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i \hat{\alpha} - \sum_{i=1}^{N} y_{i-1} \hat{\alpha} + \sum_{i=1}^{N} y_{i-1} \hat{\alpha}^2}{N(1 - \hat{\alpha})^2} \]
\[ = \frac{\sum_{i=1}^{N} y_i (1 - \hat{\alpha}) - \sum_{i=1}^{N} y_{i-1} \hat{\alpha}(1 - \hat{\alpha})}{N(1 - \hat{\alpha})^2} \]
\[ = \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_{i-1} \hat{\alpha}}{N(1 - \hat{\alpha})}. \tag{E.11} \]

Solving equation (E.8) gives

\[ \frac{\partial \mathcal{L}(\hat{\alpha}, \hat{\beta}, \hat{\mathcal{V}}^2)}{\partial \alpha} = \sum_{i=1}^{N} \left( y_i - \left( y_{i-1} \hat{\alpha} + \hat{\beta}(1 - \hat{\alpha}) \right) \right) \left( \hat{\beta} - y_{i-1} \right) \]
\[ = \sum_{i=1}^{N} y_i(\hat{\beta} - y_{i-1}) - \sum_{i=1}^{N} y_{i-1} \hat{\alpha}(\hat{\beta} - y_{i-1}) - \sum_{i=1}^{N} \hat{\beta}(1 - \hat{\alpha})(\hat{\beta} - y_{i-1}) \]
\[ = \sum_{i=1}^{N} y_i \hat{\beta} - \sum_{i=1}^{N} y_i y_{i-1} - \sum_{i=1}^{N} y_{i-1} \hat{\alpha} \hat{\beta} + \sum_{i=1}^{N} y_{i-1} \hat{\alpha} - \sum_{i=1}^{N} \hat{\beta}^2 (1 - \hat{\alpha}) + \sum_{i=1}^{N} \hat{\beta}(1 - \hat{\alpha}) y_{i-1}. \]

Using the expression we found for \( \hat{\beta} \) in the above equation gives
E.3. Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process

\[
\frac{\partial L(\hat{\alpha}, \hat{\beta}, \hat{V}^2)}{\partial \alpha} = \sum_{i=1}^{N} y_i \left( \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} \right) - \sum_{i=1}^{N} y_i y_{i-1} \\
- \sum_{i=1}^{N} y_{i-1} \hat{\alpha} \left( \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} \right) + \sum_{i=1}^{N} y_{i-1}^2 \hat{\alpha} \\
- \sum_{i=1}^{N} (1 - \hat{\alpha}) \left( \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} \right)^2 \\
+ \sum_{i=1}^{N} (1 - \hat{\alpha}) y_{i-1} \left( \frac{\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} \right)
\]

\[
= \frac{\left( \sum_{i=1}^{N} y_i \right)^2}{N(1 - \hat{\alpha})} - \frac{\hat{\alpha} \sum_{i=1}^{N} y_i \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} - \sum_{i=1}^{N} y_i y_{i-1} - \hat{\alpha} \sum_{i=1}^{N} y_i \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} \\
- \sum_{i=1}^{N} y_i y_{i-1} + \hat{\alpha}^2 \frac{\left( \sum_{i=1}^{N} y_i \right)^2}{N(1 - \hat{\alpha})} + \sum_{i=1}^{N} y_{i-1}^2 \hat{\alpha} - N(1 - \hat{\alpha}) \frac{\left( \sum_{i=1}^{N} y_i \right)^2}{N(1 - \hat{\alpha})^2} \\
+ 2N(1 - \hat{\alpha}) \frac{\sum_{i=1}^{N} y_i \sum_{i=1}^{N} y_i - \hat{\alpha}}{(N(1 - \hat{\alpha})^2) - N(1 - \hat{\alpha}) \hat{\alpha}^2 \frac{\left( \sum_{i=1}^{N} y_i - \hat{\alpha} \right)^2}{(N(1 - \hat{\alpha})^2} \\
- \hat{\alpha}(1 - \hat{\alpha}) \frac{\sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})} + (1 - \hat{\alpha}) \frac{\sum_{i=1}^{N} y_i \sum_{i=1}^{N} y_i - \hat{\alpha}}{N(1 - \hat{\alpha})}
\]

\[
= - \sum_{i=1}^{N} y_i y_{i-1} + \sum_{i=1}^{N} y_{i-1}^2 \hat{\alpha} - \hat{\alpha} \frac{\left( \sum_{i=1}^{N} y_i - \hat{\alpha} \right)^2}{N} + \sum_{i=1}^{N} y_i \frac{\sum_{i=1}^{N} y_i - \hat{\alpha}}{N - 1}.
\]

Equating this last expression to zero and solving for \( \hat{\alpha} \) gives

\[
\hat{\alpha} = \frac{N \sum_{i=1}^{N} y_i y_{i-1} - \sum_{i=1}^{N} y_i \sum_{i=1}^{N} y_i - \hat{\alpha}}{N \sum_{i=1}^{N} y_{i-1}^2 - \left( \sum_{i=1}^{N} y_{i-1} \right)^2}.
\]  

(E.12)

As a last step we set equation (E.10) equal to zero and solve for \( \hat{V} \). We obtain

\[
\hat{V}^2 = \frac{1}{N} \sum_{i=1}^{N} \left( y_i - \hat{\alpha} y_{i-1} - \hat{\beta}(1 - \hat{\alpha}) \right)^2.
\]  

(E.13)

The parameters \( \kappa, \theta \) and \( \sigma \) can be obtained by substituting back the above expressions.
Chapter E. Ornstein-Uhlenbeck Process

E.3.2 Continuous Record Maximum Likelihood

The Continuous Record likelihood method only works if the diffusion function is known and does not depend on the unknown parameters. Therefore, one can only estimate the drift parameters using this method. In the introduction of chapter 8, we derived that the Continuous Record log-Likelihood function of a general continuous diffusion process (8.1) is given by

\[
L_T(\theta) = \int_0^T \frac{\mu(\theta, X_t)}{\sigma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(\theta, X_t)}{\sigma^2(X_t)} dt.
\]

In the case of the Ornstein-Uhlenbeck process, we have that \( \mu(\theta, y_t) = (\theta - \kappa y_t) \) and \( \sigma(y_t) = \sigma \), which is assumed to be known. The log-likelihood function, therefore, becomes

\[
L_T(\theta, \kappa) = \int_0^T \frac{(\theta - \kappa y_t)}{\sigma^2} dy_t - \frac{1}{2} \int_0^T \frac{(\theta - \kappa y_t)^2}{\sigma^2} dt.
\]

The ML estimates \( \hat{\theta} \) and \( \hat{\kappa} \) can be found by putting the partial derivatives w.r.t. \( \theta \) and \( \kappa \) equal to zero and solve for \( \hat{\theta} \) and \( \hat{\kappa} \). We get the following partial derivatives:

\[
\frac{\partial L_T(\theta, \kappa)}{\partial \theta} = \int_0^T \frac{1}{\sigma^2} dy_t - \int_0^T \frac{(\theta - \kappa y_t)}{\sigma^2} dt = 0, \tag{E.14}
\]

\[
\frac{\partial L_T(\theta, \kappa)}{\partial \kappa} = -\int_0^T \frac{y_t}{\sigma^2} dy_t + \int_0^T \frac{\theta y_t}{\sigma^2} dt - \int_0^T \frac{\kappa y_t^2}{\sigma^2} dt = 0. \tag{E.15}
\]

Suppose now that we observe the process \( y_t \) at equidistant discrete time points \( 0, \Delta, 2\Delta, \ldots, N\Delta \). We replace the Lebesgue and Itô integrals by Riemann-Itô sums and we get for equation (E.14) the following:

\[
\frac{\partial L_T(\theta, \kappa)}{\partial \theta} \approx \frac{1}{\sigma^2} \sum_{i=1}^N (y_{i\Delta} - y_{(i-1)\Delta}) - \frac{\theta \Delta}{\sigma^2} + \sum_{i=1}^N \frac{\kappa y_{(i-1)\Delta}}{\sigma^2} \Delta
\]

\[
= (y_{N\Delta} - y_0) - N\Delta \theta + \kappa \Delta \sum_{i=1}^N y_{(i-1)\Delta}
\]

\[
= 0.
\]

Solving for \( \hat{\theta} \) gives 190
E.3. Maximum Likelihood Estimators for Ornstein-Uhlenbeck Process

\[ \hat{\theta} = \frac{1}{N \Delta} \left( y_N \Delta - y_0 + \hat{\kappa} \Delta \sum_{i=1}^{N} y_{(i-1)\Delta} \right). \]  

(E.16)

Equation (E.15) is now given by

\[ \frac{\partial \mathcal{L}_T(\theta, \kappa)}{\partial \kappa} \approx - \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}}{\sigma^2} (y_i \Delta - y_{(i-1)\Delta}) + \theta \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}}{\sigma^2} \Delta - \kappa \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}^2}{\sigma^2} \Delta = 0. \]

Using the expression we found for \( \hat{\theta} \) gives

\[
\begin{align*}
0 &= - \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}}{\sigma^2} (y_i \Delta - y_{(i-1)\Delta}) + \hat{\theta} \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}}{\sigma^2} \Delta - \hat{\kappa} \Delta \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}^2}{\sigma^2} - \kappa \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}^2}{\sigma^2} \Delta \\
&= - \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}}{\sigma^2} (y_i \Delta - y_{(i-1)\Delta}) + \frac{1}{N \Delta} \left( y_N \Delta - y_0 \right) \sum_{i=1}^{N} y_{(i-1)\Delta} + \hat{\kappa} \Delta \left( \sum_{i=1}^{N} y_{(i-1)\Delta} \right)^2 - \hat{\kappa} \Delta \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}^2}{\sigma^2} \\
&= - \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}}{\sigma^2} (y_i \Delta - y_{(i-1)\Delta}) + \frac{1}{N} \left( y_N \Delta - y_0 \right) \sum_{i=1}^{N} y_{(i-1)\Delta} + \hat{\kappa} \Delta \left( \frac{1}{N} \left( \sum_{i=1}^{N} y_{(i-1)\Delta} \right)^2 - \sum_{i=1}^{N} \frac{y_{(i-1)\Delta}^2}{\sigma^2} \right). 
\end{align*}
\]

Now solving for \( \hat{\kappa} \) gives

\[ \hat{\kappa} = \frac{1}{\Delta} \left( \frac{(y_N \Delta - y_0) \sum_{i=1}^{N} y_{(i-1)\Delta} - N \sum_{i=1}^{N} y_{(i-1)\Delta} (y_i \Delta - y_{(i-1)\Delta})}{N \sum_{i=1}^{N} y_{(i-1)\Delta}^2 - \left( \sum_{i=1}^{N} y_{(i-1)\Delta} \right)^2} \right). \]  

(E.17)
Appendix F

CIR Process

In this Appendix, we will discuss some properties of the CIR process, which is given by,

\[ dx_t = (\alpha - \beta x_t)dt + \sigma \sqrt{x_t}dW_t. \]  \hspace{1cm} (F.1)

Here, \( \alpha \) is the mean-reversion level, \( \beta \) the mean-reversion speed and \( \sigma \) the volatility parameter.

Within our credit-liquidity model, we are interested in expressions of the form

\[ \mathbb{E} \left[ e^{-\int_t^T cx_s ds} \mid \mathcal{F}_t \right], \]  \hspace{1cm} (F.2)

where \( c \) is a constant. Note that \( cx_t \) again follows a CIR process, since its dynamics are of the form

\[ d(cx_t) = (c\alpha - \beta cx_t)dt + \sigma \sqrt{cx_t}dW_t \]
\[ = (\tilde{\alpha} - \beta y_t)dt + \tilde{\sigma} \sqrt{y_t}dW_t, \]

where we put \( y(t) = cx(t) \) and the definitions of \( \tilde{\alpha} \) and \( \tilde{\sigma} \) are self-explanatory. If we would interpret \( cx \) as a short-rate process, expression (F.2) would be the zero-coupon bond price, and, therefore, we will refer to this expression as the bond price under the CIR process.

We know by Proposition 4.5 that the CIR dynamics give rise to the following formula

\[ \mathbb{E} \left[ e^{-\int_t^T x(s) ds} \mid \mathcal{F}_t \right] = e^{-A(t,T) - B(t,T)x(t)}, \]
where $A(t, T)$ and $B(t, T)$ satisfy the following system of ordinary differential equations:

$$
\begin{align*}
\partial_t A(t, T) &= -\alpha B(t, T), \quad A(T, T) = 0 \\
\partial_t B(t, T) &= \frac{1}{2}\sigma^2 B^2(t, T) + \beta B(t, T) - 1, \quad B(T, T) = 0.
\end{align*}
$$

If we replace $\tau = T - t$, then we get the following system of ordinary differential equations:

$$
\begin{align*}
\partial_\tau A(\tau) &= \alpha B(\tau), \quad A(0) = 0 \\
\partial_\tau B(\tau) &= -\frac{1}{2}\sigma^2 B^2(\tau) - \beta B(\tau) + 1, \quad B(0) = 0.
\end{align*}
$$

The second differential equation is known as a Riccati equation. In the next section, we will show how to solve a Riccati equation with constant coefficients in general. After that, we will use these results to find the solution of the bond price under the CIR process. We will end this Appendix with the derivation of the Continuous Record maximum likelihood estimators of the drift parameters.

### F.1 Solution for Riccati Equation With Constant Coefficients

Consider the Riccati equation

$$
y' = py^2 + qy + r, \quad (F.3)
$$

with $p, q$ and $r$ constant coefficients and where $y' = \frac{dy}{dx}$. Suppose that we have a particular solution, $y_0(x) = y_0$, to this Riccati equation. Then we know that the general solution is given by $y = y_0 + u$. Substituting $y_0 + u$ in (F.3) gives us

$$
(y_0' + u') = p(y_0 + u)^2 + q(y_0 + u) + r.
$$

Since $y_0$ is a particular solution, that is

$$
y_0' = py_0^2 + qy_0 + r,
$$

we get that $u$ satisfies the following ordinary differential equation:
F.1. Solution for Riccati Equation With Constant Coefficients

\[ u' = pu^2 + (q + 2py_0)u. \]

Substituting \( u = \frac{1}{z} \) and filling this in into the equation of \( u' \) gives

\[ z' = -(q + 2py_0)z - p. \]

So a general solution of the Riccati equation is given by \( y = y_0 + \frac{1}{z} \), where \( z \) satisfies the above linear first-order differential equation. Consider the following equation:

\[ pv^2 + qv + r = 0. \]

The roots of this equation are given by

\[ v_{\pm} = \frac{-q \pm \sqrt{q^2 - 4pr}}{2p}. \]

We will only consider the situation where the roots are real-valued (which will be the case for the CIR bond price formula). Now we have that \( v_{\pm} \) are particular solutions to the Riccati equation since

\[ v_{\pm}' = 0, \]

where \( v_{\pm}' = \frac{\partial v_{\pm}}{\partial x} \), and if we look at for example \( v_- \) (\( v_+ \) follows in a similar way), we get

\[
\begin{align*}
pv_-^2 + qv_- + r &= p \left( -q - \sqrt{q^2 - 4pr} \right)^2 + q - q - \sqrt{q^2 - 4pr} + r \\
&= p \left( \frac{q^2 - 2q\sqrt{q^2 - 4pr} + (q^2 - 4pr)}{2p} \right) + \frac{-q^2 - q \sqrt{q^2 - 4pr} + r}{2p} \\
&= \frac{1}{4p} \left( \frac{q^2 - 2q\sqrt{q^2 - 4pr} + (q^2 - 4pr)}{2p} \right) - \frac{q^2 - q \sqrt{q^2 - 4pr} + r}{2p} \\
&= 0.
\end{align*}
\]

So, if we take \( y_0 = v_- \) and we consider the first-order linear differential equation of \( z \), we obtain
\[ z' = -(q + 2py)z - p \]
\[ = -\left( q + 2p \left( \frac{-q - \sqrt{q^2 - 4pr}}{2p} \right) \right) z - p \]
\[ = -\left( q - q - \sqrt{q^2 - 4pr} \right) z - p \]
\[ = \sqrt{q^2 - 4pr}z - p \]
\[ = \gamma z - p, \]

where \( \gamma = \sqrt{q^2 - 4pr} \). We can now easily find the solution for \( z \), as a particular solution is \( z_0 = \frac{p}{\gamma} \) and a solution to the homogeneous part is \( Ce^{\gamma x} \), where \( C \) is an arbitrary constant. So we have the following general solution of \( z \):

\[ z = \frac{p}{\gamma} + Ce^{\gamma x}. \]

This means that we now also have the solution for the Riccati equation since

\[ y = y_0 + \frac{1}{z} \]
\[ = v_- + \frac{1}{\frac{p}{\gamma} + Ce^{\gamma x}} \]
\[ = \frac{v_- \left( \frac{p}{\gamma} + Ce^{\gamma x} \right) + 1}{\frac{p}{\gamma} + Ce^{\gamma x}} \]
\[ = \frac{(p + C\gamma e^{\gamma x})v_- + \gamma}{p + C\gamma e^{\gamma x}} \]
\[ = \frac{(pv_- + \gamma) + Cv_- \gamma e^{\gamma x}}{p + C\gamma e^{\gamma x}}. \]

Suppose now that we would have as initial value condition \( y(0) = 0 \) (which is the case when we want to obtain the CIR bond price formula), then we can find the value of \( C \) such that \( y(0) = 0 \). We get that this value is given by solving

\[ (pv_- + \gamma) + Cv_- \gamma = 0. \]
It appears that \( C = -\frac{p v_- + \gamma}{\gamma v_-} \). Noting that \( p v_- + \gamma = p v_+ \) gives

\[
y(x) = \frac{p v_+ - p v_+ e^{\gamma x}}{p - \frac{p v_+}{v_-} e^{\gamma x}} \]
\[
= \frac{p v_+ (1 - e^{\gamma x})}{p (1 - \frac{v_+}{v_-} e^{\gamma x})} \]
\[
= \frac{v_+ (1 - e^{\gamma x})}{1 - \frac{v_+}{v_-} e^{\gamma x}}.
\]

**F.2 Bond Price under CIR process**

We will now use the results of the previous section to derive the CIR bond price formula. In order to compute the CIR bond price formula, we had to solve the system of ordinary differential equations

\[
\begin{align*}
\partial_\tau A(\tau) &= \alpha B(\tau), \quad A(0) = 0, \\
\partial_\tau B(\tau) &= -\frac{1}{2} \sigma^2 B^2(\tau) - \beta B(\tau) + 1, \quad B(0) = 0.
\end{align*}
\]

The second equation is a Riccati equation with constant coefficients. Using the solution method of the previous section, we can solve this equation. We find that particular solutions are given by \( v_\pm \), where \( v_\pm \) are the roots of the equation

\[
-\frac{1}{2} \sigma^2 v^2 - \beta v + 1 = 0.
\]

This gives us

\[
v_\pm = \frac{\beta \pm \sqrt{\beta^2 + 2\sigma^2}}{-\sigma^2}.
\]

A quick check of \( v_- \) gives us indeed that \( v'_- = 0 \) and
Chapter F. CIR Process

\[-\frac{\sigma^2}{2}v_- - \beta v_- + 1 = -\frac{1}{2}\sigma^2 \left( -\frac{\beta + \sqrt{\beta^2 + 2\sigma^2}}{\sigma^2} \right)^2 - \beta \left( -\frac{\beta + \sqrt{\beta^2 + 2\sigma^2}}{\sigma^2} \right) + 1 \]

\[= -\frac{1}{2\sigma^2} \left( \beta^2 - 2\beta \sqrt{\beta^2 + 2\sigma^2} + \beta^2 + 2\sigma^2 \right) + \frac{\beta^2}{\sigma^2} - \frac{\beta \sqrt{\beta^2 + 2\sigma^2}}{\sigma^2} + 1 \]

\[= 0, \]

and, therefore, it is a particular solution. Furthermore, we note that \(-\frac{\sigma^2}{2} v_- + \sqrt{\beta^2 + 2\sigma^2} = -\frac{\sigma^2}{2} v_+.

By the results from the previous section, we know that the solution for \(B(\tau)\) with initial value condition \(B(0) = 0\) is given by

\[B(\tau) = v_+ \left( 1 - \frac{1}{-\frac{\beta + \gamma}{v_-} e^{\gamma\tau}} \right)\]

where \(v_\pm\) are as defined above and \(\gamma = \sqrt{\beta^2 + 2\sigma^2}\). Filling in the right expressions for \(v_\pm\) gives us

\[B(\tau) = \left( -\frac{\beta - \gamma}{\sigma^2} \right) \left( 1 - e^{\gamma\tau} \right) \frac{1 - \frac{\beta + \gamma}{\beta - \gamma} e^{\gamma\tau}}{1 - \frac{\beta + \gamma}{\beta - \gamma} e^{\gamma\tau}} \]

\[= \left( -\frac{\beta - \gamma}{\sigma^2} \right) \left( 1 - e^{\gamma\tau} \right) \left( -\beta + \gamma \right) \frac{(-\beta + \gamma) - (-\beta - \gamma) e^{\gamma\tau}}{(\beta + \gamma) (-\beta - \gamma) e^{\gamma\tau}} \]

\[= \left( -\frac{\beta - \gamma}{\sigma^2} \right) \left( 1 - e^{\gamma\tau} \right) \frac{(-\beta - \gamma) (1 - e^{\gamma\tau}) - 2\gamma}{\beta^2 - (\beta^2 + 2\sigma^2)} (1 - e^{\gamma\tau}) \]

\[= \left( -\beta - \gamma \right) (1 - e^{\gamma\tau}) - 2\gamma \]

\[= \left( -\beta - \gamma \right) (1 - e^{\gamma\tau}) - 2\gamma \]

\[= \frac{-2 (1 - e^{\gamma\tau})}{\beta + \gamma \left( e^{\gamma\tau} - 1 \right) + 2\gamma}.\]
We still have to solve the ordinary differential equation
\[ \partial_\tau A(\tau) = \alpha B(\tau), \quad A(0) = 0. \]

We obtain the solution by integration, since we can write
\[ A(\tau) = \int_t^T \alpha B(T - s) ds \]
\[ = \alpha \int_t^T \frac{2 \left( e^{\gamma(T - s)} - 1 \right)}{(\beta + \gamma)(e^{\gamma(T - s)} - 1) + 2\gamma} ds. \]

Using the change of variables
\[ x(s) = e^{\gamma(T - s)}, \]
with
\[ ds = -\frac{1}{\gamma x(s)} dx, \]
gives
\[ A(\tau) = -\frac{2\alpha}{\gamma x(s)} \int_{e^{\gamma(T - t)}}^{1} \left( \frac{x(s) - 1}{2\gamma + (\gamma + \beta)(x(s) - 1)} \right) dx \]
\[ = \frac{2\alpha}{\gamma} \int_{e^{\gamma(T - t)}}^{1} \left( \frac{1}{x(s)(2\gamma + (\gamma + \beta)(x(s) - 1))} - \frac{1}{2\gamma + (\gamma + \beta)(x(s) - 1)} \right) dx \]
\[ = \frac{2\alpha}{\gamma} \int_{e^{\gamma(T - t)}}^{1} \left( \frac{1}{(\beta - \gamma)(2\gamma + (\gamma + \beta)(x(s) - 1))} + \frac{1}{(\gamma - \beta)x(s) - 2\gamma + (\gamma + \beta)(x(s) - 1)} \right) dx. \]

Note that the last step follows from
\[ \frac{1}{(\beta - \gamma)(2\gamma + (\gamma + \beta)(x(s) - 1))} + \frac{1}{(\gamma - \beta)x(s)} \]
\[ = \frac{1}{x(s)(\beta - \gamma)(2\gamma + (\gamma + \beta)(x(s) - 1))} + \frac{1}{(\gamma - \beta)x(s)} \]
\[ = \frac{1}{x(s)(2\gamma + (\gamma + \beta)(x(s) - 1))}. \]
We again use a change of variables, namely

\[ y(x) = 2\gamma + (\gamma + \beta)(x(s) - 1), \]

with

\[ dx = \frac{1}{\gamma + \beta} dy. \]

This gives

\[
A(\tau) = \frac{2\alpha}{\gamma} \int_{2\gamma + (\gamma + \beta)(e^{\gamma(T-t)} - 1)}^{2\gamma} \left( \frac{1}{(\beta - \gamma) y(s)} - \frac{1}{(\gamma + \beta) y(s)} \right) dy + \frac{2\alpha}{\gamma} \int_{e^{\gamma(T-t)}}^{1} \frac{1}{(\gamma - \beta) x(s)} dx \\
= \frac{2\alpha}{\gamma} \left( \left[ \frac{1}{(\beta - \gamma)} \ln(y(s)) - \frac{1}{(\gamma + \beta)} \ln(y(s)) \right]_{2\gamma + (\gamma + \beta)(e^{\gamma(T-t)} - 1)}^{2\gamma} \right) \\
+ \frac{2\alpha}{\gamma} \left[ \frac{1}{(\gamma - \beta)} \ln(x(s)) \right]_{e^{\gamma(T-t)}}^{1} \\
= \frac{2\alpha}{\gamma} \left( \left[ \frac{1}{(\beta - \gamma)} \ln(2\gamma + (\gamma + \beta)(x - 1)) - \frac{1}{(\gamma + \beta)} \ln(\gamma + (\gamma + \beta)(x - 1)) \right]_{e^{\gamma(T-t)}}^{1} \right) \\
+ \frac{2\alpha}{\gamma} \left[ \frac{1}{(\gamma - \beta)} \ln(x(s)) \right]_{e^{\gamma(T-t)}}^{1} \\
= \frac{2\alpha}{\gamma} \left( \left[ \frac{1}{(\beta - \gamma)} \ln \left( 2\gamma + (\gamma + \beta) \left( e^{\gamma(T-t)} - 1 \right) \right) - \frac{1}{(\gamma + \beta)} \ln \left( \gamma + (\gamma + \beta) \left( e^{\gamma(T-t)} - 1 \right) \right) \right]_{t}^{T} \right) \\
+ \frac{2\alpha}{\gamma} \left[ \frac{1}{(\gamma - \beta)} \ln \left( e^{\gamma(T-t)} \right) \right]_{t}^{T}.
\]

Simplifying the last expressions yields the result:

200
F.3. Maximum Likelihood Estimators for CIR Process

F.3.1 Continuous Record Maximum Likelihood

The Continuous Record log-Likelihood function for the CIR process is given by

$$L_T(\alpha, \beta) = \int_0^T \frac{\alpha - \beta x_t}{\sigma^2 x_t} \, dx_t - \frac{1}{2} \int_0^T \frac{(\alpha - \beta x_t)}{\sigma^2 x_t} \, dt,$$

where

$$A(t, T) = -\frac{2\alpha}{\sigma^2} \ln \left( \frac{2\gamma e^{(\gamma+\beta)(T-t)}}{2\gamma + (\gamma + \beta)(e^{\gamma(T-t)} - 1)} \right),$$

$$B(t, T) = \frac{2(\gamma e^{(T-t)})}{(\beta + \gamma)(e^{(T-t)} - 1) + 2\gamma},$$

$$\gamma = \sqrt{\beta^2 + 2\sigma^2}.$$
where we assume that \( \sigma \) is known. Setting the partial derivatives equal to zero gives

\[
\frac{\partial L(\alpha, \beta)}{\partial \alpha} = \int_0^T \frac{1}{\sigma^2 x_t} \, dx_t - \int_0^T \frac{(\alpha - \beta x_t)}{\sigma^2} \, dt = 0,
\]

\[
\frac{\partial L(\alpha, \beta)}{\partial \beta} = \int_0^T \frac{1}{\sigma^2} \, dx_t + \int_0^T \frac{\alpha}{\sigma^2} \, dt - \int_0^T \frac{\beta x_t}{\sigma^2} \, dt = 0.
\]

If we have \( N+1 \) equidistant observations \( x_{t_i} \), with \( t_i = i\Delta \) for \( i = 0, 1, \ldots, N \), then replacing the Lebesgue and Itô integrals by Riemann-Itô sums and ignoring the factors \( \frac{1}{\sigma^2} \) gives the following approximations:

\[
\frac{\partial L(\alpha, \beta)}{\partial \alpha} \approx \sum_{i=1}^N \frac{x_i \Delta - x_{(i-1)\Delta}}{x_{(i-1)\Delta}} - \sum_{i=1}^N \frac{\alpha}{x_{(i-1)\Delta}} + \sum_{i=1}^N \frac{\beta \Delta}{x_{(i-1)\Delta}}
\]

\[
= \sum_{i=1}^N \frac{x_i \Delta}{x_{(i-1)\Delta}} - N - \alpha \Delta \sum_{i=1}^N \frac{1}{x_{(i-1)\Delta}} + \beta \Delta N
\]

\[
= 0.
\]

Solving for \( \hat{\alpha} \) gives

\[
\hat{\alpha} = \frac{1}{\Delta \sum_{i=1}^N \frac{1}{x_{(i-1)\Delta}}} \left( \sum_{i=1}^N \frac{x_i \Delta}{x_{(i-1)\Delta}} - N + \hat{\beta} \Delta N \right).
\]

(F.4)

Furthermore, we get

\[
\frac{\partial L(\alpha, \beta)}{\partial \beta} \approx - \sum_{i=1}^N (x_i \Delta - x_{(i-1)\Delta}) + \alpha N \Delta - \sum_{i=1}^N x_{(i-1)\Delta} \Delta
\]

\[
= 0.
\]

Using (F.4) for \( \hat{\alpha} \) gives

\[
\frac{\partial L(\alpha, \beta)}{\partial \beta} = - \sum_{i=1}^N (x_i \Delta - x_{(i-1)\Delta}) + \sum_{i=1}^N \frac{1}{x_{(i-1)\Delta}} \left( \sum_{i=1}^N \frac{x_i \Delta}{x_{(i-1)\Delta}} - N + \hat{\beta} \Delta N \right) - \hat{\beta} \Delta \sum_{i=1}^N x_{(i-1)\Delta}
\]

\[
= -(x_N \Delta - x_0) + \sum_{i=1}^N \frac{\hat{\alpha} \Delta N^2}{x_{(i-1)\Delta}} - \sum_{i=1}^N \frac{1}{x_{(i-1)\Delta}} - \sum_{i=1}^N \frac{1}{x_{(i-1)\Delta}} - \hat{\beta} \Delta \sum_{i=1}^N x_{(i-1)\Delta}.
\]
Solving for $\hat{\beta}$ gives

$$\hat{\beta} = \frac{1}{\Delta} \left( \frac{N \sum_{i=1}^{N} \frac{x_{i\Delta}}{x_{(i-1)\delta}} - N^2 - (x_{N\delta} - x_0) \sum_{i=1}^{N} \frac{1}{x_{(i-1)\Delta}}}{\sum_{i=1}^{N} x_{(i-1)\Delta} \sum_{i=1}^{N} \frac{1}{x_{(i-1)\Delta}} - N^2} \right). \quad (F.5)$$
References


REFERENCES


F.3. REFERENCES


