Pricing Credit Risky Swaps In An Affine Framework

Authors: Adam Polack & Adam Schydt

Superviser: Mads Stenbo-Nielsen

Thesis submitted: August 4, 2011

Cand.merc(mat.)

Master Thesis in partial fulfilment of the Master degree in Business Administration and Management Science

Department of Finance

Abstract

This master's thesis focuses on the pricing of counterparty credit risky claims. The subject has grown increasingly popular in recent years due to the global recognition that participants in the over-the-counter market indeed may default.

We argue that the price of *any* risky claim can be *linearly decomposed* into the price of a risk free claim minus a certain risk premium called the Credit Value Adjustment. We choose to take a closer look at *interest rate swaps*, and show that under the assumption of independence between counterparty default risk and interest rates, the risky swap price is given as an infinite sum of *swaption prices* multiplied by marginal default probabilities. So, while traditional pricing of a counterparty risk free swap requires nothing more than a zero coupon term structure, we argue that the pricing of a *risky* swap requires a lot more.

We choose to explore the swaption pricing model proposed in Pelsser and Schrager (2006). By using *Fourier inversion* techniques, we show that the model is able to generate prices possibly stipulating the dynamics of *any* affine interest rate model. In terms of modelling the default risk, we choose to examine the *intensity model* proposed in Lando (1998). Since the driver of the intensity model, the *Cox-process*, can *also* be assumed affine, we choose to use affine models for both purposes. This has the benefits of providing semi-analytical solutions to zero coupon rates, default probabilities, and especially the *characteristic function* which is the focal point in terms of the Fourier inversion which ultimately provides the swaption price.

In our applicational part, we choose to specify both of the affine models in compliance with the one-factor CIR model (Cox et al., 1985) to set a basic example. However, we emphasize that extending the framework to a higher number of risk factors is fairly straightforward due to the *flexibility* of both models. We use different estimation methods and study the interplay between the two models in different economical data that showcases tendencies pre- and post crisis, respectively. This investigation is conducted for two different counterparties; the global bank HSBC and the automobile manufacturer Fiat.

Preface

We would like to thank our supervisor, Mads Stenbo-Nielsen, for his excellent guidance and his willingness to participate in discussions throughout the process. For that, we are grateful. In addition, we would like to thank Ketil Skotte for proofreading this master's thesis.

Copenhagen, August 2011
Adam Polack & Adam Schydt

Contents

Chapter 1

Introduction

Financial interaction takes place in different market places. While many standardized products are traded through an exchange which helps create a liquid, transparent and almost credit risk free market, the need for tailored contracts has given rise to another market; the over-the-counter (OTC) market. In this market, *derivatives* are traded *bilaterally* between counterparties. According to the Bank of International Settlements, the notional amount outstanding in the global OTC market reached USD $601,048$ billion at the end of 2010. Even though the OTC market decreased a bit during the recent financial crisis, the market was still bigger than ever at the end of 2010.

The crisis, however, shed new light on one of the possible downsides of trading in the OTC market; the counter party credit risk. While pre-crisis market practice typically was to believe that major counterparties would not (be allowed to) default, the default of corporations such as Lehman Brothers in 2008 ruptured this illusion. Participants in the OTC market now have to take counterparty credit risk more seriously in their risk management as well as in their assessment of market prices. Furthermore, market *regulators* have increasingly required OTC participants to keep reserves in the case of defaults of clients.

The most traded contract in the OTC market is the plain vanilla fixed-for-floating *interest rate swap*. Notionals of USD $364,378$ billion, or more than 60% of the entire OTC market, were placed in the market of interest rate swaps (henceforth simply *swaps*) in late 2010. Unlike many other derivatives, e.g. options, a swap has the feature of causing counterparty credit risk to *both* sides of the contract since neither counterparty knows if the swap will become an asset or a liability. This great uncertainty of the future value distinguishes the swap market from e.g. the bond market, and generally causes the associated credit risk in the swap market to be lower compared to the bond market since only assets can carry default risk. Even though notional outstanding amounts provide a measure of market size and *not* of the market value of the contracts, the numbers clearly indicate that just a small risk premium in the market could provide significant changes in the pricing of interest rate swaps since these have traditionally been priced in a way that has been neglecting counterparty credit risk. Considering the financial turmoil of the last couple of years, along with the tremendous market size, the *quantification* of counterparty risk is of greater importance than ever.

1.1 Formulation of the Problem

In this exposition we seek to quantify asset pricing. We wish to give a thorough introduction to *risk-neutral* pricing and pricing involving other measures than the risk-neutral one. In particular, we wish to show how *change of numeraire* techniques allow shifting between different price measures since pricing under the *swap measure* becomes necessary in order to apply the swaption model proposed in Pelsser and Schrager (2006). By carefully laying down the very general asset pricing framework, we can extrapolate to the field of counterparty risk since the pricing of *risky* swaps is our ultimate goal.

While traditional swap pricing simply requires a zero coupon bond term structure, it becomes evident that a credit risky swap requires a lot more. Following Brigo and Masetti (2005), we show that the price of a generic credit risky claim can be *linearly decomposed* into the *risk-free* price minus a *Credit Value Adjustment* (henceforth simply a CVA). Turning to the task of swap pricing, we subsequently show that by assuming independence between interest rates and default probability, the generic case reduces to the task of pricing a risk-free swap minus the sum of an infinite amount of *swaption prices*, each multiplied with the probability of default during the lifetime of the swaption. By using standard approximation techniques, our definitive task thus becomes twofold; we need to *both* develop the toolbox necessary in order to price swaptions efficiently as well as derive default probabilities on a given counterparty. With this goal in mind the challenge (beauty) of this master's thesis therefore becomes to explore the interplay of two very different subjects in the financial literature.

The question is now which models to use. In our pursuit of maintaining generality, we choose to explore the larger class of *affine* models. This model family has several advantages. First and foremost, it is applicable both in the swaption pricing framework *and* in the default framework. Secondly, by making the relevant assumptions it provides (semi-) analytical solutions to i) zero coupon bond prices, ii) default probabilities and iii) the *characteristic function* of the probability distribution in question. The latter being an essential focal point in our specific choice of swaption pricing model which is the one proposed in Pelsser and Schrager (2006). The model indeed uses an affine assumption of the unobservable short-rate to derive the dynamics of the *swap rate*. By using *change of numeraire* techniques along with low variance martingale (LVM) approximations, it is possible to derive the *affine* dynamics under the so-called *swap measure*. Finally, by applying an option pricing method using a *Fourier inversion* involving the characteristic function, as suggested in Carr and Madan (1998), we are able to establish a swaption pricing formula. We emphasize that by using this approach, many classical interest rate models may be used to specify the affine model, e.g. the models by Merton, Vasicek and Cox, Ingersroll and Ross are all *directly* applicable.

Turning to the task of deriving default probabilities, we choose to examine the class of intensity models as suggested in Lando (1998) and Lando (2004). Since the underlying driver of the intensity models, the *Cox process*, represents a greater class of processes than the affine ones, it allows for an affine specification of the intensity process. Recalling that our goal for the model is to obtain default probabilities, we show how it can be applied to the pricing of *Credit Default Swaps* (CDS) taking a zero coupon term structure and a *default term structure* as input. More importantly, we show how the model allows for an *inversion* so that a default term structure may be produced given a *CDS term structure*. We underline that by explicitly specifying the intensity process, one may extrapolate out of the timespan of the observed CDS term structure in terms of deriving default probabilities.

When the theoretical setting for the two models is established, we will show how the models may be used to quantify a CVA for two given counterparties, HSBC and Fiat, in different economic scenarios. For both models, we choose to rely on the *one-factor CIR model* as proposed in Cox et al. (1985). We rely on a *maximum likelihood method* in order to estimate the central interest rate model as proposed by Kladivko (2004). Following Mortensen (2006), we use a more classical *RMSE minimization routine* in order to estimate our default model.

1.2 Delimitation

We have chosen to focus on a rather theoretical approach to the pricing of CVA's to attach emphasis to the *interplay* between the two models. As a consequence, less weight will be given to the empirical investigation of the CVA's on swaps and more weight will be given to model frameworks and the *implementation* of the models. Our master's thesis should be regarded as a display of how one *might* do in practice. With this in mind, we will in the following section walk through some of the major decisions and consequently their attached delimitations.

We choose to rely on the independence assumption between the interest rates and default probabilities so that two independent models may be established. By incorporating *interdependence*, one would have to apply one *joined* model which would obviously require a more complex setup. We will, however, *discuss* the incorporation of interdependence in the final part of the exposition where concepts such as wrong-way and right-way risks are central.

We undertake the approach of looking at a single swap contract. This might seem counterintuitive in relation to the fact that portfolios of swaps may mitigate some of the counterparty credit risk due to *netting effects*. Our argument is, however, that by computing the CVA on every single swap in the portfolio, one may easily take into account the possible netting effects by offsetting different exposures since the CVA is *additive*[1]. Incorporating *one* big model to single-handedly quantifying the CVA of a portfolio of swaps, as seen in e.g. Brigo and Masetti (2005), would of course complicate matters significantly since one would have to address correlations, etc.

Our choice of counterparty risk assumptions will be confined to the *unilateral* case, i.e. only a single counterparty would be assumed to have a positive probability of default. While every entity is indeed subject to the risk of default, we undertake this assumption as a result of several reasons. As argued in Chapter 2, *bilateral* default risk may be difficult for the counterparties to agree upon since the larger counterparty may see a possibility of calling the shots and since *information asymmetry* may play a role. Furthermore, expanding to a bilateral model should essentially only require an assessment of the remaining counterparty's default risk, which for our case would be a repetitive use of the intensity model.

We have chosen the one-factor CIR interest rate model. Since the setup adapted from Pelsser and Schrager (2006) allows *direct* usage of *any* affine model, we could easily have applied a more complex model with more risk factors. This would, however, complicate the estimation and the *interpretation* of the model and focus would probably have been shifted away from counterparty risk and be pushed more towards which interest rate model to use. As an *example*, we choose to apply the CIR model which has been popular due to a sound degree of interpretation and various other key properties which will be discussed in Chapter 6. Using the CIR model, we put

[1] See e.g. Gregory (2010).

an effort in testing the LVM approximation. Furthermore, we use *Monte Carlo* methods to verify the analytical characteristic function.

Finally, we find it appropriate to state that our approach to interest rate model estimation might differ from the typical routines used in practice. This applies particularly in conjunction with our swaption pricing model for which market practice probably would be to fit the model to *observed* swaption prices[2]. However, since *both* the swaption- and the default model rely on the modelled interest rates we choose to estimate according to observed interest rates and *not* to observed swaption prices.

1.3 Structure

Our exposition is structured in three different parts. The first part, *The Financial Background*, has the purpose of *describing* key concepts and central contracts. The second part, *Quantitative Methods*, turn to a general pricing framework focussing on swaption pricing and intensity modelling. Finally, the third part, *Applications*, displays implementations of the frameworks substantiated in the second part.

Each part is divided into several chapters. Part I consists of two chapters. Chapter 2 introduces counterparty credit risk and Chapter 3 presents an overview of key interest rates, bonds, and derivatives.

Moving on to Part II, the reader will find six chapters. In Chapter 4 we introduce the general framework for valuating contingent claims and in Chapter 5 we turn to more complex pricing methods using Fourier inversion involving the characteristic function. Chapter 6 concentrates on the class of affine models and Chapter 7 addresses the quantitative aspects of pricing counterparty credit risk. In Chapter 8 we focus on the swaption pricing framework and in Chapter 9 we elaborate on intensity models and the theoretical derivation of default probabilities.

Part III consists of the final seven chapters. Chapter 10 briefly introduces our plans for the third part. Chapter 11 and Chapter 12 concern the estimation of the interest rate and the default model, respectively. Chapter 13 deals with implementation and verification of the swaption pricing model and Chapter 14 assesses the actual CVA prices. In Chapter 15 we will discuss possible intuitive extensions of the framework and finally Chapter 16 concludes upon the master's thesis as a whole.

[2]Instead of fitting to actual prices, one would typically fit to the *implied Black volatilities*, which carry the same information as market prices.

Part I

The Financial Background

Chapter 2

Introducing Counterparty Risk

Taking a glance at Part I, we will begin by discussing aspects of counterparty credit risk. Our first objective is to give a concise feed to the subject of counterparty risk. The subject is thoroughly described in the recent book by Gregory (2010). The book is a milestone since counterparty risk is generally very sparsely described in the literature due to the world's very late acknowledgment of the importance of the field. Other more concise introductory sources do however exist to some extent, see e.g. Pykhtin and Zhu (2007), Canabarro and Duffie (2003), Boettcher (2010) Shahram et al. (2010) and Stein and Lee (2010).

After the imposition in counterparty risk is completed, we will present ways to control and mitigate counterparty risk. Subsequently, we will turn to our main focal point in counterparty risk; the Credit Value Adjustment. Finally, we will outline the counterparty risk aspects that apply to interest rate swaps.

2.1 Defining Counterparty Credit Risk

Credit risk is defined as the risk of a counterparty not meeting all of his obligations. Credit risk is key in, for instance, the lending markets where the lending period and the credit quality of a *debt issuer* typically determine the price of the loan, i.e. the interest rate *demanded* by the investor to compensate for the credit risk. The longer the lending period and the worse the credit quality of the issuer, the higher the required interest rate. By shifting point of view to the issuers side, we emphasize that the credit quality of *the investor* does not affect the interest rate. This is due to the fact that the bond issuer has no *credit exposure* in the agreement, i.e. the bond issuer is not affected by a default by the investor since the investor will never possess any debt to the bond issuer.

Counterparty risk is according to Gregory (2010) defined as the overall risk generated by two parties trading with each other. So while credit risk in the lending market context of the above example is solely possessed by the investor who bears the risk of not receiving his claims, the counterparty risk consists of all risk sources of a trade and affects both parties. Gregory (2010) argues, however, that the main component of counterparty risk indeed is the credit risk. Hence, the expressions counterparty risk and counterparty credit risk will be used interchangeably.

While in a two-party contract as in e.g. the loan market, there is no doubt where the credit risk lies, things get more complicated when considering contracts involving three or more parties. For a derivative, e.g. a call option, the credit risk is said to lie in the risk of a default by the entity which the call option is written on. The *counterparty* credit risk, however, lies in the risk of the counterparty not being able to deliver the underlying asset when the option is exercised. Counterparty credit risk thus always lies between the two *trading* parties, regardless of what sort of contract they agree upon.

Continuing the example regarding the call option, one may argue that the size of the counterparty risk depends on which *market* the option is traded on. If the option is traded through a central counterparty, e.g. an exchange, one will typically regard the counterparty risk as being non-existing since the only cause of concern is the solvency of the exchange itself. However, if the option is traded OTC, the significance of counterparty risk is of a much larger magnitude. According to Pykhtin and Zhu (2007), the following classes of financial products are the primary subject to real counterparty risk:

- OTC derivatives: forwards, swaps, options, credit derivatives, etc;
- Securities financing transactions: repos, reverse repos, securities borrowing and lending.

Comparing the two classes, the former category is perhaps the more significant one due to market size and product diversity. However, incorporating counterparty risk in the derivatives market has only grown increasingly popular in recent years after default-free counterparty illusion was shattered.

There are two major reasons why counterparty credit risk on derivatives distinguishes itself from the credit risk on bonds:

- The exposure is unknown.
- The risk is often bilateral.

Commenting on the first point, it is important to state that the exposure also is not exactly known in a bond trade because of possible changes in the interest term structure. It is however known to some extent since the principal (and sometimes also the coupon) is fixed. In derivatives, however, the exposure can often be limitless and–as the second point states–with unknown sign. Considering some examples, a call option has unknown (and limitless) exposure, but the credit risk is (typically) only carried by the option investor. A forward or a swap will however both be subject to unknown exposure and unknown sign of the credit risk since the future value might be either positive or negative. It is albeit important to realize, that while the exposure in derivatives might be both unknown and without limit, the counterparty risk in each contract is still often of a smaller magnitude compared to the bond market. This stems from the fact that the *whole* principal is at risk in a loan which is contrary to, say, a swap in which the principal is never exchanged.

2.2 Asset or Liability?

In the context of counterparty risk on derivatives, the paramount attribute of a derivative is whether it is an asset or a liability since only an asset can carry counterparty risk. To obtain more insight into the loss arising from a default, we position ourselves in a derivative contract between a bank and a counterparty and assume that the counterparty has just defaulted. After the default event, the bank must immediately close its positions with the defaulting counterparty. We assume that the bank enters a similar contract with another counterparty in order to maintain its market position, i.e. the market position of the bank is unchanged. We then ask ourselves if the contract, from the banks point of view, is an asset or a liability, i.e. if the contract value is negative or positive.

i) If the contract has a *negative* value, the bank

- closes out the position by paying the defaulting counterparty the market value of the contract;
- enters into a similar contract with another counterparty and receives the market value of the contract since buying a contract with negative value makes one *receive* a cashflow;
- has a net loss of zero.

ii) If the contract has a *positive* value, the bank

- closes out the position but receives nothing from the defaulting counterparty (by assuming zero recovery);
- enters into a similar contract with another counterparty and pays the market value of the contract;
- has a net loss equal to the market value of the contract.

In summary, the credit exposure of a bank having a single derivative contract with a counterparty is the maximum of the contract value and zero. By denoting the time t value of a contract as $\pi(t)$, the contract level exposure is then given by

$$\mathrm{Ex}(t) = \max\Big(\pi(t), 0\Big). \tag{2.1}$$

Since $\pi(t)$ evolves with uncertainty over time, only the current exposure is known with certainty as stated previously. Given this uncertainty, (2.1) shows that the exposure of a contract exhibits an option-like expression. This expression is pivotal in *quantifying* counterparty risk and will be investigated further in Chapter 7.

2.3 Mitigating Counterparty Risk

There are many ways to mitigate or limit counterparty risk. As mentioned in Gregory (2010), the most common method historically has been to trade only with the most financially sound banks. The method is unfortunately dangerous if one thereby assumes that some counterparties cannot (or will not be allowed to) default. Other forms of risk mitigation focus on reducing and controlling the exposure. The most important tools are:

- *Diversification.* Spreading exposure across different counterparties to lower the risk of many defaults.
- *Netting/set-off.* Offsetting positive and negative contract values with the same counterparty in case of a default. Netting requires that a legally binding netting agreement[1] has been made and that the derivative can have both positive and negative values. Netting is essentially "free" but it gives preferential benefit to derivatives counterparties at the expense of other creditors, e.g. bond- and shareholders. Contrary to the diversification philosophy, netting effects are actually an argument supporting that one should trade *more* with a given counterparty in order to obtain a greater *mirroring* effect.

[1] A netting agreement is an optional part of the *ISDA master agreement* discussed further in Chapter 3.

- *Collateralization/Margining.* Holding cash or securities against an exposure. Essentially, the counterparties sign an agreement which states that collateral must be posted when exposure is different from zero[2]. To keep operational costs under control, posting of collateral will occur at a time frequency specified in the agreement. While collateralization (in principle) can completely neutralize an exposure. It is important to note that the method typically incepts new risk sources in form of *liquidity risk*[3] and *legal risk*[4]. Furthermore, aspects such as *rehypothecation*[5] may complicate matters, which it did in e.g. the bankruptcy of Lehman Brothers in 2008. It is however not uncommon that many counterparties choose not to trade on an uncollateralized basis[6]. Collateral is also key in the increasingly popular credit support annex (CSA) contracts in which the difference between the market value of the contracts and a mandatory margin account continuously determines if more collateral must be posted to neutralize the difference. The CSA agreement is examined in detail in Shahram et al. (2010).

- *Hedging.* Trading credit derivatives, typically credit default swaps, to reduce exposure. This option is a very direct way of trying to eliminate counterparty risk, essentially allowing an exposure to reach zero. The method comes with a direct cost (the premium in the credit derivatives) and requires that a credit derivative written on the counterparty in question exists. While the credit derivative might neutralise a given exposure, the instrument itself is also vitiated with the default risk of the third party (the credit derivative seller).

- *Close-outs.* Letting settlements occur at their *mark-to-market* value more frequently than they normally would. A classic example of a built-in close-out feature is in the futures market where settlement occurs daily. By using a close-out agreement, the exposure is limited to overnight effects. The downside is an increase in transactions and thereby presumably an increase in transaction costs.

- *Credit triggers.* Agreeing with the counterparty to set a pre-specified *credit-trigger* rating below the current rating of the counterparty. If the counterparty's rating deteriorates to the trigger level (or below) before default or maturity, the investor has the right to settle the deal with the counterparty at mark-to-market. Credit triggers can be set to both sides of the investor and the counterparty and are studied in detail in Yi (2010).

As argued in Gregory (2010), one should be careful to not blindly commend *any* mitigation of default risk. Reason being, that risk mitigation may cause the market to develop too fast and thereby reaching a dangerous size. By using the above mentioned methods, the risk *may* be mitigated and thus neutralized. The risk *sources*, however, are not neutralized. Furthermore, many of the stated techniques can be applied without even assessing the size of the counterparty credit risk of the company[7], and methods such as diversification may be used to obtain other goals than mitigating counterparty risk.

[2] Strictly speaking, this is only the case in a *two-way* agreement. In a one-way agreement, only one of the parties is required to post collateral when the exposure (driven by the market value) changes in favor of the other counterparty.

[3] Liquidity risk is the risk that a given security or asset cannot be traded quickly enough in the market.

[4] Legal risk is the risk that a counterparty is not legally able to enter into a contract.

[5] Rehypothecation occurs when banks re-use the collateral posted by clients to back the bank's own trades and borrowings.

[6] By the end of 2010 the total amount of collateral used in all OTC derivatives transactions was reported to be $3.5 trillion, see ISDA's margin survey, ISDA (2010).

[7] Collateralization and hedging may work best when counterparty risk is properly monitored

2.4 The Credit Value Adjustment

While the previously mentioned methods of mitigating counterparty risk indeed dim the counterparty risk, they are all based upon a binary decision making process and they all have their flaws. By pricing counterparty risk, one can move beyond a decision making process and instead concentrate on whether a derivative transaction is profitable once the counterparty risk has been priced in. As we will show in Chapter 7, the price of a counterparty risky derivative can be *linearly decomposed* into the (counterparty) risk free price minus a credit value adjustment; the CVA. The downside of this approach is clearly (as we will also see in Chapter 7) that the pricing of the CVA might be complicated to assess and/or model dependent. Furthermore, a precise CVA pricing should account for all aspects affecting the counterparty risk, including (but not necessarily limited to)

- the default probability of the counterparty;
- the transaction in question;
- netting of other transactions with the same counterparty;
- collateralization and hedging aspects;
- recovery;
- the default probability of the bank itself.

The last bullet may appear conspicuous. The reason for its importance stems from the *bilateral* nature of counterparty risk; it is natural to assume that both counterparties will/should assess the credit quality of the parties. This may give birth to a *credit game* in which the two counterparties try to come to an agreeable assessment of each other's credit quality which (if agreed) allows a transaction to take place. This *puzzle* can be difficult to solve since there might be *information asymmetry*. Hence, it might be the case that the first counterparty will propose one CVA but that the second counterparty will propose a different one. This could ultimately result in a negation of a trade. Note that the CVA can presume both positive and negative values in the bilateral case which can indeed cause a risky derivative to be more expensive than a risk free one! In reality, a critique of the bilateral setup has furthermore been linked to the fact that the price of a derivative should be associated with a *hedge*. But there is (in June 2011) no realistic hedge an investor can establish for his own default risk. As mentioned in Stein and Lee (2010), accounting boards have been lobbied to reject bilateral CVA as an acceptable approach since the hedging issues argue against the bilateral CVA as a reasonable market price. The bilateral setup is examined in detail in e.g. Brigo and Capponi (2009), Brigo et al. (2010), Haase et al. (2010), and specifically in a rating-based setup in Huge and Lando (1999).

By assuming that the bank is regarded as being risk free by the counterparty, the approach is reduced to the *unilateral* case, which will constitute our choice of approach. A reason for this approach could be justified by either a significant difference in the parties' credit quality which "negates" the bank's credit quality or simply by excluding the bilateral approach due to the discussed reasons above. We will generally assume that a given position is *uncovered*, i.e. established without use of collateral.

2.5 Counterparty Risk in Interest Rate Swaps

The literature on the default risk in swaps is again scarce. The sources that we have used are all written before 2007, but most findings are actually just more underlined in the economy of 2011, i.e. the default risk in a typical swap is probably assumed greater now than it was in the time span prior to the crisis, say 1990-2007. The imperative articles are (to our knowledge) Sorensen and Bollier (1994), Cooper and Mello (1991), Duffie and Huang (1996), and Mozumdar (1999).

Turning to the risk attributes of a swap, a contract that is described in detail in Chapter 3, we can characterize the contract using the characteristics mentioned in the previous section. First, since a swap participant is simultaneously long (short) the receiving leg and short (long) the paying leg, the swap can be both an asset or a liability. Second, a swap has unknown exposure, i.e. *if* the swap becomes an asset we do not know how much profit it will provide since the profit moves with future interest rates. Third, the risk is bilateral, i.e. both parties in the swap are vulnerable to the other party's default upon inception since the swap has the *potential* of being an asset. As time passes, the counterparty risk will still be bilateral since both parties have a chance of receiving a net payment at the next settlement day. Fourth, as noted in Cooper and Mello (1991), the *recovery* in swaps (as defined in the ISDA Master agreement) differs from recovery in the bond market since swap recovery is paid on the whole market value, whereas recovery on a defaulted bond is paid only on the principle (future coupon payments are lost).

To control counterparty risk in swap agreements, the earlier mentioned tools can all be applied. Especially the netting agreements are obvious due to the bilateral nature of swaps.

Moving on to the CVA of a swap, we can either characterize it as an *upfront* payment or a fixed-rate adjustment that compensates the swap parties for default risk. The major considerations for pricing the default risk should include a combination of the credit conditions of the parties, the actual level of interest rates in order to find the present value of a future exposure, and their existing swap portfolios with each other in case of a netting agreement being reached. However, as noted in Sorensen and Bollier (1994), swap exposure is also highly dependent on the *shape* and the *volatility* of the swap rate–a size that we derive in Chapter 3. The influence of the volatility can be verified simply by considering the option-like expression given by (2.1) on page 8; if the floating rate volatility is high, the future swap value also has high volatility and the probability of a high exposure rises. On the contrary, if the volatility of the floating rate is low, then the probability of a high exposure also becomes low. Pricing of a swap CVA should thus *jointly* model the probability of the counterparty defaulting and the cost of the default to the solvent party. However, assuming independence between exposure and default probability allows for a separation into two autonomous models. As mentioned, we will investigate the mathematics of risky swap pricing in Chapter 7.

Chapter 3

Interest Rates and Derivatives

In this chapter we will construct the necessary toolbox in order to precisely define and understand counterparty *risk free* interest rate swaps. The toolbox consists of basic financial assets and terms such as zero coupon rates, zero coupon bonds, forward rates, and xIBOR rates. After completion of the toolbox, we will turn to swaps which constitute the foundation of our exposition. This is the reason why the concept of such products will be described in detail. We will expand this chapter so that it elaborates on swaptions as well since swaption pricing is necessary in order to price a credit *risky* swap. Furthermore, Credit Default Swaps will be covered since they are fundamental in deriving default probabilities which again are required to price a credit risky swap.

3.1 Interest Rate Basics

3.1.1 Discount Factors and Zero Coupon Rates

A (credit) risk free *zero coupon bond* (ZCB) is defined as a contract that, with certainty, pays *a single* amount in a given currency at maturity. Unless stated otherwise, we will assume that the face value of a ZCB is equal to 1. Formally, we denote the price at time t for an observed ZCB maturing at time T by $D(t,T)$. Because of the absence of credit risk, we have that $D(T,T) = 1$. If we further assume that money at this point of time is worth more than money tomorrow, it must hold that $D(t,T) \leq 1 \; \forall \; t \leq T$. Since we assume that the ZCB is risk free, the price is a pure measure of the *value* of a future unit payment that can be scaled to fit any future cash flow. Hence, the price of a ZCB is often referred to as a *discount factor*. Fixing time t and varying maturity T thus give us all the discount factors at time t making it possible to value *all* future cash flows.

A simple non-arbitrage argument can show that

$$D(t,T) = \frac{D(0,T)}{D(0,t)}. \tag{3.1}$$

Hence, the *forward price* $D(t,T)$ for $t > 0$ can still be calculated at time 0, assuming that we can observe $D(0,T)$ for different values of T in the market. Given the risk free interest rate $r(t,T)$ between time t and T, we have the following price of a ZCB:

$$D(t,T) = \exp\Big(-r(t,T)(T-t)\Big) \tag{3.2}$$

Here, we have assumed that interest rates are continuously compounded[1]. Since $D(t,T)$ is monotonically diminishing in $r(t,T)$, there is a unique one-to-one relationship between ZCBs and *zero coupon interest rates*–also known as *spot rates*. If we (again) fix t and vary T, we can now look at different interest rates that correspond to the equivalent discount factors. This mapping $T \to r(t,T)$ is called a zero coupon *yield curve* or a zero coupon *term structure*. Note, that the functions $T \to D(t,T)$ and $T \to r(t,T)$ carry exactly the same information.

3.1.2 Forward Rates

While a ZCB reflects the price on a loan between today and a given future date, a *forward rate* reflects the price on a loan between two future dates. Zero coupon rates can be thought of as the average rate of return over some time period. Decomposing such a period return thus leaves us with a collection of marginal returns. These returns are known as the forward rates since they measure the incremental interest over some future period.

To define a forward rate mathematically, we will fix three time values t, T and S, so that $t \leq T \leq S$. Looking at a ZCB maturing at time S, discounted in two steps, we must have the following relationship:

$$
\begin{aligned}
D(t,S) &= D(t,T)D(T,S) && (3.3)\\
&= D(t,T)\exp(-f(t;T,S)(S-T)) && (3.4)
\end{aligned}
$$

From this expression it is clear that $f(t;T,S)$ is a forward interest rate observed at time t and applied between time T and S. Isolating $f(t;T,S)$ in equation (3.4) we get the relationship between forward rates and ZCBs which can be altered to the relationship between forward rates and spot rates:

$$
\begin{aligned}
f(t;T,S) &= -\frac{\log(D(t,S)) - \log(D(t,T))}{S-T} && (3.5)\\
&= \frac{r(t,S)(S-t) - r(t,T)(S-T)}{S-T} && (3.6)
\end{aligned}
$$

By letting $S \to T$ in equation (3.5), we get (by definition) the minus of the logarithm of the derivative of $D(t,T)$:

$$
f(t,T,T) = -\frac{\partial \log(D(t,T))}{\partial T} \tag{3.7}
$$

This limit expression is called the *instantaneous forward rate* seen from time t maturing at time T since it measures the interest rate that applies in an infinitesimal future time span.

Considering equation (3.5) it can be observed that forward rates carry the same information as the ZCB prices since the knowledge of ZCB prices implies the forward rates and vice versa. Hence, the information given in a term structure is equivalent to the information given in the corresponding ZCB prices which again equals the information given by the corresponding *forward curve*.

[1] We will always assume that interest rates and returns are continuously compounded unless stated otherwise. In forthcoming chapters we will often further assume that the interest rate is stochastic.

3.1.3 xIBOR Rates and Coverages

Most interest rate derivatives are written on a set of official floating interest rates called xIBOR rates. xIBOR is an abbreviation for the *x Interbank Offered Rate*, which is a reference rate at which banks offer to lend *unsecured* funds to other banks. The xIBOR rate is a filtered average of interbank deposit rates. The x is *usually*[2] linked to the capital of the country where the entity that fixes the rate resides. Typical examples are LIBOR, EURIBOR, and CIBOR that are interest rates fixed respectively by the British Bankers Association, The European Central Bank and the Danish Central Bank. Hence, the letter L is short for *London* and C is short of *Copenhagen.* The fixing methodology can differ from bank to bank, but all xIBOR rates are used in the same way as underlying in interest rate derivatives. Thus, we will (in order to follow standard practice in the financial literature) use the term xIBOR and LIBOR interchangeably. A LIBOR rate often has a code written after it, e.g. LIBOR6M. This number is a description of the maturity of the interest rate in question. The maturity of a LIBOR rate can range from a single business day to 12 months.

A LIBOR rate is reported using the *Money Market convention*[3], which means that the interest rate being paid at time T on some notional N simply is ΔNL, where L is the LIBOR rate. Here, Δ denotes the *coverage* which is sometimes simply called the *year fraction.* It is defined according to the given *day count convention.* We could calculate the coverage according to e.g. the Act/360 convention so that $\Delta = \frac{T-t}{360}$ for a time span of actual size $T - t$, where T and t are measured in days. Coverage may differ from country to country and from product to product. We will discuss coverage further in Section 3.2 regarding swaps.

Formally, a LIBOR rate can be defined as either a spot or a forward interest rate. Given the existence of a family of zero coupon bond prices for all relevant maturities, a simple non arbitrage argument can show that the *spot* LIBOR rate $L(0, T)$ between reset date $t = 0$ and settlement date T is

$$L(0,T) = \frac{1}{\Delta}\left(\frac{1}{D(0,T)} - 1\right). \tag{3.8}$$

By letting $L(t; T, T + \Delta)$ denote the *forward LIBOR rate* at time t between T and $T + \Delta$, a similar non-arbitrage argument can show that the so called Δ-*tenor* forward rate can be stated as follows:

$$L(t;T,T+\Delta) = \frac{1}{\Delta}\left(\frac{D(t,T)}{D(t,T+\Delta)} - 1\right) \tag{3.9}$$

Note, that the spot LIBOR rate is a special case of the forward LIBOR rate–just as we saw regarding zero coupon interest rates.

3.2 Interest Rate Swaps

3.2.1 Definition and Risk Free Pricing

Since swaps are traded OTC, they can be specified in a huge number of varieties to meet the specific needs of the counterparties. We will however confine ourselves to look at *plain vanilla fixed for floating rate* swaps that are defined in the following.

[2]The Norwegian interbank rate *NIBOR* is a counterexample.
[3]This is also sometimes called *simple interest.*

A swap is an agreement between two parties to exchange a stream of fixed payments for a stream of floating payments of interest rates on a prespecified notional between two dates at some prespecified frequency. The two payment streams are usually referred to as the *fixed*- and *floating leg*, respectively. The floating leg is typically, but not necessarily, linked to some LIBOR rate. For each counterparty, the position in the swap is denoted *relatively to the fixed leg*, e.g. a party that pays the floating rate, *receives* the fixed rate and has thus entered a receiver swap. In a swap agreement, the floating rate is set *before* each forthcoming period, i.e. the floating rate is fixed *in advance*. These dates are called the reset dates. Dates where there is an actual exchange of cash flows are called settlement days and settlements occur *in arrears*. We will assume that time intervals between payments are equally spaced. Another important remark on the swap jargon is the *price* which is defined as the *swap rate* since swaps are priced so that the initial value of each leg is zero. Figure 3.1 sums up the structure of a swap.

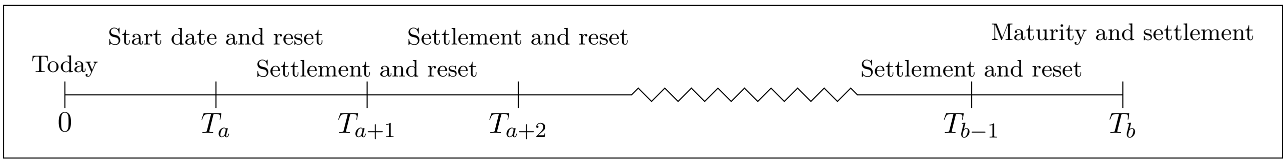


Figure 3.1: *Structure of an Interest Rate Swap.*

Formally, a *payer* swap with notional N, fixed interest rate κ, length n, start date T_a, maturity date T_b, reset dates T_i , $i = a, \dots, b-1$, settlement days T_i , $i = a+1, \dots, b$ written on the floating rate $L(t, T_i)$ leads to *receiving* the following series of cash flows:

$$X_{T_i} = N\Delta\Bigg(L(T_{i-1}, T_i) - \kappa\Bigg) \text{ for } i = a+1, \dots, b. \tag{3.10}$$

While most swaps are traded with spot start (a=0), the above equation applies to forward starting swaps as well.

Note that the cash flows of a payer swap resemble the cash flows of a portfolio consisting of a long position in a floating rate bond indexed on a Δ-tenor LIBOR rate and a short position in a non-callable fixed (annuity/bullet) bond with coupon rate κ (the converse holds for a receiver swap). This implies that we can replicate the portfolio by pricing a fixed and a floating rate bond. We remind ourselves[4] that the price $\pi_{fix}(t)$ of a fixed rate bond with constant coupon κ and payments at time $a+1, \dots, b$ is

$$\pi_{fix}(t) = N\left(D(t, T_n) + \kappa\Delta \sum_{i=a+1}^{b} D(t, T_i)\right), \tag{3.11}$$

and that the price $\pi_{float}(t)$ of a floating rate bond written on the LIBOR rate $L(T_{i-1}, T_i)$ is simply

$$\pi_{float}(t) = D(t, T_a). \tag{3.12}$$

By fixing the notional to some value, say $N = 1$, we denote the price of the payer swap with *some*

[4]See Björk (2009) page 358-359.

fixed rate κ, running between time T_a and T_b by

$$\pi^{pay}_{a,b}(t;\Delta,\kappa).$$

By *replicating* the payer swap, the value becomes

$$\pi^{pay}_{a,b}(t;\Delta,\kappa) = \underbrace{D(t,T_a)}_{\text{price of a floating rate bond}} - \underbrace{D(t,T_b) - \kappa\Delta\sum_{i=a+1}^{b} D(t,T_i)}_{\text{price of a fixed rate bond}}\ ,\ t \leq T_a. \tag{3.13}$$

We can now derive the *forward par swap rate* (henceforth simply the *swap rate*), which by definition is the rate κ that renders the value of the payer swap zero upon inception and simultaneously cause the initial value of the corresponding receiver swap to be zero. The derivation is done by isolating κ in the above equation. We hereby obtain the classic result

$$\kappa = \frac{D(t,T_a) - D(t,T_b)}{\Delta\sum_{a+1}^{b} D(t,T_i)} \equiv \frac{D(t,T_a) - D(t,T_b)}{P_{a+1,b}(t)} \equiv y_{a,b}(t). \tag{3.14}$$

It is thus stated that the swap rate, which is a function of t, T_a and T_b, is $y_{a,b}(t)$. The denominator in the above equation, $P_{a+1,b}(t)$, is called the *annuity factor* of the swap. This stems from the fact that the value of the denominator equals the value of an annuity with face value of one.

As will be a central point in this paper, we note that although future cash flows in the floating leg are uncertain, traditional risk free swap pricing only requires a zero coupon term structure which essentially is *observable* and thus provides *model independent* swap pricing.

3.2.2 Market Practice and Conventions

In practice, handling interest rate swaps is not quite as easy as the subsection above might suggest. This mainly stems from the fact that settlements cannot occur each and every day, and that different countries might have different conventions regarding the use of dates.

First and foremost, swap payments can only take place on *business days*. The definition of a business day is a non-weekend non-holiday, where the latter might differ from country to country. Furthermore, inception of the interest accrual is not always set as the trading date (time t), but is typically offset a couple of business days according to the convention in question, so that the accrual starts at the so called *spot date*.

The frequency of payments may also differ, not only from country to country, but often also between the fixed and floating leg. E.g. the payments from the fixed leg may be annual, whereas the payments from the floating leg are paid semi-annually. Equivalently, day-count conventions may also differ between countries and payment legs.

In the event of a settlement hitting a non-business day, we have to make use of some rule in order to adjust the non-business day to a business day. For that purpose a *rolling convention* is applied. The most used convention is the so-called *Modified Following* (MF) that *rolls* the holiday *forward* to the next business day, unless if this falls in the forthcoming month. In that case, the date is rolled *back* to the latest business day.

In summary, handling swaps (and most other financial contracts) requires a carefully planned schedule and an insight into the relevant conventions. In order to maintain a smooth working market, the ISDA[5] organization was founded in 1985. The organization determines, among other activities, the legal and policy rules of the OTC market. When two counterparties want to make a deal, they fill in an *ISDA Master Agreement* that, combined with a schedule, sets out the basic trading terms between the parties. Each subsequent trade is then recorded in a confirmation that references the Master Agreement and the schedule. One of the advantages of the Master Agreement is that once it is fulfilled, no further standard agreements have to be made when additional trades are made–the Master agreement automatically applies. Furthermore, the agreement makes sure that legal risk is kept at a minimum.

Letting "B" denote *business days*, "Q" denote *quarterly*, "S" denote *semi-annually* and "A" denote *annually*, the standard market conventions according to ISDA can be seen in Table 3.1. It is observed that GBP is in fact the only (major) currency in which fixed and floating leg conventions are the same. Furthermore, GBP is the only (major) currency for which spot start equals the actual trading day.

				Floating Leg		Fixed Leg	
Currency	Index Name	Spot Start	Roll	Frequency	Daycount	Frequency	Daycount
EUR	EURIBOR	2B	MF	S	Act/360	A	30/360
USD	USD LIBOR	2B	MF	Q	Act/360	S	30/360
GBP	GBP LIBOR	0B	MF	S	Act/365	S	Act/365
JPY	JPY LIBOR	2B	MF	S	Act/360	S	Act/365
SEK	STIBOR	2B	MF	Q	Act/360	A	30/360
NOK	NIBOR	2B	MF	S	Act/360	A	30/360
DKK	CIBOR	2B	MF	S	Act/360	A	30/360

Table 3.1: *Interest rate swap conventions according to ISDA.*

For more on dates, schedules and conventions, see Dalskov (2007).

3.3 Swaptions

3.3.1 Definition, notation and payoffs

The OTC contract known as a swap option or simply a *swaption* is an option on an interest rate swap, i.e. the swaption gives the owner the right, but not the obligation, to enter into a swap agreement with a predetermined fixed rate. Depending on whether the swaption gives the investor the right to pay fixed or floating in the underlying swap agreement, the swaption is referred to as being either a *payer* or *receiver* swaption. As with all options, the contract specifies when the swaption is exercisable. This could e.g. be only at maturity (a *European* swaption) or at multiple times (a *Bermudan* swaption). European swaptions are market standard, but Bermudan swaptions are also traded–especially in relation to the mortgage bond market. For our purpose, we will confine ourselves to only looking at European swaptions, and thus *a swaption* will refer to a European one. Upon inception, the swaption investor can choose between physical or cash settlement. If the settlement is physical, there is a true initiation of a swap if the swaption is exercised. If the settlement is cash, only a cash replication of the *swap value* will be exchanged upon swaption exercition. Swaptions are most often traded *at-the-money forward* (ATMF), which

[5]The International Swaps and Derivatives Association

means that the strike equals the *forward* swap rate. The concentration of trading around ATMF swaptions makes the instrument one of the most liquid derivatives traded in the financial markets. Furthermore, it is interesting that payer and receiver swaptions must have the same price when traded ATMF[6].

Formally, we will look at an option maturing at time T_n to enter into a swap maturing at time T_N. Such an instrument is called a "T_n *in* $(T_N - T_n)$" swaption, so that the two pronounced numbers indicate the time span of each of the two swaption elements; the option and the swap. Note how the option maturity equals the swap initiation time, so that the final time span of the contract is $T_n + T_N$. Fixing the payment days of the underlying swap as T_i , $i = n + 1, \ldots, N$[7], we can derive the time T_n value of a swaption, using the *swap* value given by equation (3.13) on page 16 and the swap rate given by (3.14):

$$\begin{aligned}\left[\pi^{pay}_{n,N}(T_n;\Delta,K)\right]^+ &= \left[D(T_n,T_n) - D(T_n,T_N) - K\Delta\sum_{i=n+1}^{N} D(t_n,T_i)\right]^+ \\ &= \left[y_{n,N}(T_n)P_{n+1,N}(T_n) - KP_{n+1,N}(T_n)\right]^+ \\ &= P_{n+1,N}(T_n)\left[y_{n,N}(T_n) - K\right]^+ \qquad (3.15)\end{aligned}$$

So, while the expression resembles a *swap* value at time T_n, it becomes pivotal in quantifying *swaptions* when placing ourselves somewhere *before* time T_n. One can say that we have set $K = \kappa$ since the swap rate is the strike on the swaption. Equation (3.15) shows that payoffs from a payer swaption can be seen as *proportional* to those from a *call option* on the swap rate. Equivalently it can be shown that the payoffs from a receiver swaption is proportional to a *put option* on the swap rate.

3.4 Credit Derivatives

The afore mentioned derivatives have payoffs that are directly linked to some interest rate. The value of credit derivatives is however not derived directly from an interest rate but instead from *the risk* of a *credit event* of an entity or asset. The definition of a credit event may vary between contracts but *a default*, i.e. a failure to meet a debt obligation, will always trigger a credit event. Other types of credit events *may be* restructuring of debt (the financial liabilities of the borrowing entity are changed) or acts from the government that somehow changes the timing of debt. By the 2003 ISDA Credit Derivatives Definitions, a restructuring credit event occurs if there is: (i) a reduction in the interest rate or in the amount of principal, (ii) a postponement or other deferral of dates for the payment of interest, principal, or premium, (iii) a change in the ranking

[6] This can be seen from the so-called *Swaption-Swap Parity* which basically states that

$$\pi_{payer} - \pi_{receiver} = \pi_{swap}$$

where the swaptions (left side) and the swap (right side) have exactly the same properties in terms of underlying interest rate, time to maturity, settlement days etc.

[7] This notation of swap payments is not the same as stated earlier. The notation is different to indicate that we are in a swaption setting.

in priority of payment of any obligation that causes subordination of it to other obligations or (iv) any change in the currency or composition of any payment of interest or principal.

While there is no doubt about ISDAs definition of a restructuring event, an actual contract may still be more or less independent of the restructuring legislation. Essentially, in specifying a certain contract one must declare which relationship to restructuring the contract should have. There are four possibilities:

1. *No Restructuring:* A restructuring is *not* regarded as a default.
2. *Full Restructuring:* In case of a restructuring, bonds of *any* maturity may be delivered.
3. *Modified Restructuring:* Bonds of a maturity up to *30 months* after the restructuring event may be delivered.
4. *Modified Modified Restructuring:* This is a softer version of Modified Restructuring in the sense that bonds of a maturity up to *60 months* after the restructuring event may be delivered. This option is often simply called *mod-mod*.

To the authors' knowledge, the Modified Modified Restructuring option is the most popular choice in Credit Default Swaps. This means that for these products restructuring will most often be regarded as a default. Furthermore, the credit products that we are using in our empirical investigations will be using the mod-mod option. For more details on restructuring see e.g. Whetten (2004).

The largest players in the credit markets are commercial banks. Traditionally, the business of a bank has involved credit risk as it originates loans to corporations. The credit market offers banks a way to transfer risk without removing assets from its balance sheet and without involving borrowers. Furthermore, a bank may use credit products to diversify its portfolio, which is often concentrated in certain industries or geographic areas. Surveys indicate (see e.g. Fitch, 2003) that banks are the *net buyers* of credit derivatives[8].

3.4.1 Credit Defaults Swaps

A single name Credit Default Swap (henceforth simply CDS) is the most fundamental credit derivative. It is agreed between two parties and provides insurance against a credit event of a third entity or *reference bond* on a prespecified face value. The two counterparties in a CDS are called the *protection buyer* (PB) and the *protection seller* (PS), respectively. The PS is thus required to *compensate* the PB in the case of a credit event by the given entity. On the other hand, the PB has to pay a (typically quarterly) running premium until a default happens, or the contract reaches maturity. The premium will often be called the *CDS price* or the *CDS quote* which indeed characterizes the value of the CDS since it (as other swaps) is priced so that the value upon initiation is zero. The compensation that the PB is given by the PS in case of a default can be specified as either a *physical* or a *cash* settlement. With a physical delivery, the PB is paid the face value of the reference asset and the PB must simultaneously deliver the defaulted asset. If a CDS is specified with a cash settlement, the PS must pay the difference between the face value and the market value right after the credit event. In case of both methods, one can say that the PS (in a credit event) buys the underlying bond for par rate. The two methods appear

[8]In 2003, global banks held gross bought positions of $1,553 billion in credit derivatives, with gross sold positions of $1,324 billion.

equally fair but the latter grew increasingly popular during the credit crisis since there can be more notional outstanding in CDS contracts than in the reference asset when CDS's are used for speculative purposes. Summing up, CDS's are differentiated with respect to maturity, type of settlement and the definition of a credit event.

Our use of CDS's will be limited to the goal of inferring default probabilities. In that context, it is important to keep in mind that a credit event is not solely linked to the probability of default. Debt restructuring will typically also trigger a credit event, even though the debt restructuring is not affecting the bond holders. In fact Berndt et al. (2006) find that the average premium for restructuring risk represents 6% to 8% of the CDS rate without restructuring. One argument supporting that CDS quotes are higher than what pure default risk can explain is as follows. When a restructuring happens, the bond rate can either be below par rate or not. If the bond rate is not below par rate, the PS buys the bond at its current market value and neither PS nor PB gains anything. However, if the bond trades below par rate the PS must still buy the bond at par rate. Consequently, the PB has a positive option-like relationship to a restructuring and the PS requires a higher premium.

A second argument that also supports larger CDS quotes than inferred by pure credit risk stems from the impact of counterparty credit risk in the CDS itself. The PS's future premiums will be lost if the PB defaults, but if PS defaults, the PB can just enter a new CDS with a maturity that equals the remaining maturity of the defaulted CDS. A simultaneous default of the PS and the reference entity could of course be disastrous for the PB, but that risk should be extremely low. In the CDS market, a running use of collateral is often required in order to compensate for the changing market value of the CDS so that losses for the PS are minimized.

Part II

Quantitative Methods

Chapter 4

The Theory of Derivatives Pricing

4.1 The Basic Framework and Important Definitions

4.1.1 The Framework

We consider an economy with continuous trading, taking place inside a finite horizon $[0, \tilde{T}]$ where $\tilde{T}$ is greater than any date of interest. No restrictions are made on short selling and we assume that there are no transaction costs associated with trading i.e. the market is frictionless. Uncertainty and information arrival are modelled by a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}(t), \mathbb{P})$ where Ω is the sample space, $\mathcal{F}$ is the σ-algebra on Ω and $\mathbb{P}$ denotes the probability measure on $(\Omega, \mathcal{F})$. Information is developed over time according to the filtration $\{\mathcal{F}(t), t \in [0, T]\}$ which represents a family of sub-σ-algebras of $\mathcal{F}$ satisfying $\mathcal{F}(s) \subseteq \mathcal{F}(t)$ for every $s \leq t$. So $\mathcal{F}(t)$ can be interpreted as the information available at time t. The economy consists of d non-dividend-paying assets with the price vector process $S(t) = (S_1(t), \dots, S_d(t))^\intercal$ and we assume that process $S(t)$ is adapted to $\mathcal{F}(t)$, so that $S(t)$ is observable at time t. We also require that the filtration satisfies the usual conditions.[1] In our setup the information is generated by a d-dimensional Standard Brownian motion $W^{\mathbb{P}}(t) = (W_1^{\mathbb{P}}(t), \dots, W_d^{\mathbb{P}}(t))^\intercal$ under $\mathbb{P}$ and the filtration we consider will be the one generated by $W^{\mathbb{P}}$, $\mathcal{F}(t) = \sigma\{W^{\mathbb{P}}(u), 0 \leq u \leq t\}$.

We assume that the asset price dynamics follow certain stochastic processes known as Itô processes[2] under the probability measure $\mathbb{P}$. Hence, the asset process $S_i(t)$ is the solution to the following stochastic differential equation (SDE)

$$dS_i(t) = \mu_i(S(t), t)dt + \sigma_i(S(t), t)^\intercal dW^{\mathbb{P}}(t), \tag{4.1}$$

which is just a short version of the following integral equation

$$S_i(t) = S_i(0) + \int_0^t \mu_i(S(u), u)du + \int_0^t \sigma_i(S(u), u)^\intercal dW^{\mathbb{P}}(u), \tag{4.2}$$

where $\mu_i : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\sigma_i : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^d$ for all i. Hence, μ_i is a scalar process and σ_i is a $\mathbb{R}^d$-valued process. The measure $\mathbb{P}$ is called the objective (or real-world) probability measure, since it is intended to describe the real-world probabilities. This implies that the SDE's in (4.1) describe the empirical dynamics of asset prices. Furthermore, we denote the instantaneous

[1] $\mathcal{F}(t)$ is right-continuous for all t and $\mathcal{F}(0)$ is the completion of the trivial σ-algebra, $\{\emptyset, \Omega\}$.

[2] For a full definition of a Itô process, see Shreve (2004).

covarians between assets as

$$\Sigma_{ij}(S(t),t) = \sigma_i(S(t),t)^\intercal \sigma_j(S(t),t),\ i,j = 1,\dots,d \tag{4.3}$$

We assume that μ and σ are adapted to $\mathcal{F}(t)$ and that the following regularity conditions hold for all $t \in [0,T]$,

$$\int_0^t |\mu(S(u),u)|du < \infty, \tag{4.4}$$
$$\int_0^t |\sigma(S(u),u)|^2 du < \infty, \tag{4.5}$$

a.s.[3] From (4.2) we see that the sample paths of S are continuous with probability one, so we assume that there are no jumps in asset prices.

4.1.2 Results Related to the Framework

An important concept closely connected to the theory of stochastic processes and one that has laid the foundation of modern financial theory is the concept of *martingales*:[4]

Definition 4.1. *(DEFINITION OF MARTINGALES)*
Let $(\Omega, \mathcal{F}, \mathcal{F}(t), \mathbb{P})$ be a filtered probability space and let $S(t)$ be a continuous (or discrete) time stochastic process adapted to the filtration $\mathcal{F}(t)$ with $\mathbb{E}^{\mathbb{P}}\Big[|S(t)|\Big] < \infty$. Then $S(t)$ is a martingale under the measure $\mathbb{P}$ if for every u and t with $0 \leq u \leq t$ it holds that,

$$S(u) = \mathbb{E}^{\mathbb{P}}\left[S(t)|\mathcal{F}(u)\right],\ a.s.$$

Technical comments. In this thesis paper we will sometimes use the notation $\mathbb{E}_t^{\mathbb{P}}\left[\cdot\right] = \mathbb{E}^{\mathbb{P}}\left[\cdot|\mathcal{F}(t)\right]$ depending on the specific situation.

If a stochastic process, according to the definition, satisfies the martingale property then, at all points in time, the expected change in the value of the process over any given future period is equal to zero. Hence, the expected future value of the process is equal to the current value of the process. Also notice that the concept of a martingale is connected to a specific probability measure.

Returning to the integral equation in (4.2), we see that the uncertainty in the price process $S_i(t)$ is driven by an integral w.r.t. a d-dimensional Brownian motion. This specific integral is called a stochastic integral or an Itô integral[5]. Stochastic integrals can be defined for very general processes. However, in this thesis paper we will only consider stochastic integrals where the integrator is a Brownian motion. Now, for given $s < t$ the stochastic integral is a random variable with the property that

$$\mathbb{E}\left[\int_s^t g(S(u),u)^\intercal dW^{\mathbb{P}}(u)\,\middle|\,\mathcal{F}(s)\right] = 0 \tag{4.6}$$

[3]In (4.5) we have defined $|\sigma(S(t),t)|^2 = tr\left(\sigma(S(t),t)\sigma(S(t),t)^\intercal\right)$.
[4]See Björk (2009), page 504.
[5]See Shreve (2004), chapter 4 for a complete discussion on Itô integrals.

for some integrand g satisfying the condition stated in (4.5). From this property, one can show that any stochastic process on the form $dS(t) = g(S(t), t)^\intercal dW^{\mathbb{P}}(t)$ where $g(S(t), t)$ satisfies the condition in (4.5) is a martingale. Hence, the process is a martingale if it has no drift-term.

More generally, let us place ourselves in the setup described above with a d-dimensional Brownian motion adapted to the filtered probability space and fix a vector process $h(t) = (h_1(t), \dots, h_d(t))^\intercal$ where $h(t)$ satisfies the condition in (4.5). If we then define a process $M(t)$ by

$$M(t) = M(0) + \int_0^t h(s)^\intercal dW^{\mathbb{P}}(s), \text{ for } t \in [0, T], \tag{4.7}$$

we know that $M(t)$ is a martingale. So, given certain integrability conditions, every stochastic integral w.r.t. a Brownian motion is a $\mathcal{F}(t)$-martingale. It turns out that the converse holds, i.e that every $\mathcal{F}(t)$-martingale $M(t)$ can be written on the form (4.7). This fact is given by the following theorem:[6]

Theorem 4.1. *(THE MARTINGALE REPRESENTATION THEOREM)*
Let $W^{\mathbb{P}}(t) = (W_1^{\mathbb{P}}(t), \dots, W_d^{\mathbb{P}}(t))^\intercal$ be a d-dimensional Brownian motion and assume that the filtration $\{\mathcal{F}(t), t \in [0, T]\}$ is defined as

$$\mathcal{F}(t) = \mathcal{F}^W(t), \; t \in [0, T].$$

Let M be any $\mathcal{F}(t)$-adapted martingale. Then there exists a uniquely determined $\mathcal{F}(t)$-adapted process $h(t) = (h_1(t), \dots, h_d(t))^\intercal$ such that M has the representation

$$M(t) = M(0) + \int_0^t h(u)^\intercal dW^{\mathbb{P}}(u), \; t \in [0, T].$$

We say that $M(t)$ has a stochastic integral representation w.r.t. the Brownian motion. So, from Theorem 4.1 we can conclude that in a framework where uncertainty is driven by Brownian motions, every martingale can be written as a stochastic integral w.r.t. the underlying Brownian motions.

When pricing derivatives we often have to work with functions of Itô processes and knowing the differential of these functions is therefore of critical importance. The differential of such functions is determined by the famous *Itô formula* which, loosely speaking, is the stochastic calculus counterpart of the chain rule in ordinary calculus. Since we will use this formula extensively we will now state the multidimensional version[7]:

[6] See Björk (2009), page 161.
[7] See Björk (2009), page 58.

Theorem 4.2. *(ITÔ'S FORMULA)*
Let $S(t) = (S_1(t), \dots, S_d(t))^\intercal$ where $S_i(t)$ is of the form (4.1). Define the process $Z(t) = f(t, S(t))$ where $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ is a $C^{1,2}$ mapping. Then the stochastic differential df is given by

$$df(S(t),t) = \frac{\partial f(S(t),t)}{\partial t}dt + \sum_{i=1}^{d} \frac{\partial f(S(t),t)}{\partial S_i}dS_i(t) + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 f(S(t),t)}{\partial S_i \partial S_j}dS_i(t)dS_j(t)$$

where

$$\begin{aligned} (dt)^2 &= 0 \\ dt \cdot dW^{\mathbb{P}}(t) &= 0 \\ (dW_i^{\mathbb{P}}(t))^2 &= dt,\ i = 1, \dots, d \\ dW_i^{\mathbb{P}}(t) \cdot dW_j^{\mathbb{P}}(t) &= 0,\ i \neq j. \end{aligned}$$

Technical comments. In this Theorem, we have assumed that $W_1^{\mathbb{P}}(t), \dots, W_d^{\mathbb{P}}(t)$ are independent Brownian motions, hence the condition $dW_i^{\mathbb{P}}(t) \cdot dW_j^{\mathbb{P}}(t) = 0,$ for $i \neq j$.

4.1.3 Self-Financing Trading Strategies and Arbitrage

We will now introduce the concept of a self-financing trading strategy and arbitrage, both fundamental concepts in the theory of derivatives. We place ourselves in the continuous time economy discussed in Subsection 4.1.1 and therefore assume that the asset dynamics evolve according to (4.1) along with the regularity conditions. This presentation follows the ideas in Glasserman (2003).

We define a portfolio as a vector $\theta \in \mathbb{R}^d$ where each θ_i represents the number of units held of the ith asset. From this definition, the value of the portfolio at time t is

$$\theta_1 S_1(t) + \dots + \theta_d S_d(t) = \theta^\intercal S(t). \tag{4.8}$$

We specify a trading strategy by a stochastic process $\theta(t) = (\theta_1(t), \dots, \theta_d(t))$ where $\theta_i(t)$ is the number of units of asset i held just before time t trading. Any decision regarding rebalancing the portfolio at time t must be based on the information available up until time t, which is why we require that $\theta(t)$ is measurable with respect to $\mathcal{F}(t)$. This requirement leaves out the possibility of using future information as it would impose arbitrage possibilities.

The capital gain of a trading strategy is the change in the portfolio value when trading in the underlying assets. We want to describe the capital gain for a general class of trading strategies and we do this using a stochastic integral, so that the gain from trading over $[0, t]$ is

$$\int_0^t \theta^\intercal(u) dS(u). \tag{4.9}$$

This allows trading of any size, both positive or negative, in the underlying assets. Notice that this is obviously not consistent with real world trading, but nevertheless necessary in order to price a large class of derivatives.

In the economy considered, we will only allow so called *self-financing* trading strategies where no exogenous infusion or withdrawal of money is permitted other than the initial amount at time zero. In other words, the purchase of a new portfolio must be financed solely by selling assets already in the portfolio. This is also in line with our initial assumption about a frictionless market where there are no transactions costs when rebalancing a portfolio. The concept of a self-financing trading strategy is so important that we will state the definition[8]:

Definition 4.2. *(SELF-FINANCING TRADING STRATEGY)*
Let the d-dimensional price process $\{S(t), t \geq 0\}$ be given. A self-financing trading strategy θ is a stochastic process $\theta(t) = (\theta_1(t), \ldots, \theta_d(t))$ such that:

1. *θ is $\mathcal{F}(t)$-measurable.*
2. *θ has the self-financing property*

$$\theta(t)^\intercal S(t) = \theta(0)^\intercal S(0) + \int_0^t \theta(u)^\intercal dS(u).$$

So with an initial investment of $V(0) = \theta(0)^\intercal S(0)$ we can obtain a portfolio value of $V(t) = \theta(t)^\intercal S(t)$ by following the trading strategy θ over $[0, t]$.

Naturally, there should be limitations to the profits that self-financed trading strategies can make. One reasonable limitation is that it should be impossible to create something out of nothing. This idea leads us to the main assumptions; that the market is free of arbitrage possibilities. We will now state this very important concept in terms of self-financed trading strategies:[9]:

Definition 4.3. *(DEFINITION OF ARBITRAGE)*
An arbitrage possibility on a financial market is a self-financed portfolio $V(t) = \theta(t)^\intercal S(t)$ such that

1. *$V(0) = 0$,*
2. *$P(V(T) \geq 0) = 1$,*
3. *$P(V(T) \geq 0) > 0$.*

We define the market as arbitrage free if there are no arbitrage possibilities.

So, an arbitrage possibility is the possibility of making a positive amount of money out of nothing without taking risk, and such strategies cannot exist in economic equilibrium. Therefore, precluding arbitrage is a fundamental consistency requirement on the asset processes.

[8]See Hunt and Kennedy (2000), page 145.
[9]See Björk (2009), page 96.

4.2 Martingale Pricing

The goal of this section is to present the martingale approach to pricing derivatives. To this day, this method is the most general approach for arbitrage pricing and is a cornerstone of modern financial mathematics. As the name partly indicates, the method is based on the concepts of martingales and equivalent measures and leads to pricing formulas in the form of expectations. These expectations can be solved using numerical methods such as Monte Carlo simulation and is therefore extremely efficient from a computational point of view. The main problem that concerns us is to find out under which conditions the market is free of arbitrage. As we shall see, these conditions are stated in the *First Fundamental Theorem of Asset Pricing.* In order to be able to understand the main results in the martingale approach, we first need to introduce the concept of equivalent measures. Our presentation follows the ideas in Björk (2009) and Andersen and Piterbarg (2007).

4.2.1 Equivalent Measures

We wish to specify under which conditions the market is free of arbitrage. A way to do this involves the concept of *equivalent martingale measures.* In order to be able to understand this concept we first need to specify what an equivalent measure is. Let us assume that we have two probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ defined on the same measurable space $(\Omega, \mathcal{F})$. Then $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are said to be *equivalent* if $\mathbb{P}(A) = 0 \Leftrightarrow \tilde{\mathbb{P}}(A) = 0,\ \forall A \in \mathcal{F}$. So equivalent measures agree on which sets are impossible (and therefore also which are possible), but they do not necessarily agree on how probable the events are.

A very important result in measure theory and a result used extensively in the theory of derivatives pricing is the *Radon-Nikodym Theorem.* This theorem states that equivalent measures are uniquely connected through an almost surely positive random variable known as the *the Radon-Nikodym derivative*[10]. According to the theorem, the only way to construct a probability measure $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}$ is through the Radon-Nikodym derivative.

Now, consider a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}(t), \mathbb{P})$ on the interval $[0, T]$ where $L(T)$ is some non-negative random variable in $\mathcal{F}(T)$. We define a new measure $\tilde{\mathbb{P}}$ on $\mathcal{F}(T)$ by setting

$$d\tilde{\mathbb{P}} = L(T)d\mathbb{P}, \text{ on } \mathcal{F}(T). \tag{4.10}$$

If $\mathbb{E}^{\mathbb{P}}\left[L(T)\right] = 1$, then the new measure will also be a probability measure where $L(T)$ will be the Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ w.r.t. $\mathbb{P}$ on $\mathcal{F}(T)$, so that $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are equivalent. Thus we will also have that $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are equivalent on $\mathcal{F}(t)$ for all $t \leq T$, so by The Radon-Nikodym Theorem, there will exist a random process $\{L(t), 0 \leq t \leq T\}$ defined by

$$L(t) = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}, \text{ on } \mathcal{F}(t). \tag{4.11}$$

The process L is known as the *Likelihood process* with the property that the process is a $\mathbb{P}$-martingale on $\{\mathcal{F}(t), t > 0\}$. By applying a result known as *Bayes Abtract rule*[11] on any $\mathcal{F}(T)$-

[10] A precise statement of the theorem and a sketch of the proof can be found in Björk (2009).
[11] See Björk (2009), page 501.

measurable random variable $S(T)$, we get the following relation between to equivalent measures

$$\mathbb{E}^{\tilde{\mathbb{P}}}\left[S(T)|\mathcal{F}(t)\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[S(T)L(T)|\mathcal{F}(t)\right]}{\mathbb{E}^{\mathbb{P}}\left[L(T)|\mathcal{F}(t)\right]} \tag{4.12}$$

$$= \frac{1}{L(t)}\mathbb{E}^{\mathbb{P}}\left[S(T)L(T)|\mathcal{F}(t)\right] \tag{4.13}$$

$$= \mathbb{E}^{\mathbb{P}}\left[S(T)\frac{L(T)}{L(t)}\,\middle|\,\mathcal{F}(t)\right]. \tag{4.14}$$

This relation will especially be useful when we introduce the concept of changing numeraires.

Now, the relation described in equations (4.12) to (4.14) explains the connection between equivalent measures and how the transformation from one probability measure to another is linked to the likelihood process. However, it does not explain the effect it has on the underlying process $S(t)$ when changing measures. This effect is described by the *Girsanov theorem*, which also plays a significant role in derivatives pricing. Because of its important role and since we shall use the result numerous times throughout this thesis paper, we will state the theorem according to Björk (2009)[12]:

Theorem 4.3. *(THE GIRSANOV THEOREM)*
Let $W^{\mathbb{P}}(t) = (W_1^{\mathbb{P}}(t), \ldots, W_k^{\mathbb{P}}(t))^\intercal$ be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}(t), \mathbb{P})$. Let $\varphi(t) = (\varphi_1(t), \ldots, \varphi_k(t))^\intercal$ be any d-dimensional adapted process. Choose a fixed T and define the process L on $[0, T]$ by

$$L(t) = \exp\left(-\int_0^t \varphi(u)^\intercal dW^{\mathbb{P}}(u) - \frac{1}{2}\int_0^t \|\varphi(u)\|^2 du\right),$$
$$L(0) = 1,$$

where $\|\varphi(u)\|$ denotes the Euclidean norm $\|\varphi(u)\| = \left(\sum_{i=1}^d \varphi_i^2(u)\right)^{\frac{1}{2}}$. Assume that

$$\mathbb{E}^{\mathbb{P}}[L(T)] = 1,$$

and define the new probability measure $\tilde{\mathbb{P}}$ on $\mathcal{F}(T)$ by

$$L(T) = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}, \text{ on } \mathcal{F}(T).$$

Then

$$dW^{\tilde{\mathbb{P}}}(t) = dW^{\mathbb{P}}(t) + \varphi(t)dt,$$

where $W^{\tilde{\mathbb{P}}}$ is a d-dimensional Brownian motion.

Technical comments. The process φ is often referred to as the *Girsanov kernel* of the measure transformation. Note that the theorem differs slightly when compared with the version in Björk (2009), where the form of the likelihood process is different, and as a consequence the Girsanov

[12]See Björk (2009), page 164-165.

theorem states that

$$dW^{\mathbb{P}}(t) = dW^{\tilde{\mathbb{P}}}(t) + \varphi(t)dt.$$

For a proof of the Girsanov Theorem, the reader is referred to Björk (2009), Chapter 11.

So, the Girsanov Theorem explains the connection between Brownian motion under two equivalent measures. We see that the theorem has the attractive consequence that the effects on a stochastic process of changing between measures are captured by a simple adjustment to the drift. To see this, let us consider the simple case where we have a one-dimensional Itô process under the measure $\mathbb{P}$

$$dS(t) = \mu(S(t), t)dt + \sigma(S(t), t)dW^{\mathbb{P}}(t).$$

Now, assume the existence of an equivalent measure $\tilde{\mathbb{P}}$ with the associated kernel φ. Then by the Girsanov theorem we have that

$$\begin{aligned} dS(t) &= \mu(S(t), t)dt + \sigma(S(t), t)(dW^{\tilde{\mathbb{P}}}(t) - \varphi(t)dt) \\ &= (\mu(S(t), t) - \varphi(t)\sigma(S(t), t))dt + \sigma(S(t), t)dW^{\tilde{\mathbb{P}}}(t). \end{aligned}$$

Thus, $\mu(S(t), t) - \varphi(t)\sigma(S(t), t)$ is the new drift under $\tilde{\mathbb{P}}$. However, it is important to notice that the volatility remains unchanged during the measure transformation.

4.2.2 Equivalent Martingale Measures and No-Arbitrage

We now introduce the concept of a *deflator*, a strictly positive Itô process used to normalize asset prices. When an asset is chosen as a deflator we call it a *numeraire*. We now expand our market model with one extra asset, a strictly positive Itô process S_0 which leaves us with $d+1$ assets in total. So, instead of looking at the price vector process $S(t) = (S_0(t), \dots, S_d(t))^\intercal$, we will look at the relative price vector process $\frac{S(t)}{S_0(t)}$, where $S_0(t)$ acts as the numeraire. We define the concept of a *normalized economy* where the price vector process Z defines the normalized asset prices processes by

$$Z(t) = (Z_0(t), \dots, Z_d(t))^\intercal = \left(1, \frac{S_1(t)}{S_0(t)}, \dots, \frac{S_d(t)}{S_0(t)}\right)^\intercal. \tag{4.15}$$

Notice that in the normalised economy we have a risk free asset $Z_0 = 1$, with zero rate of return. We call a measure $\mathbb{Q}^0$ an *equivalent martingale measure induced by* S_0 if $Z(t)$ is a martingale w.r.t $\mathbb{Q}^0$.[13]

Let us consider a portfolio based on a self-financing trading strategy θ, i.e. $dV^S(S(t), t) = \theta(t)^\intercal dS(t)$. We call this portfolio S-self-financing. Then by using Itô's product rule[14] we can

[13] The choice of $\mathbb{Q}$ (in our case $\mathbb{Q}^0$) as notation for an equivalent martingale measure is standard practice in most text books on mathematical finance.

[14] Let $X(t)$ and $Y(t)$ be Itô processes. Then

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t).$$

derive the dynamics of a portfolio $V^Z = \frac{V^S}{S_0}$ as

$$dV^Z(S(t),t) = \theta(t)^\intercal S(t) d\left(\frac{1}{S_0(t)}\right) + \frac{1}{S_0(t)}\theta(t)^\intercal dS(t) + \theta(t)^\intercal dS(t) d\left(\frac{1}{S_0(t)}\right) \tag{4.16}$$

$$= \theta^\intercal \left(S(t) d\left(\frac{1}{S_0(t)}\right) + \frac{1}{S_0(t)} dS(t) + dS(t) d\left(\frac{1}{S_0(t)}\right)\right) \tag{4.17}$$

$$= \theta^\intercal d\left(\frac{S(t)}{S_0(t)}\right) = \theta^\intercal dZ(t) \tag{4.18}$$

and by integration

$$V^Z(S(t),t) = V^Z(S(0),0) + \int_0^t \theta(u)^\intercal dZ(u). \tag{4.19}$$

This implies that if a portfolio based on a trading strategy θ is S-self-financing, then the portfolio based on θ is also Z-self-financing. The logic behind this result is that a portfolio based on a self-financing trading strategy should not depend on a specific choice of numeraire. This result is formulated more precisely in the *Invariance Lemma*[15]. If we assume that $\mathbb{Q}^0$ is an equivalent martingale measure, then V^Z is a $\mathbb{Q}^0$-martingale if the gain process

$$\int_0^t \theta(u)^\intercal dZ(u), \text{ for all } t \in [0,T] \tag{4.20}$$

is a $\mathbb{Q}^0$-martingale.

Assume that there exists an equivalent martingale measure $\mathbb{Q}^0$ (equivalent to $\mathbb{P}$) where the normalised asset price process $Z(t)$ are $\mathbb{Q}^0$-martingales. We also assume that the gain process (4.20) is a $\mathbb{Q}^0$-martingale. Given these assumptions, we want to show that no arbitrage possibilities exist. We assume that θ is bounded, which makes integration over θ possible. Since we do not wish to get too technical we will not discuss the case when θ is possibly unbounded.

We consider a self-financing trading strategy θ with the corresponding portfolio value $V^Z(S(t),t) = \theta^\intercal Z(t)$ and assume that the portfolio satisfies the relations

$$\mathbb{P}(V^Z(S(t),t) \geq 0) = 1, \tag{4.21}$$

$$\mathbb{P}(V^Z(S(t),t) > 0) > 0. \tag{4.22}$$

Hence, θ constitutes a possible arbitrage strategy according to Definition 4.3 on page 26 and we therefore have to show that $V^Z(S(0),0) > 0$ in order to rule out arbitrage.
Since $\mathbb{P}$ and $\mathbb{Q}^0$ are equivalent, we must have that

$$\mathbb{Q}^0(V^Z(S(t),t) \geq 0) = 1, \tag{4.23}$$

$$\mathbb{Q}^0(V^Z(S(t),t) > 0) > 0, \tag{4.24}$$

and since θ is self-financing and (4.20) is a $\mathbb{Q}^0$-martingale, we know that $V^Z(S(t),t)$ is a $\mathbb{Q}^0$-martingale (since we assume that θ is bounded). In particular we have that

$$V^Z(S(0),0) = \mathbb{E}^{\mathbb{Q}^0}\left[V^Z(S(T),T)\right]. \tag{4.25}$$

[15]See Björk (2009), page 147-148.

However, (4.23)-(4.24) imply that $\mathbb{E}^{\mathbb{Q}^0}\left[V^Z(S(T),T)\right] > 0$, which means that $V^Z(S(0),0) > 0$. And since $\mathbb{P}$ and $\mathbb{Q}^0$ are equivalent we have shown that (4.21)-(4.22) imply $V^Z(S(0),0) > 0$ and therefore that the arbitrage trading strategy cannot exist.

There exists a general result known as *The First Fundamental Theorem of Asset Pricing* and as the name indicates, it is of huge importance for asset pricing theory. Essentially, the theorem can be divided into two parts. The first part states that *if* an equivalent measure exists where the normalized asset prices w.r.t a numeraire are martingales, then the prices are arbitrage free. The second part explains the conditions under which such a measure exists. The proof of the theorem is very technical and especially the proof of the second part requires very demanding mathematics. This is indeed outside the scope of this thesis paper, and therefore we conclude from the discussion that the absence of arbitrage is equivalent to the existence of a martingale measure. For further discussion, see Björk (2009), chapter 10.

4.2.3 The General Pricing Formula

We define a derivative, or a contingent claim, with maturity at time T as any $\mathcal{F}(T)$-measurable random variable. We let X denote the contract function of the derivative.[16] Our goal is to establish a general pricing formula for the "fair" price of a derivative, denoted $\pi_X(t)$, based on the no-arbitrage framework described above, where $\pi_X(T) = X$.
Again, we consider a market consisting of $d+1$ assets with price vector process $S(t) = (S_0(t), \dots, S_d(t))$ and assume that the market is free of arbitrage. We fix a derivative with maturity at time T and contract function X. The price of the derivative should be consistent with the prices of the underlying assets, and we therefore demand that the market consisting of $\pi_X(t)$ and $S(t)$ should be free of arbitrage.

Now, by the First Fundamental Theorem of Asset Pricing, we know that the existence of an equivalent martingale measure ensures that the market admits no arbitrage. So, by choosing S_0 as the numeraire, we demand that there exists an equivalent martingale measure induced by S_0 for the market consisting of $\pi_X(t)$ and $S(t)$. Denoting this measures $\mathbb{Q}^0$ and using the fact that $\frac{\pi_X(t)}{S_0(t)}$ is a $\mathbb{Q}^0$-martingale, we can establish the following general pricing formula

$$\begin{aligned}\frac{\pi_X(t)}{S_0(t)} &= \mathbb{E}^{\mathbb{Q}^0}\left[\frac{\pi_X(T)}{S_0(T)} \,\middle|\, \mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{Q}^0}\left[\frac{X}{S_0(T)} \,\middle|\, \mathcal{F}(t)\right] \Leftrightarrow \\ \pi_X(t) &= S_0(t)\mathbb{E}^{\mathbb{Q}^0}\left[\frac{X}{S_0(T)} \,\middle|\, \mathcal{F}(t)\right]. \qquad (4.26)\end{aligned}$$

Now, let us assume that the derivative can be replicated by a self-financing trading strategy θ with the portfolio value dynamics $dV(S(t),t) = \theta(t)^\intercal dS(t)$ and that the terminal value of the portfolio equals the payoff of the derivative i.e.

$$V(S(T),T) = \pi_X(T). \qquad (4.27)$$

[16] In the case of a standard European call option written on an underlying asset $S(t)$ with strike K, the contract function is given by

$$X = \Big(S(T) - K\Big)^+.$$

In this case, we call the derivative *attainable* with θ. Furthermore, if we assume the existence of an equivalent martingale measure $\mathbb{Q}^0$ where the gain process for $V^Z(S(t),t) = \frac{V(S(t),t)}{S_0(t)}$ is a $\mathbb{Q}^0$-martingale, we know from Subsection 4.2.2 that $V^Z(S(t),t)$ is a $\mathbb{Q}^0$-martingale. From (4.27) this implies that $\frac{\pi_X(t)}{S_0(t)}$ is also a $\mathbb{Q}^0$-martingale and we can again establish the general pricing formula where we, for an attainable derivative, have that

$$V(S(t),t) = S_0(t)\mathbb{E}^{\mathbb{Q}^0}\left[\frac{X}{S_0(T)}\,\middle|\,\mathcal{F}(t)\right], \tag{4.28}$$

which holds for any replicating self-financing trading strategy under any martingale measure.

If all derivatives are attainable, then the market is *complete* and (4.26) and (4.28) coincide. So, in a complete market the price of any derivative is uniquely determined by the demand of absence of arbitrage. The uniqueness stems from the fact that the derivative could just as well be replaced by any replicating trading strategy. Whether a market is complete or not depends on the number of risky traded assets compared to the number of random sources in the model, where completeness is achieved if there are *at least* as many risky traded assets as there are random sources. In our case, the market is complete if there are at least as many risky assets, $S_i(t)$, as there are Brownian motions. For more on complete markets see Björk (2009), Chapter 8.

If the market is not complete we call it *incomplete*. In an incomplete market the demand for absence of arbitrage is not enough to determine unique prices for derivatives. In this case, there exist many equivalent martingale measures under which no-arbitrage prices for derivatives can be established according to (4.26). So, in order to use (4.26), one has to choose a specific equivalent martingale measure. This measure should be the one chosen by the market, which can be achieved by calibrating to market prices. For more on incomplete markets see Björk (2009), Chapter 15.

The surprisingly simple formula in (4.26) is at the heart of asset pricing theory and makes pricing derivatives very flexible. In particular, it establishes a foundation for pricing derivatives using simulation schemes such as Monte Carlo Methods. In the next section, we will turn our focus towards a specific measure that has certain nice properties. This measure is known as the *Risk Neutral measure* and will be used extensively throughout this thesis paper.

4.3 Risk Neutral Pricing

Let us assume that there exists one risk free asset $M(t)$ in the sense that it evolves according to an Itô process with $\sigma_i(S(t),t) = 0$ and $\mu_i = r(t)$. Here $r(t)$ is a one-dimensional $\mathcal{F}(t)$-adapted stochastic process describing the evolution of the instantaneous, continuously compounded short rate of interest. The short rate may be interpreted as a risk free interest rate and we can think of M as a money market account with dynamics

$$dM(t) = r(t)M(t)dt \Leftrightarrow M(t) = M(0)\exp\left(\int_0^t r(u)du\right) \tag{4.29}$$

where $M(0) = 1$. Now since $M(t)$ is a stricty positive (Itô) process, it can be used as a numeraire. So by letting $S_0(t) = M(t)$ in (4.26) we arrive at the well-known *risk neutral pricing formula*

$$\pi_X(t) = M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{M(T)}\,\middle|\,\mathcal{F}(t)\right] \Leftrightarrow \tag{4.30}$$

$$\pi_X(t) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)X\,\middle|\,\mathcal{F}(t)\right] \tag{4.31}$$

where $\mathbb{Q}$ is called the *risk neutral martingale measure* with the money market account as numeraire. Hence, the time t price of a derivative under $\mathbb{Q}$ can be computed as an expected discounted value. The measure $\mathbb{Q}$ is referred to as the risk-neutral measure since discounting is done by using the risk free interest rate among other things.

Next, we want to define the asset dynamics under the $\mathbb{Q}$-measure. For this purpose, let us assume that the economy is described by a system of strictly positive price processes

$$dS_i(t) = S_i(t)\left(\mu_i(S(t),t)dt + \sigma_i(S(t),t)^{\intercal}dW^{\mathbb{P}}(t)\right), \text{ for } i = 1,\dots,d \tag{4.32}$$

where $W^{\mathbb{P}}$ is a d-dimensional Brownian motion under the objective measure $\mathbb{P}$, μ_i is a one-dimensional $\mathcal{F}(t)$-adapted process and σ_i is an $\mathcal{F}(t)$-adapted process taking values in $\mathbb{R}^d$. The Girsanov Theorem stated in Subsection 4.2.1 will be essential in order to describe the connection between $\mathbb{P}$ and $\mathbb{Q}$.

Now, we say that a probability measure $\mathbb{Q}$ is *risk neutral* if

1. $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, and
2. under $\mathbb{Q}$, the discounted asset prices $\frac{S_i(t)}{M(t)}$ are martingales.

Let us begin by finding the dynamics of discounted asset prices $\frac{S_i(t)}{M(t)}$. From Itô's formula, one can show that $d\frac{1}{M(t)} = -r(t)\frac{1}{M(t)}dt$[17]. Then, by combining this result with Itô's product rule, we have that

$$\begin{aligned}
d\left(\frac{S_i(t)}{M(t)}\right) &= S_i(t)\left(-r(t)\frac{1}{M(t)}dt\right) + \frac{1}{M(t)}S_i(t)\left(\mu_i(S(t),t)dt + \sigma_i(S(t),t)^{\intercal}dW^{\mathbb{P}}(t)\right)\\
&\quad - r(t)\frac{1}{M(t)}dt\left(S_i(t)\left(\mu_i(S(t),t)dt + \sigma_i(S(t),t)^{\intercal}dW^{\mathbb{P}}(t)\right)\right)\\
&= \frac{S_i(t)}{M(t)}\left((\mu_i(S(t),t) - r(t))dt + \sigma_i(S(t),t)^{\intercal}dW^{\mathbb{P}}(t)\right)
\end{aligned} \tag{4.33}$$

In order to convert the discounted asset prices into martingales we would like to rewrite (4.33) and then use the Girsanov Theorem so that

$$d\left(\frac{S_i(t)}{M(t)}\right) = \frac{S_i(t)}{M(t)}\sigma_i(S(t),t)^{\intercal}(\varphi(t)dt + dW^{\mathbb{P}}(t)) \tag{4.34}$$

$$= \frac{S_i(t)}{M(t)}\sigma_i(S(t),t)^{\intercal}dW^{\mathbb{Q}}(t) \tag{4.35}$$

If we can specify the process $\varphi(t)$ that makes (4.34) hold, we can then use the Girsanov theorem to construct an equivalent measure $\mathbb{Q}$ under which $W^{\mathbb{Q}}$ is a d-dimensional Brownian motion to

[17] For a derivation, see Appendix A.

obtain (4.35). Thus, making $\frac{S_i(t)}{M(t)}$ a martingale under $\mathbb{Q}$. This means that finding a risk neutral measure comes down to specifying the process φ that makes (4.33) and (4.34) agree. Simple manipulations show that they agree if and only if

$$\mu_i(S(t), t) = r(t) + \sigma_i(S(t), t)^\intercal \varphi(t). \tag{4.36}$$

From this relation, we can see that φ characterizes a scaled risk premium in the sense that it determines the amount by which the drift rate exceeds the risk free rate. If we consider the scalar case and rearrange (4.36) to $\varphi(t) = \frac{\mu_i(S(t),t)-r(t)}{\sigma_i(S(t),t)}$ we can interpret φ as *the market price of risk*, which measures the excess return demanded by investors per unit of risk. In the multidimensional setup, each component of φ_i may also be interpreted as the the market price of risk associated with the ith risk factor.

So, how are the undiscounted dynamics of assets affected under the risk neutral measure? It turns out that $S_i(t)$ has a mean rate of return (drift) equal to the risk free rate under $\mathbb{Q}$. This can be verified by making the replacement $dW^{\mathbb{P}} = -\varphi(t)dt + dW^{\mathbb{Q}}$ in (4.32) along with (4.36). With these substitutions, the dynamics of $S_i(t)$ under $\mathbb{Q}$ becomes

$$\begin{aligned} dS_i(t) &= S_i(t) \left((r(t) + \sigma_i(S(t), t)^\intercal \varphi(t))dt + \sigma_i(S(t), t)^\intercal (-\varphi(t)dt + dW^{\mathbb{Q}}) \right) \\ &= S_i(t) \left(r(t)dt + \sigma_i(S(t), t)^\intercal dW^{\mathbb{Q}} \right). \end{aligned} \tag{4.37}$$

Consequently, when we change measures from $\mathbb{P}$ to $\mathbb{Q}$ the drift in the dynamics is shifted with the risk premium. This makes sense, since in a risk neutral world all assets will command a rate of return equal to the risk free interest rate and therefore have a risk premium of 0. Also, the present value of any future stochastic payment would be equal to the expected value of the net payments discounted using the risk free interest rate. Since the money market account is not considered a risky asset, we have equally many risky assets and brownian motions which make the market complete. Hence, all risk associated with derivatives can be perfectly replicated and we can disregard risk premiums when pricing, since all assets evolve according to the risk free interest rate. The reason why the risk-neutral pricing formula does not depend on knowing the mean rate of return of the underlying assets is because this information is already included in the asset prices and the derivative is priced according to these prices. However, it is important to emphasise that just because we are pricing derivatives under the risk-neutral measure it does not mean that we assume that investors are risk-neutral. The formula only states that derivatives can be priced *as if* investors were risk-neutral. In fact (4.30) is preference free in the sense that the formula holds regardless of the investors individual preferences.

It turns out that in certain situations, due to modeling considerations, it is convenient to change the benchmark asset to which other assets are measured. For instance, if one chooses to model the short rate as a stochastic process, the money market account becomes stochastic. Consequently, calculating (4.30) gets complicated since it would require taking the expectation of the joint distribution (computing a double integral) of $M(T)$ and X under the $\mathbb{Q}$ measure. In this case and in many other cases, it is possible to reduce the computational complexities in models by changing the numeraire. So, computing prices within a specific model can be complicated or simple depending on the choice of the numeraire in the model. This interesting fact, which was formally described by Geman, Karoui and Rochet (1995), leads us to the next important topic in derivatives pricing.

4.4 Change of Numeraire

As we have already mentioned, a deflator is a strictly positive Itô process that acts as a standard by which other asset values are measured. If the deflator is chosen from one of the assets, we call the deflator a *numeraire*. In the risk-neutral setup the money market account acted as the numeraire. By choosing the money market account as a numeraire we defined the risk-neutral probability measure $\mathbb{Q}$. The idea of changing numeraire is thus to measure the value of derivatives and assets relative to a different asset than the one first chosen. Our goal is therefore to show how changing numeraire affects pricing formulas and the underlying dynamics. In this section, we follow the ideas in Poulsen (1999) and Björk (2009).

We begin by placing ourselves in the risk-neutral setup where asset dynamics are specified under the $\mathbb{Q}$-measure by

$$dS_i(t) = S_i(t)\left(r(t)dt + \sigma_i(S(t),t)^{\intercal} dW^{\mathbb{Q}}(t)\right), \text{ for } i = 1,\dots,d \tag{4.38}$$

along with the dynamics of the money market account described in (4.29). We remind ourselves that $W^{\mathbb{Q}}$ is a d-dimensional Brownian motion, $r(t)$ is a one-dimensional $\mathcal{F}(t)$-adapted process and σ_i is an $\mathcal{F}(t)$-adapted process taking values in $\mathbb{R}^d$. We want to use the risky asset S_d as numeraire which means that we will define a new probability measure $\mathbb{Q}^d$.

First, let us consider the following process

$$L(T) = \frac{S_d(T)M(0)}{S_d(0)M(T)} = \frac{S_d(T)}{S_d(0)M(T)} \tag{4.39}$$

As a consequence, we have that

$$L(T) > 0 \text{ and } L(0) = 1. \tag{4.40}$$

Furthermore, notice that under the $\mathbb{Q}$-measure, $L(T)$ is a martingale, i.e.

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{S_d(T)}{S_d(0)M(T)} \,\middle|\, \mathcal{F}(t)\right] = \frac{1}{S_d(0)}\mathbb{E}^{\mathbb{Q}}\left[\frac{S_d(T)}{M(T)} \,\middle|\, \mathcal{F}(t)\right] = \frac{S_d(t)}{S_d(0)M(t)} \tag{4.41}$$

where we use that fact that $\frac{S_d}{M}$ is a $\mathbb{Q}$-martingale. Therefore, from Subsection 4.2.1, we know that we can define a new equivalent probability measure $\mathbb{Q}^d$ by the likelihood process

$$L(t) = \frac{d\mathbb{Q}^d}{d\mathbb{Q}} \text{ on } \mathcal{F}(t). \tag{4.42}$$

We wish to establish a link between a general pricing formula using the risky asset S_d as numeraire and the risk neutral formula in (4.30) and thereby show how one can change numeraire from M to S_d. Following the relation between equivalent measures we derived in (4.12)-(4.14) on page 28 we have that

$$S_d(t)\mathbb{E}^{\mathbb{Q}^d}\left[\frac{\pi_X(T)}{S_d(T)}\,\middle|\,\mathcal{F}(t)\right] = S_d(t)\frac{\mathbb{E}^{\mathbb{Q}}\left[\frac{\pi_X(T)}{S_d(T)}L(T)\,\middle|\,\mathcal{F}(t)\right]}{\mathbb{E}^{\mathbb{Q}}\left[L(T)\,\middle|\,\mathcal{F}(t)\right]} \tag{4.43}$$

$$= S_d(t)\frac{1}{L(t)}\mathbb{E}^{\mathbb{Q}}\left[\frac{\pi_X(T)}{S_d(T)}L(T)\,\middle|\,\mathcal{F}(t)\right] \tag{4.44}$$

$$= S_d(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{\pi_X(T)}{S_d(T)}\frac{S_d(T)}{S_d(0)M(T)}\frac{S_d(0)M(t)}{S_d(t)}\,\middle|\,\mathcal{F}(t)\right] \tag{4.45}$$

$$= M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{\pi_X(T)}{M(T)}\,\middle|\,\mathcal{F}(t)\right] \tag{4.46}$$

$$= \pi_X(t). \tag{4.47}$$

This means that $\frac{\pi_X(t)}{S_d(t)}$ is a $\mathbb{Q}^d$-martingale and the following relation holds

$$\pi_X(t) = S_d(t)\mathbb{E}^{\mathbb{Q}^d}\left[\frac{\pi_X(T)}{S_d(T)}\,\middle|\,\mathcal{F}(t)\right]. \tag{4.48}$$

Notice that if we had begun the construction from the $\mathbb{Q}$-measure, we could have achieved the same result by defining the inverse likelihood process $L^{-1}(T) = \frac{S_d(0)M(T)}{S_d(T)}$ and then used that $L^{-1}(t)$ is a $\mathbb{Q}^d$-martingale.

So, when using S_d as numeraire, we see from (4.48) that all derivatives and assets are martingales under the $\mathbb{Q}^d$-measure. In order to derive (4.48) we had to begin the construction from a martingale measure, in this case the $\mathbb{Q}$-measure. This confirms that we can change between martingale measures.

The relation (4.48) does not provide any information about the asset dynamics under the $\mathbb{Q}^d$-measure, and this information is necessary in order to use (4.48). We know from the Girsanov Theorem that in order to change between equivalent measures, we need to find the connection between the Brownian motions under the two measures. We also know that $\frac{M(t)}{S_d(t)}$ is a $\mathbb{Q}^d$-martingale. With this in mind, we will now try to derive the $\mathbb{Q}^d$-dynamics.

First, we begin with $\frac{1}{S_d(t)}$ and by Itô's formula we have that

$$d\left(\frac{1}{S_d(t)}\right) = -\frac{1}{S_d^2(t)}dS_d(t) + \frac{1}{S_d^3(t)}(dS_d(t))^2 \tag{4.49}$$

$$= -\frac{1}{S_d(t)}\Big(r(t)dt + \sigma_d(S(t),t)^\intercal dW^Q(t)\Big) + \frac{1}{S_d(t)}\Big(\sigma_d(S(t),t)^\intercal\sigma_d(S(t),t)dt\Big) \tag{4.50}$$

$$= \frac{1}{S_d(t)}\Big(-r(t) + \sigma_d(S(t),t)^\intercal\sigma_d(S(t),t)\Big)dt - \frac{1}{S_d(t)}\sigma_d(S(t),t)^\intercal dW^Q(t). \tag{4.51}$$

Then by Itô's product rule we have that

$$d\left(\frac{M(t)}{S_d(t)}\right) = \frac{M(t)}{S_d(t)}\Big(\Big(-r(t) + \sigma_d(S(t),t)^\intercal \sigma_d(S(t),t)\Big)dt - \sigma_d(S(t),t)^\intercal dW^Q(t)\Big) \tag{4.52}$$

$$+ \frac{r(t)M(t)}{S_d(t)}dt \tag{4.53}$$

$$= \frac{M(t)\sigma_d(S(t),t)^\intercal}{S_d(t)}\Big(\sigma_d(S(t),t)dt - dW^Q(t)\Big). \tag{4.54}$$

Since we know that $\frac{M(t)}{S_d(t)}$ is a $\mathbb{Q}^d$-martingale, we must have that the process $\sigma_d(S(t),t)dt - dW^Q(t)$ is also a $\mathbb{Q}^d$-martingale. Then from Girsanov it follows that $dW^{\mathbb{Q}^d} = dW^{\mathbb{Q}} - \sigma_d(S(t),t)dt$ is a standard Brownian motion under $\mathbb{Q}^d$ and the dynamics of S_i under $\mathbb{Q}^d$ is defined as

$$dS_i(t) = S_i(t)\Big(r(t)dt + \sigma_i(S(t),t)^\intercal \left(dW^{\mathbb{Q}^d} + \sigma_d(S(t),t)dt\right)\Big) \tag{4.55}$$

$$= S_i(t)\Big(\Big(r(t) + \sigma_i(S(t),t)^\intercal \sigma_d(S(t),t)\Big)dt + \sigma_i(S(t),t)^\intercal dW^{\mathbb{Q}^d}\Big) \tag{4.56}$$

where the dynamics of $M(t)$ is unchanged since it does not contain a Wiener term. This implies that when we change measure from $\mathbb{Q}$ to $\mathbb{Q}^d$ we have to correct the drift of the underlying assets by adding the instantaneous covarians between asset i and the numeraire.

In order to completely describe the impact of changing measures, we will now derive the dynamics of the normalized price process, $Z_i(t) = \frac{S_i(t)}{S_d(t)}$. We know that $Z_i(t)$ is a martingale under $\mathbb{Q}^d$ and since $S_i(t)$ is on proportional form, we have that (here we just write σ_i in order to ease notation)

$$\frac{S_i(t)}{S_d(t)} = \frac{S_i(0)}{S_d(0)}\exp\left(\left(r(t) - r(t) - \frac{1}{2}\|\sigma_i\|^2 + \frac{1}{2}\|\sigma_d\|^2\right)T + (\sigma_i - \sigma_d)^\intercal W^{\mathbb{Q}}(T)\right) \tag{4.57}$$

$$= \frac{S_i(0)}{S_d(0)}\exp\left(\left(\frac{1}{2}\|\sigma_d\|^2 - \frac{1}{2}\|\sigma_i\|^2\right)T + (\sigma_i - \sigma_d)^\intercal W^{\mathbb{Q}}(T)\right) \tag{4.58}$$

$$= \frac{S_i(0)}{S_d(0)}\exp\left(\left(-\frac{1}{2}\|\sigma_d\|^2 - \frac{1}{2}\|\sigma_i\|^2 + \sigma_i^\intercal\sigma_d\right)T + (\sigma_i - \sigma_d)^\intercal (W^{\mathbb{Q}}(T) - \sigma_d T)\right) \tag{4.59}$$

$$= \frac{S_i(0)}{S_d(0)}\exp\left(-\frac{1}{2}\|\sigma_i - \sigma_d\|^2 T + (\sigma_i - \sigma_d)^\intercal (W^{\mathbb{Q}}(T) - \sigma_d T)\right). \tag{4.60}$$

Now, if we let $\frac{S_i(t)}{S_d(t)} = \exp(h)$ where $h = \ln\left(\frac{S_i(t)}{S_d(t)}\right)$, we can use Itô's formula to derive the dynamics as

$$d\left(\frac{S_i(t)}{S_d(t)}\right) = \frac{S_i(t)}{S_d(t)}(\sigma_i - \sigma_d)^\intercal (dW^{\mathbb{Q}}(t) - \sigma_d dt). \tag{4.61}$$

Thus, by Girsanov, we can show that $Z_i(t) = \frac{S_i(t)}{S_d(t)}$ is a $\mathbb{Q}^d$-martingale with the dynamics (reintroducing proper notation)

$$dZ_i(t) = Z_i(t)\Big(\sigma_i(S(t),t) - \sigma_d(S(t),t)\Big)^\intercal dW^{\mathbb{Q}^d}(t). \tag{4.62}$$

The above results show the consequences of changing measures from the risk-neutral martingale measure to a new measure induced by the numeraire S_d. It is important to stress that one does not need to initiate the construction from the risk-neutral measure. We could just as easily have initiated the construction from any martingale measure using the same procedure described above. The consequences of changing from the risk-neutral measure is that the dynamics of the money market account does not change, since it is a risk free investment and does not contain a Wiener driven term. This would not have been the case if we had changed from a martingale measure induced by a risky numeraire.

The results above can be compressed into the following important theorem.

Theorem 4.4. *(CHANGE OF MEASURE)*
Consider a market model with asset dynamics given by (4.38) *and a money market account given by* (4.29) *under the risk-neutral martingale measure* $\mathbb{Q}$. *For any fixed numeraire process* $S_d(t)$ *there exists an equivalent martingale measure* $\mathbb{Q}^d$ *such that*

$$M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{\pi_X(T)}{M(T)}\,\middle|\,\mathcal{F}(t)\right] = S_d(t)\mathbb{E}^{\mathbb{Q}^d}\left[\frac{\pi_X(T)}{S_d(T)}\,\middle|\,\mathcal{F}(t)\right].$$

In particular, we have that the price of any traded asset is found as

$$\pi_X(t) = S_d(t)\mathbb{E}^{\mathbb{Q}^d}\left[\frac{\pi_X(T)}{S_d(T)}\,\middle|\,\mathcal{F}(t)\right],$$

with the dynamics under $\mathbb{Q}^d$ *given by*

$$\begin{aligned} dM(t) &= r(t)M(t)dt, \\ dS_i(t) &= S_i(t)\Big(\Big(r(t) + \sigma_i(S(t),t)^\intercal\sigma_d(S(t),t)\Big)dt + \sigma_i(S(t),t)^\intercal dW^{\mathbb{Q}^d}\Big) \text{ for all } i, \end{aligned}$$

where $W^{\mathbb{Q}^d}(t) = W^{\mathbb{Q}}(t) - \sigma_d(S(t),t)$ *is a* $\mathbb{Q}^d$*-Brownian motion. If we define*

$$Z_i(t) = \frac{S_i(t)}{S_d(t)} \text{ for all } i,$$

we have that

$$dZ_i(t) = Z_i(t)\Big(\sigma_i(S(t),t) - \sigma_d(S(t),t)\Big)^\intercal dW^{\mathbb{Q}^d}(t) \text{ for all } i.$$

It is not immediately obvious how a change of measure can simplify the valuation of a given derivative. All we have shown are the implications on pricing formulas and underlying dynamics when a new measure is introduced. So in order to show that the change of a numeraire can be a useful tool to reduce the computational time, we now proceed with an example concerning swaptions. Since swaption pricing plays a key role in the valuation of a credit risky swap, we will, in this example, show how changing numeraire can reduce the complexity of the pricing formula significantly. The end results will be applied later on in the thesis paper. We choose to focus solely on the swaption pricing formula and not the underlying dynamics since this will be covered later on.

4.4.1 Pricing Swaptions under the Swap Measure

We assume that a frictionless market exists for ZCBs for every maturity $T \leq \tilde{T}$. As in Section 3.1.1, we denote the time t price of a ZCB with maturity at T as $D(t,T)$ where $D(T,T) = 1$. We place ourselves under the risk-neutral measure and assume that the interest rate evolves according to some unspecified stochastic process. According to Subsection 4.3, we can then express the time t price of a ZCB as

$$D(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right) \,\middle|\, \mathcal{F}(t)\right]. \tag{4.63}$$

From Subsection 3.3.1 we know that the arbitrage free forward swaprate at time t for a swap contract starting at time T_n lasting until T_N, with the first payment at time T_{n+1}, is given by

$$y_{n,N}(t) = \frac{D(t,T_n) - D(t,T_N)}{\sum_{i=n+1}^{N} \Delta D(t,T_i)} = \frac{D(t,T_n) - D(t,T_N)}{P_{n+1,N}(t)} \tag{4.64}$$

where Δ is the daycount fraction (here assumed constant) and $P_{n+1,N}(t)$ is the so-called annuity factor. Now, according to (3.15) on page 18, the payoff of a payer swaption can be seen as proportional to that from a call option on the swaprate. Hence, under the risk-neutral measure the price of a payer swaption at time t with strike K, denoted $\pi_{PS}(t)$, is obtained as

$$\pi_{PS}(t) = M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{P_{n+1,N}(T_n)}{M(T_n)}(y_{n,N}(T_n) - K)^+ \,\middle|\, \mathcal{F}(t)\right] \tag{4.65}$$

Since we assume a stochastic interest rate, both the money market account and the annuity factor are random variables at time T_n. Therefore, computing (4.65) is not a simple task, since it involves computing the expectation of the joint distribution of three random variables which implies computing a triple integral. We will now show how changing the numeraire from the money account to the annuity factor can simplify the computation greatly.

The idea is to use $P_{n+1,N}(t)$ as numeraire, and we therefore consider the following likelihood process

$$L(T_n) = \frac{P_{n+1,N}(T_n)M(t)}{P_{n+1,N}(t)M(T_n)} \tag{4.66}$$

where $L(t) = 1$ and $L(T_n) > 0$. Since $L(T_n)$ is a $\mathbb{Q}$-martingale, we can define an equivalent measure $\mathbb{Q}^{n+1,N}$ to $\mathbb{Q}$ by its Radon-Nikodym derivative $L(t)$. We will call this measure the *Swap measure*. When applying $P_{n+1,N}$ as numeraire, we know that $\frac{M(T_N)}{P_{n+1,N}(T_N)}$ is $\mathbb{Q}^{n+1,N}$-martingale, since[18]

$$\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[\frac{M(T_n)}{P_{n+1,N}(T_n)} \,\middle|\, \mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{L(T_n)}{L(t)}\frac{M(T_n)}{P_{n+1,N}(T_n)} \,\middle|\, \mathcal{F}(t)\right], \tag{4.67}$$

$$= \mathbb{E}^{\mathbb{Q}}\left[\frac{P_{n+1,N}(T_n)M(t)}{P_{n+1,N}(t)M(T_n)}\frac{M(T_n)}{P_{n+1,N}(T_n)} \,\middle|\, \mathcal{F}(t)\right], \tag{4.68}$$

$$= \mathbb{E}^{\mathbb{Q}}\left[\frac{M(t)}{P_{n+1,N}(t)} \,\middle|\, \mathcal{F}(t)\right] \tag{4.69}$$

$$= \frac{M(t)}{P_{n+1,N}(t)}. \tag{4.70}$$

[18]Here we use Bayes Abtract Theorem.

We now introduce the inverse likelihood process

$$L^{-1}(T_n) = \frac{P_{n+1,N}(t)M(T_n)}{P_{n+1,N}(T_n)M(t)}. \tag{4.71}$$

From (4.67)-(4.70), one can show that $L^{-1}(T_n)$ is a $\mathbb{Q}^{n+1,N}$-martingale. Therefore we have that

$$\begin{aligned}
\pi_{PS}(t) &= M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{P_{n+1,N}(T_n)}{M(T_n)}(y_{n,N}(T_n) - K)^+ \,\middle|\, \mathcal{F}(t)\right] && (4.72)\\
&= M(t)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[\frac{L^{-1}(T_n)}{L^{-1}(t)}\frac{P_{n+1,N}(T_n)}{M(T_n)}(y_{n,N}(T_n) - K)^+ \,\middle|\, \mathcal{F}(t)\right] && (4.73)\\
&= M(t)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[\frac{P_{n+1,N}(t)M(T_n)}{P_{n+1,N}(T_n)M(t)}\frac{P_{n+1,N}(T_n)}{M(T_n)}(y_{n,N}(T_n) - K)^+ \,\middle|\, \mathcal{F}(t)\right] && (4.74)\\
&= P_{n+1,N}(t)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(y_{n,N}(T_n) - K)^+ \,\middle|\, \mathcal{F}(t)\right]. && (4.75)
\end{aligned}$$

Since the payoff of a receiver swaption can be seen as proportional to that from a put option, the time t price of the receiver swaption under the swap measure, denoted $\pi_{RS}(t)$, can be derived in exactly the same manner as the payer swaption. This gives us the following two swaption pricing formulas as seen from time 0,

$$\begin{aligned}
\pi_{PS}(0) &= P_{n+1,N}(0)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(y_{n,N}(T_n) - K)^+\right], && (4.76)\\
\pi_{RS}(0) &= P_{n+1,N}(0)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(K - y_{n,N}(T_n))^+\right]. && (4.77)
\end{aligned}$$

These pricing formulas will play an instrumental part when we turn to the pricing of credit risky swaps.

From (4.76) and (4.77) we see that the valuation formulas for swaptions have been greatly simplified. At time zero, the annuity factor is known since the sum of ZCBs can be extracted from market data. Consequently, we only have to compute the expectation of one random variable, the swap rate. Also notice that under the swap measure the swap rate is a martingale. This is realized by looking at the definition of the swap rate and remembering that ZCBs relative to the annuity factor are martingales under the swap measure. The change of numeraire also shows explicitly why swaptions can be viewed as options on swap rates.

Chapter 5

Pricing Using Fourier Inversion

5.1 Introduction

When pricing options as expectations, knowledge about the distribution of the underlying asset is essential. In certain cases where the terminal distribution is known, closed form expressions for European option prices can be derived. A famous example is the Black-Scholes formula for a call option written on a stock. In the Black-Scholes model, the stock evolves according to a geometric Brownian motion and thereby has a lognormal distribution. However, in some cases when the underlying stochastic process is very complex, determining the distribution of the underlying asset is difficult and in special cases requires significant computational power. It turns out that there is an alternative approach to modeling the option payoff directly by a stochastic process. The idea of this alternative approach is to exploit a link between the *Characteristic function* of the underlying asset's density function and the option payoff. This approach enables a derivation of the option price by an *inverse Fourier transform*. In this way, option prices can usually be computed much more easily for complex processes.

There is an increasing interest in applying methods using characteristic functions and Fourier transforms since modern pricing models are often constructed using complex non-Gaussian processes that are more easily characterized through a characteristic function rather than a probability distribution. This is because the characteristic function *completely* defines the probability distribution. Therefore, we wish to present a method of pricing European call and put options using Fourier inversion. We end the chapter by deriving pricing formulas for swaptions according to the formulas stated in Pelsser and Schrager (2006). These formulas will be applied when we turn to computing prices for credit risky swaps.

5.2 The Characteristic Function

We begin with an introduction to characteristic functions. Our introduction is inspired by the ideas in Grimmett and Stirzaker (2001) and Schmelzle (2010).

5.2.1 Definition and Basic Properties

The Characteristic function (CF) plays a very important role in probability theory since it completely defines the probability distribution of any random variable. Even if the density of a random variable does not have an analytical expression, the CF always exists. Thus, every random variable possesses a unique CF. An interesting fact is that there is a one to one relationship between the CF and the probability density function and hence to the distribution function. This implies that any knowledge about the CF also gives information about the distribution.

In order to give a precise definition of the CF, one must use the operator known as the *Fourier transform* (FT). There are many ways to define the FT, so instead of stating many different versions we have chosen a definition that makes the connection to CFs very clear. We define the FT of a function $f : \mathbb{R} \to \mathbb{R}$ satisfying conditions specified below, as

$$\hat{f}(u) \equiv \int_{-\infty}^{\infty} \exp(iux) f(x) dx, \text{ for } u \in \mathbb{R}, \tag{5.1}$$

where $i = \sqrt{-1}$ is the imaginary unit. From this definition we can see that the FT describes a transformation of a real valued function f from $\mathbb{R}$ into the complex space $\mathbb{C}$. Given the FT of f, the function f can be recovered by *Fourier Inversion* (FI), so that

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-iux) \hat{f}(u) du, \text{ for } u \in \mathbb{R}. \tag{5.2}$$

The FT method is a widely used and well understood mathematical tool used extensively in physics and engineering disciplines as for example a method for solving partial differential equations. In finance, the FT method has become a very important tool. It was first applied in the early 90s in Heston (1993) to obtain an analytical pricing formula for European call options with stochastic volatility of the underlying asset.

When applying the theory of FTs, one must be familiar with the notion of absolute integrability. A function is absolutely integrable if the integral of its absolute value in $\mathbb{R}$ is finite,

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty. \tag{5.3}$$

The reason as to why this notion is so important is that in order for the FT and its inverse to exist, the function f must be absolutely integrable and (5.3) must therefore hold. FTs can also be extended to square integrable functions which satisfy

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty. \tag{5.4}$$

These conditions must always be kept in mind when using the operator.

With the definition and properties of the FT in mind we will now state the definition of the CF:

Definition 5.1. *(THE CHARACTERISTIC FUNCTION)*
The characteristic function of a random variable X with density f_X is the function $\phi_X : \mathbb{R} \to \mathbb{C}$ defined by

$$\begin{aligned} \phi_X(u) &= \int_{\mathbb{R}} \exp(iux) f_X(x) dx \\ &= \mathbb{E}\Big[\exp(iuX)\Big] \end{aligned}$$

where $i = \sqrt{-1}$ is the imaginary unit.

Comments. We can see from the definition that the CF of a random variable X is defined as the FT of the density of X.

Certain obvious observations can be made from the definition stated above; for instance, it is clear that $\phi_X(0) = 1$. Also, the CF of a deterministic variable is given as $\phi_X(u) = \exp(iux)$ for $\mathbb{P}(X = x) = 1$. The geometric interpretation of the CF is closely connected to the unit circle in the complex plane which implies that there exists an upper bound of $|\phi_X(u)| \leq 1$, i.e. the norm of the CF is never bigger than one. If the CF is absolutely integrable and therefore satisfies condition (5.3), then X has a continuous probability distribution. Also, the fact that $|\exp(iuX)|$ is a continuous and bounded function for all finite real u and x ensures that $\phi_X(u)$ always exists.

Since the trigonometric functions *cosine* and *sine* are even and odd respectively, we can, by applying Eulers formula[1] to Definition 5.1, show that

$$\begin{aligned}\phi_X(-u) &= \mathbb{E}\Big[\cos(-uX) + i\sin(-uX)\Big] \\ &= \mathbb{E}\Big[\cos(uX) - i\sin(uX)\Big] \\ &= \mathbb{E}\Big[\overline{\exp(iuX)}\Big] \\ &= \overline{\phi_X(u)}. \end{aligned} \tag{5.5}$$

Hence, $\phi_X(-u)$ is the complex conjugate of $\phi_X(u)$[2]. This symmetry around $u = 0$ suggests that we only need to consider the characteristic function for $u > 0$ in order to describe the distribution. Finally, if X and Y are two independent random variables, then for $Z = X + Y$ it holds that

$$\phi_Z(u) = \mathbb{E}\Big[\exp(iu(X+Y))\Big] = \mathbb{E}\Big[\exp(iuX)\Big]\mathbb{E}\Big[\exp(iuY)\Big] = \phi_X(u)\phi_Y(u),$$

and for $a, b \in \mathbb{R}$ such that $Y = a + bX$ we have that

$$\phi_Y(u) = \mathbb{E}\Big[\exp(iu(a+bX))\Big] = \exp(iua)\mathbb{E}\Big[\exp(iubX)\Big] = \exp(iua)\phi_X(ub).$$

5.2.2 The Inversion Theorem

As mentioned, there is a one-to-one relationship between the CF and the distribution of a random variable. This means that we can evaluate the density function and thereby also the distribution function at any point by knowing the CF. This is done via the inverse FT stated in the following theorem:

[1] $\exp(iuX) = \cos(uX) + i\sin(uX)$

[2] The complex conjugate of a complex variable $z = x + iy$ is defined as

$$\overline{z} = x - iy.$$

Theorem 5.1. *If X is a continuous random variable with the CF*

$$\phi_X(u) = E\Big[\exp(iuX)\Big]$$

then the density function of X, denoted by $f_X(x)$, is given by the inverse FT such that

$$f_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \exp(-iux)\phi_X(u)du. \tag{5.6}$$

Then, denoting the distribution function for X by $F_X(x)$, we have that

$$F_X(x) = \mathbb{P}(X \leq x) = \frac{1}{2} - \frac{1}{2\pi}\int_{\mathbb{R}} \frac{\exp(-iux)\phi_X(u)}{iu}du. \tag{5.7}$$

So, in order to calculate the distribution and density function of a random variable, one has to evaluate complex integrals. However, it turns out that (5.6) and (5.7) can be simplified due to the properties of the CF. This can be realized by observing that the FT of a real-valued function is a transformation from $\mathbb{R}$ to $\mathbb{C}$. Hence, it can be written in its complex form where if $z \in \mathbb{C}$ we have $z = a + ib$ for $a, b \in \mathbb{R}$. Here $\Re[z] = a$ is the real part and $\Im[z] = b$ is the imaginary part and we have that

$$\Re[z] = \frac{z + \overline{z}}{2} \text{ and } \Im[z] = \frac{z - \overline{z}}{2i} \tag{5.8}$$

If we consider the CF, which we defined as the FT of a real valued density function, we can use the result in (5.5) to show that

$$\Re[\phi_X(u)] = \frac{\phi_X(u) + \phi_X(-u)}{2} \text{ and } \Im[z] = \frac{\phi_X(u) - \phi_X(-u)}{2i}. \tag{5.9}$$

Combining the above result with the fact that the complex function $f(uX) = \exp(iuX)$ also satisfies $\overline{f(uX)} = f(-uX)$[3] and that the product of two complex conjugated numbers equals the

[3]This can also be realised by applying (5.5) on a degenerate distribution.

conjugate of the product[4], we can simplify the integration in (5.6) to

$$
\begin{aligned}
\int_{-\infty}^{\infty} \exp(-iux)\phi_X(u)du &= \int_{-\infty}^{0} \exp(-iux)\phi_X(u)du + \int_{0}^{\infty} \exp(-iux)\phi_X(u)du && (5.10)\\
&= \int_{0}^{\infty} \exp(iux)\phi_X(-u)du + \int_{0}^{\infty} \exp(-iux)\phi_X(u)du && (5.11)\\
&= \int_{0}^{\infty} \overline{\exp(-iux)\phi_X(u)}du + \int_{0}^{\infty} \exp(-iux)\phi_X(u)du && (5.12)\\
&= \int_{0}^{\infty} \Big(\overline{\exp(-iux)\phi_X(u)} + \exp(-iux)\phi_X(u)\Big)du && (5.13)\\
&= 2\int_{0}^{\infty} \Re\Big[\exp(-iux)\phi_X(u)\Big]du && (5.14)
\end{aligned}
$$

where in (5.13)-(5.14) we have used the results in (5.9). Thus, the density in (5.6) is reduced to

$$
f_X(x) = \frac{1}{\pi}\int_{0}^{\infty} \Re\Big[\exp(-iux)\phi_X(u)\Big]du. \tag{5.15}
$$

Similar calulations for the cumulative distribution function in (5.7) yield

$$
F_X(x) = \frac{1}{2} - \frac{1}{\pi}\int_{0}^{\infty} \Re\left[\frac{\exp(-iux)\phi_X(u)}{iu}\right]du. \tag{5.16}
$$

Thus, calculating the distribution and density function by FI only requires evaluation of real integrals, which are easier to handle numerically.

Often in financial models the underlying asset evolves according to some stochastic process that implies a time-dependent price process expressed by conditional expectations. So, in order to apply the CF in these models we must make the CF time dependent. Thus, for $t < T$ we define the *Conditional Characteristic Function* (CCF) of a stochastic process $\big(X(t)\big)_{t\geq 0}$ as

$$
\phi_X(u,t,T) = \mathbb{E}_t\Big[\exp\left(iuX(T)\right)\Big], \tag{5.17}
$$

making the CCF a time dependent version of the CF.

5.3 Pricing European Options

The goal of this section is to derive a semi-analytical price of an European option using Fourier inversion techniques. To do this, we follow the approach of Carr and Madan (1998) and Lee (2004) where the idea is to calculate the FT of a *modified* call option price w.r.t. the *logarithmic* strike price. A semi-analytical formula for the option price can then be obtained by inverting the FT of the modified call option price. By taking this approach, an efficient algorithm named Fast Fourier Transform (FFT) can be applied in order to numerically evaluate the integral in the formula. With this specification and the FFT routine, a wide range of option prices can be obtained. Although we are not interested in using the FFT routine, we wish to use the same approach in order to establish the semi-analytical formula for the option price.

[4]If $u, w \in \mathbb{C}$, then for $z = uw$ we have that

$$
\overline{z} = \overline{uw}.
$$

Now, if we let $X(t)$ denote the price of the underlying process and K denote the strike price, we define

$$x(t) = \log X(t), \tag{5.18}$$
$$k = \log K. \tag{5.19}$$

Let $\pi_C(t,k)$ denote the time t value of a call option with strike $\exp(k)$ maturing at time T. Now, if we define $q_{t,T}(x)$ as the (conditional) risk neutral density of $x(T)$ at time t, we know that the CCF of $q_{t,T}(x)$ is given by

$$\phi_x(u,t,T) = \int_{-\infty}^{\infty} \exp(iux) q_{t,T}(x) dx. \tag{5.20}$$

For the sake of simplicity, let us assume that the short rate is constant so that the money market account $M(t) = \exp(rt)$ evolves deterministically. Then from Section 4.3 we know that value of the call option at time zero, $\pi_C(0,k)$, can be obtained under the risk-neutral measure as

$$\begin{aligned}
\pi_C(0,k) &= \mathbb{E}_0^{\mathbb{Q}}\Big[\exp(-rT)\pi_C(T,k)\Big] &(5.21)\\
&= \int_{-\infty}^{\infty} \exp(-rT)\Big(\exp(x) - \exp(k)\Big)^+ q_{0,T}(x)dx &(5.22)\\
&= \int_{k}^{\infty} \exp(-rT)(\exp(x) - \exp(k)) q_{0,T}(x)dx &(5.23)
\end{aligned}$$

Notice that when we express the call option in terms of the logarithm of the strike, we have that

$$\begin{aligned}
\lim_{k\to-\infty} \pi_C(0,k) &= \int_{-\infty}^{\infty} \exp(-rT)\exp(x) q_{0,T}(x)dx &(5.24)\\
&= \mathbb{E}_0^{Q}\Big[\exp(-rT)\exp(x(T))\Big] &(5.25)\\
&= X(0) &(5.26)
\end{aligned}$$

since $\exp(-rT)X(T)$ is a martingale under the $\mathbb{Q}$-measure. So $\pi_C(0,k)$ tends to $X(0)$ as k goes to $-\infty$ and therefore the limit of $\pi_C(0,k)$ does not converge to zero. Hence, $\pi_C(0,k)$ does not satisfy condition (5.3) on page 42 and is therefore not absolutely integrable. This means that an FT of $\pi_C(0,k)$ does not exist and a FT of the call option price is therefore not possible. However, if we introduce an exponential *damping* factor $\exp(\alpha k)$ where $\alpha > 0$, we can make $\pi_C(0,k)$ integrable by considering the modified call price $\tilde{\pi}_C(0,k)$ defined by

$$\tilde{\pi}_C(0,k) = \exp(\alpha k)\pi_C(0,k). \tag{5.27}$$

Since $\exp(\alpha k) \to 0$ for $k \to -\infty$, we force the modified call price to converge to zero in the limit. Consequently,

$$\int_{-\infty}^{\infty} |\exp(\alpha k)\pi_C(0,k)| dk < \infty, \tag{5.28}$$

for some suitable choice of α, making the modified call option price absolutely integrable.
The FT of $\tilde{\pi}_C(0,k)$ is then given by

$$\psi_C(u) = \int_{-\infty}^{\infty} \exp(iuk)\tilde{\pi}_C(0,k)dk. \tag{5.29}$$

Next step is to develop an analytical expression for $\psi_C(u)$ in terms of the CF in (5.29). The calculations can be found in Carr and Madan (1998) and we will therefore just state the result here:

$$\psi_C(u) = \frac{\exp(-rT)\phi_x(u-(\alpha+1)i,0,T)}{\alpha^2+\alpha-u^2+i(2\alpha+1)u}, \tag{5.30}$$

where $\phi_x(u-(\alpha+1)i,t,T)$ denotes the CCF of the underlying price process $(x(t))_{t\geq 0}$ measured in $u-(\alpha+1)i$. Now, by Fourier inversion the undampened call option price can be obtained as

$$\tilde{\pi}_C(0,k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-iuk)\psi_C(u)du \tag{5.31}$$

$$\Leftrightarrow \pi_C(0,k) = \frac{\exp(-\alpha k)}{2\pi}\int_{-\infty}^{\infty} \exp(-iuk)\psi_C(u)du \tag{5.32}$$

$$= \frac{\exp(-\alpha k)}{\pi}\int_{0}^{\infty} \Re\Big[\exp(-iuk)\psi_C(u)\Big]du. \tag{5.33}$$

That (5.33) holds can be realised using the same line of reasoning as in (5.10)-(5.14) on page 45, since $\psi_C(u)$ consists of a FT of a real-valued function and therefore has the property

$$\psi_C(-u) = \frac{\exp(-rT)\phi_x(-u-(\alpha+1)i,0,T)}{\alpha^2+\alpha-u^2-i(2\alpha+1)u} = \overline{\psi_C(u)}. \tag{5.34}$$

The call price formula in (5.33) is semi-analytical, so in order to apply the formula one has to compute the real integral numerically. Also, one has to know *either* the FT of the modified call price *or* the CCF for the terminal distribution of the logarithm of the asset process. However, knowledge about the terminal distribution of the underlying asset is not necessary.

Note that the method only makes sense when α is chosen so that the modified call price is well behaved. For positive values of α, using $\exp(\alpha k)$ to modify the call price makes it integrable along the negative log strike axis, but worsens the integrability condition along the positive axis. A sufficient condition required for the modified call price to be integrable along the negative and positive axis, and hence squared integrable, is given by $\psi_C(u) = 0$ being finite[5]. For $u = 0$ we have that

$$\psi_C(0) = \frac{\exp(-rT)\phi_x(-(\alpha+1)i,0,T)}{\alpha^2+\alpha}, \tag{5.35}$$

and in order for (5.35) to be finite we must have that $\phi_x(-(\alpha+1)i,0,T) < \infty$. Using the definition of the CCF this is equivalent to

$$\phi_x(-(\alpha+1)i,0,T) = \mathbb{E}_0\Big[\exp\Big(i(-(\alpha+1)i)x(T)\Big)\Big] = \mathbb{E}_0\Big[\exp\Big((\alpha+1)x(T)\Big)\Big] \tag{5.36}$$

$$= \mathbb{E}_0\Big[X(T)^{\alpha+1}\Big] < \infty. \tag{5.37}$$

Hence, the modified call price $\tilde{\pi}_C(0,k)$ is squared integrable when the conditional moments of order $1+\alpha$ of $X(T)$ exist and are finite.

If we consider the corresponding put option, i.e an option with payoff $\Big(\exp(k)-\exp(x(T))\Big)^+$,

[5] See Carr and Madan (1998) page 64 for further details.

and denote the value at time zero as $\pi_P(0,k)$, we have

$$\pi_P(0,k) = \int_{-\infty}^{\infty} \exp(-rT)\Big(\exp(k) - \exp(x)\Big)^+ q_{0,T}(x)dx. \tag{5.38}$$

Now, if we define the modified put price as

$$\tilde{\pi}_P(0,k) = \exp(-\alpha k)\pi_P(0,k), \tag{5.39}$$

one can show that the FT of the modified put price is given by

$$\psi_P(u) = \frac{\exp(-rT)\phi_x(u-(-\alpha+1)i,0,T)}{\alpha^2 - \alpha - u^2 + i(-2\alpha+1)u}. \tag{5.40}$$

Then by Fourier inversion, the put price can be found as

$$\pi_P(0,k) = \frac{\exp(\alpha k)}{\pi}\int_0^{\infty} \Re\Big[\exp(-iuk)\psi_P(u)\Big]du. \tag{5.41}$$

5.3.1 Pricing Swaptions Using Fourier Inversion

We now turn to the task of deriving the pricing formulas for swaptions stated in Pelsser and Schrager (2006) using the framework described above[6]. If we let $y_{n,N}(t)$ denote the forward starting par swap rate of a swap contract starting at time T_n, paying out for the first time at time T_{n+1} and lasting until time T_N, we have already shown that the time t price of a payer and a receiver swaption with strike K can be expressed according to (4.76) and (4.77) on page 40. Introducing a slightly different notation, we restate the pricing formulas as

$$\begin{aligned}
\pi_{PS}(t,K) &= P_{n+1,N}(t)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[\Big(y_{n,N}(T_n) - K\Big)^+ \,\Big|\, \mathcal{F}(t)\right] &(5.42)\\
&= P_{n+1,N}(t)\int_{-\infty}^{\infty}\Big(y-K\Big)^+ q_{t,T_n}(y)dy, &(5.43)\\
\pi_{RS}(t,K) &= P_{n+1,N}(t)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[\Big(K - y_{n,N}(T_n)\Big)^+ \,\Big|\, \mathcal{F}(t)\right] &(5.44)\\
&= P_{n+1,N}(t)\int_{-\infty}^{\infty}\Big(K-y\Big)^+ q_{t,T_n}(y)dy, &(5.45)
\end{aligned}$$

where $\mathbb{Q}^{n+1,N}$ denotes the swap measure and $q_{t,T_n}(y)$ denotes the (conditional) density of $y_{n,N}(T_n)$ under the swap measure at time t. Now, the CCF of the swap rate is defined as

$$\phi_y(u,t,T_n) = \mathbb{E}_t^{\mathbb{Q}^{n+1,N}}\Big[\exp(iuy_{n,N}(T_n))\Big] = \int_{-\infty}^{\infty}\exp(iuy)q_{t,T_n}(y)dy. \tag{5.46}$$

Following the same approach as in the previous section, we define the modified payer and receiver swaption price as

$$\begin{aligned}
\tilde{\pi}_{PS}(t,K) &= \exp(\alpha K)\pi_{PS}(t,K), &(5.47)\\
\tilde{\pi}_{RS}(t,K) &= \exp(-\alpha K)\pi_{RS}(t,K), &(5.48)
\end{aligned}$$

[6] Pelsser and Schrager (2006) only states the payer swaption price whereas we will state both the payer and the receiver swaption price.

for some $\alpha \in \mathbb{R}_+$. This means that we can establish the FT of the modified swaption prices as

$$\psi_P(u,t,T_n) = \int_{-\infty}^{\infty} \exp(iuK)\tilde{\pi}_{PS}(t,K)dK, \tag{5.49}$$
$$\psi_R(u,t,T_n) = \int_{-\infty}^{\infty} \exp(iuK)\tilde{\pi}_{RS}(t,K)dK. \tag{5.50}$$

Now, the price of a payer and a receiver swaption can be stated in the following proposition:

Proposition 5.1. *Let $y_{n,N}(t)$ denote the forward par swap rate under the swap measure $\mathbb{Q}^{n+1,N}$ for a swap contract starting at time T_n, paying out for the first time at time T_{n+1} and lasting until time T_N. Assume that an $\alpha \in \mathbb{R}_+$ exists such that* (5.47) *and* (5.48) *are absolutely integrable.*
The time t value of a payer swaption then equals

$$\pi_{PS}(t,K) = \frac{\exp(-\alpha K)}{\pi}\int_0^{\infty} \Re\Big[\exp(-iuK)\psi_P(u,t,T_n)\Big]du, \tag{5.51}$$
$$\psi_P(u,t,T_n) = P_{n+1,N}(t)\frac{\phi_y(u-i\alpha,t,T_n)}{(iu+\alpha)^2} \tag{5.52}$$

and the corresponding receiver swaption

$$\pi_{RS}(t,K) = \frac{\exp(\alpha K)}{\pi}\int_0^{\infty} \Re\Big[\exp(-iuK)\psi_R(u,t,T_n)\Big]du, \tag{5.53}$$
$$\psi_R(u,t,T_n) = P_{n+1,N}(t)\frac{\phi_y(u+i\alpha,t,T_n)}{(iu-\alpha)^2} \tag{5.54}$$

where ϕ_y denotes the CCF of $y_{n,N}(T_n)$ defined in (5.46).

Proof. We will only state the proof for the payer swaption, since the approach in the proof for the receiver swaption is exactly the same. We begin the proof by deriving (5.52):

$$\psi(u,t,T_n) = \int_{-\infty}^{\infty} \exp(iuK)\tilde{\pi}_{PS}(t,K)dK \tag{5.55}$$
$$= \int_{-\infty}^{\infty} \exp(iuK)\exp(\alpha K)\pi_{PS}(t,K)dK \tag{5.56}$$
$$= \int_{-\infty}^{\infty} \exp(iuK)\exp(\alpha K)P_{n+1,N}(t)\int_{-\infty}^{\infty}(y-K)^+ q_{t,T_n}(y)dydK \tag{5.57}$$
$$= \int_{-\infty}^{\infty} P_{n+1,N}(t)\int_{-\infty}^{\infty}\exp(K(iu+\alpha))(y-K)^+ q_{t,T_n}(y)dydK \tag{5.58}$$
$$= \int_{-\infty}^{\infty} P_{n+1,N}(t)q_{t,T_n}(y)\int_{-\infty}^{\infty}\exp(K(iu+\alpha))(y-K)^+ dKdy \tag{5.59}$$
$$= \int_{-\infty}^{\infty} P_{n+1,N}(t)q_{t,T_n}(y)\int_{-\infty}^{y}\exp(K(iu+\alpha))(y-K)dKdy \tag{5.60}$$

In (5.58)-(5.59) we have used Fubini's rule[7] to interchange the integration order, which is justified

[7] Consider a function $f: X\times Y \to \mathbb{R}$ and assume that f is integrable over $X\times Y$. Then

$$\int_X\int_Y f(x,y)dydx = \int_Y\int_X f(x,y)dxdy.$$

since the modified payer swaption is absolutely integrable. In (5.60) we have changed the upper bound of the inner integral to y, since any strike larger than the swap rate would provide a payoff of zero. Now, for the inner integral in (5.60), we have that

$$\int_{-\infty}^{y} \exp(K(iu+\alpha))(y-K)dK \tag{5.61}$$

$$= y\int_{-\infty}^{y} \exp(K(iu+\alpha))dK - \int_{-\infty}^{y} \exp(K(iu+\alpha))KdK \tag{5.62}$$

$$= \frac{y}{iu+\alpha}\Big[\exp(K(iu+\alpha))\Big]_{-\infty}^{y} \tag{5.63}$$

$$- \left(\frac{1}{iu+\alpha}\Big[\exp(K(iu+\alpha))K\Big]_{-\infty}^{y} - \int_{-\infty}^{y} \frac{1}{(iu+\alpha)} \exp(K(iu+\alpha))dK\right) \tag{5.64}$$

where we, in the last equality, have used integration by parts. Since $\alpha > 0$, we have that $\lim_{K\to-\infty} \exp(K(iu+\alpha)) = 0$ and $\lim_{K\to-\infty} K\exp(K(iu+\alpha)) = 0$. Therefore (5.64) reduces to

$$\frac{1}{(iu+\alpha)^2}\exp(y(iu+\alpha)). \tag{5.65}$$

Hence,

$$\psi(u,t,T_n) = \int_{-\infty}^{\infty} P_{n+1,N}(t)q_{t,T_n}(y)\frac{1}{(iu+\alpha)^2}\exp(y(iu+\alpha))dy \tag{5.66}$$

$$= \frac{P_{n+1,N}(t)}{(iu+\alpha)^2}\int_{-\infty}^{\infty} \exp(y(iu+\alpha))q_{t,T_n}(y)dy \tag{5.67}$$

$$= \frac{P_{n+1,N}(t)}{(iu+\alpha)^2}\int_{-\infty}^{\infty} \exp(i(u-i\alpha)y)q_{t,T_n}(y)dy \tag{5.68}$$

$$= P_{n+1,N}(t)\frac{\phi_y(u-i\alpha,t,T_n)}{(iu+\alpha)^2}, \tag{5.69}$$

which proves (5.52). Now, by Fourier inversion we get

$$\pi_{PS}(t,K) = \frac{\exp(-\alpha K)}{2\pi}\int_{-\infty}^{\infty} \exp(-iuK)\psi(u,t,T_n)du \tag{5.70}$$

$$= \frac{\exp(-\alpha K)}{\pi}\int_{0}^{\infty} \Re\Big[\exp(-iuK)\psi(u,t,T_n)\Big]du. \tag{5.71}$$

□

So, in order to price swaptions using FI as in Proposition 5.1, one needs to stipulate the CCF of the swap rate. This in turn depends on the specific dynamics chosen for the swap rate. Once the CCF is derived based on the dynamics, a proper value of the parameter α needs to be determined in order to ensure integrability. In Lee (2004) general formulas combined with upper and lower bounds are derived for α depending on the specific option payoff. One can also use other well known models in order to choose α, for example the Black-Scholes model. Since the Black-Scholes model has an analytical solution, we know that the numerical implementation must converge to the same theoretical price for a standard call or put option. Hence, a suitable choice of α can be provided by comparing the two prices.

Chapter 6

Affine Models

In order to price swaptions we will need to specify an interest rate model that enables both pricing of ZCBs as well as a specification of the characteristic function in order to apply Proposition 5.1 on page 49. It turns out that a certain tractable subclass of models, namely the *affine models*, manages to satisfy *both* criteria in a very convenient and explicit way. The subclass also allows for a wide variety of classic models to choose among, e.g. the Vasicek model (Vasicek, 1977) or CIR model (Cox et al., 1985).

In order to derive default probabilities, the affine models will also be key. They will enable the derivation of a complete set of default probabilities from CDS quotes without just interpolating between the bootstrapped default probabilities.

The chapter is inspired by Duffie and Kan (1996), Munk (2005), Dai and Singleton (2000), Singleton and Umantsev (2002a), Duffee (1998), Duffie et al. (2000) and Pelsser and Schrager (2006).

6.1 Construction

Consider a vector of N state variables $X(t)$ where each state variable is called a *factor*. $X(t)$ is assumed to be representative of the *state* of the economy. The dynamics, i.e. the *instantaneous increments*, are then described by an *N-factor* diffusion model

$$dX(t) = \Psi\left(\Theta - X(t)\right)dt + \Sigma\sqrt{V(t)}dW^{\mathbb{Q}}(t) \tag{6.1}$$

where Ψ and Σ are $N \times N$ matrices and where Θ is an N-dimensional vector. $V(t)$ is a diagonal matrix and $W^{\mathbb{Q}}(t)$ is an N-dimensional independent standard Brownian motion under the risk neutral measure $\mathbb{Q}$. The specification could albeit in principle be under some other probability measure. In the most general setting, the matrices are allowed to be time varying[1]. To ensure existence of the process $X(t)$, equation (6.1) must have a unique solution that requires $\Psi, \Psi\Theta$ and Σ to satisfy certain regulatory conditions[2]. Restrictions on $V(t)$ are also necessary since the square root computation[3] might have multiple solutions if $V(t)$ is not positive definite. However, if $V(t)$ *is* positive definite, a unique solution is guaranteed.

[1]By following Pelsser and Schrager (2006) in using an affine framework to price swaptions, we actually have to specify dynamics under a certain measure so that Ψ, Θ and Σ are time varying.

[2]See Øksendal (1998).

[3]The square root of a square matrix S is a square matrix R so that $RR = S$.

Definition 6.1. *A stochastic process $r(t)$ is driven by an affine model if (and only if) the following two relationships hold*

$$r(t) = \omega_0 + \omega_X^\intercal X(t)$$
$$V(t)_{ii} = \alpha_i + \beta_i^\intercal X(t) \geq 0$$

where

- $X(t)$ *is specified according to equation* (6.1)
- ω_0 *and* α_i *are scalars*
- ω_X *and* β_i *is* N*-dimensional vectors.*

Note how *both* the relationship between $r(t)$ and the state vector $X(t)$ and the relationship between the *variances* $V(t)_{ii}$ and $X(t)$ are affine. The latter relationship is particularly interesting since it allows for square root specification of the *volatility term* that is used in e.g. the CIR model. The definition implies that $V(t)$ by matrix notation is given as

$$V(t) = \text{diag}(\alpha + \beta X(t)) \tag{6.2}$$

where the matrix β is defined as $[\beta_1 \cdots \beta_N]^\intercal$ and the vector α is defined as $[\alpha_1 \cdots \alpha_N]^\intercal$.

The stochastic process $r(t)$ could typically model the short interest rate (which it mainly will for our purpose), but it could also model some other variable e.g. a default intensity (see Chapter 9).

6.2 Implications

6.2.1 Transitioning into the $\mathbb{P}$-measure

In the construction of the affine setup, we have deliberately defined the underlying processes under the risk neutral measure $\mathbb{Q}$. It is however possible to backtrack to the real world measure $\mathbb{P}$ by assuming that the market prices of risk, $\Lambda(t)$, are given by

$$\Lambda(t) = \sqrt{V(t)}\lambda \tag{6.3}$$

where λ is an N-dimensional vector of constants. Straightforward calculations can then show[4] that the dynamics for $X(t)$ under $\mathbb{P}$ are given by an *affine form*

$$dX(t) = \Psi_\mathbb{P} \left(\Theta_\mathbb{P} - X(t)\right) dt + \Sigma\sqrt{V(t)}dW^\mathbb{P}(t) \tag{6.4}$$

where $\Psi_\mathbb{P} = \Psi - \Sigma\Phi$ and $\Theta_\mathbb{P} = \Psi_\mathbb{P}^{-1} \left(\Psi\Theta + \Sigma\psi\right)$, given that the ith row of Φ is given by $\lambda_i\beta_i^\intercal$ and ψ is an N-dimensional vector whose ith element is given by $\lambda_i\alpha_i$.

[4] See Duffee (1998)

The important statement is that under *both* measures ($\mathbb{P}$ and $\mathbb{Q}$) the *affine form* of the diffusion process given by equation (6.1) is maintained, and only the drift is truly changed. This fact stems from the Girsanov Theorem, see equation (4.15) on page 28.

6.2.2 The Price of a Zero Coupon Bond

The affine framework has a very explicit way of characterizing the cardinal expression

$$D(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(s)ds\right) \middle| \mathcal{F}(t)\right] \tag{6.5}$$

where $r(t)$ denotes a short rate process[5]:

> **Proposition 6.1.** *Assume a short rate process $r(t)$ so that Definition 6.1 is fulfilled and the ZCB price is given as the expectation in equation* (6.5)*. Then the following closed-form ZCB price, which is completely determined by the specification of the risk-neutral short rate dynamics given by relationship* (6.1) *and Definition 6.1, exists:*
>
> $$D(t,T) = \exp\left(A(t,T) - B(t,T)^\intercal X(t)\right)$$
>
> *where the* one-*dimensional function $A(t,T)$ and the N-dimensional function $B(t,T)$ satisfy the ordinary differential equations (ODEs)*
>
> $$\frac{dA(t,T)}{dt} = -\Theta^\intercal\Psi^\intercal B(t,T) + \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma^\intercal B(t,T)\right]_i^2 \alpha_i - \omega_0, \tag{6.6}$$
>
> $$\frac{dB(t,T)}{dt} = -\Psi^\intercal B(t,T) \quad - \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma^\intercal B(t,T)\right]_i^2 \beta_i + \omega_X, \tag{6.7}$$
>
> *with the boundary conditions $A(T,T) = 0$ and $B_i(T,T) = 0 \ \forall\ i$.*

Proof. See Duffie and Kan (1996). The idea is to show that

$$D(t,T) = \exp\left(A(t,T) - B(t,T)^\intercal X(t)\right) \tag{6.8}$$

is a solution to the *fundamental PDE*[6]. By inserting the derivatives of (6.8) in the PDE, one can show that the PDE only holds if the *exact* ODEs given by (6.6) and (6.7) hold. ☐

The Proposition states that the ZCB price is given on semi-analytical *exponential affine* form. The second ODE is of the *Ricatti* type, which means that it is quadratic in the unknown function[7].

[5] Later we will see that the expression given by (6.5) is also central in chapter 9: *Intensity Modelling*, where $r(s)$ will be substituted by a default intensity $\lambda(s)$.

[6] See e.g. Björk (2009) Chapter 13.

[7] Technically speaking a Ricatti equation is an ODE on the form

$$\frac{dy(x)}{dx} = q_0(x) + q_1(x)y(x) + q_2(x)y(x)^2$$

The ODEs will typically be solved using numerical integration. However, for the one-dimensional cases considered by Vasicek (1977) and Cox et al. (1985) there are explicit solutions for (A, B).

6.2.3 The Zero Coupon Bond Price Process

In order to price swaptions, we have to know the dynamics of the ZCB price. It turns out that this process is (also) remarkably simple in an affine setup.

From Proposition 6.1 we have that the ZCB price is given on exponential affine form. Applying Itô's Formula on this expression and using the dynamics of $X(t)$, given by equation (6.1), yield

$$\begin{aligned} dD(t,T) &= \left(\frac{\partial A}{\partial t}(t,T) - \frac{\partial B^\intercal}{\partial t}(t,T)X(t)\right) dt - B(t,T)^\intercal D(t,T) dX(t) \\ &\quad + \frac{1}{2} B(t,T)B(t,T)^\intercal D(t,T)(dX(t))^2 \qquad (6.9) \\ &= \left(\frac{\partial A}{\partial t}(t,T) - \frac{\partial B^\intercal}{\partial t}(t,T)X(t)\right) dt \\ &\quad - B(t,T)^\intercal D(t,T) \left((\Psi(\Theta - X(t)))dt + \Sigma\sqrt{V(t)} dW^{\mathbb{Q}}(t) \right) \\ &\quad + \frac{1}{2} B(t,T)B(t,T)^\intercal D(t,T)\Sigma V(t)\Sigma^\intercal dt \qquad (6.10) \\ &= \Bigg(\left(\frac{\partial A}{\partial t}(t,T) - \frac{\partial B^\intercal}{\partial t}(t,T)X(t)\right) \\ &\quad - B(t,T)^\intercal \left(\Psi(\Theta - X(t))\right) + \frac{1}{2}\Sigma V(t)\Sigma^\intercal \Bigg) D(t,T) dt \\ &\quad - B(t,T)^\intercal D(t,T)\Sigma\sqrt{V(t)} dW^{\mathbb{Q}}(t). \qquad (6.11) \end{aligned}$$

Assuming no arbitrage, we know the drift from Section 4.3 (especially equation (4.37) on page 34) and can thus obtain the simple dynamics:

$$dD(t,T) = r(t)D(t,T)dt - B(t,T)^\intercal D(t,T)\Sigma\sqrt{V(t)} dW^{\mathbb{Q}}(t) \qquad (6.12)$$

We can observe that the volatility term is negative so that the dynamics are consistent with the fact that ZCB prices should diminish when interest rates rise.

6.2.4 The Characteristic Function

We remind ourselves that the (conditional) characteristic function given by equation (5.17) on page 45 in terms of $X(T)$ can by stated as

$$\phi_X(v,t,T) = \mathbb{E}\left[\exp\left(ivX(T)\right) \,\middle|\, \mathcal{F}(t)\right]. \qquad (6.13)$$

Duffie et al. (2000) then show that when assuming an affine framework as described in the above section, the characteristic function is *also* given on exponential affine form where the terms in the affine exponent are given as the solution to a set of *complex* ODEs. This result is of utmost

where $q_0(x) \neq 0$ and $q_2(x) \neq 0$.

importance since it is a cornerstone in the swaption pricing framework given in Pelsser and Schrager (2006).

Proposition 6.2. *Assume an affine framework described according to Definition 6.1. The characteristic function of the random variable X is then given as*

$$\phi_X(v,t,T) = \exp\Big(\gamma(t) + \delta(t)\cdot X(t)\Big)$$

where $\delta(t)$ and $\gamma(t)$ are given as the solution to the following set of complex ODEs

$$\begin{aligned}\frac{d\delta(t)}{dt} &= \Psi^\intercal\delta(t) - \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma^\intercal\delta(t)\right]_i^2\beta_i,\\ \frac{d\gamma(t)}{dt} &= -\Theta^\intercal\Psi\delta(t) - \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma^\intercal\delta(t)\right]_i^2\alpha_i,\end{aligned}$$

with the terminal conditions

$$\begin{aligned}\delta(T) &= iv,\\ \gamma(T) &= 0.\end{aligned}$$

Proof. We will only provide a brief sketch of the proof–see Duffie et al. (2000) for a rigorous version. The idea is to assume that the solution is correct and then show that this is the case when the two ODEs are fulfilled.

The characteristic function is by equation (6.13) given as

$$\phi_X(v,t,T) = \mathbb{E}\left[\exp\left(ivX(T)\right)\,\Big|\,\mathcal{F}(t)\right]$$

so that

$$\phi_X(v,T,T) = \exp\left(ivX(T)\right)$$

since ϕ is adapted to $\mathcal{F}(t)$. By assuming that

$$\phi_X(v,t,T) = \exp\Big(\gamma(t) + \delta(t)\cdot X(t)\Big)$$

we can derive the dynamics for $\phi_X(u,t,T)$ by Itô's Formula. By translating the dynamics into the corresponding integral equation and evaluating at time T, we then have two expressions for $\phi_X(u,T,T)$ that must be equal. It is easily seen that at time T, γ must be zero and δ must be iu in order to equal the time T value of equation (6.13) and equation (6.14). By using these terminal conditions, the ODEs can be established.

□

Technical comments. First of all, we would like to emphasise that in their original paper Duffie et al. (2000) include jumps in their specification of $X(t)$ which gives each of the ODEs an extra term. Secondly, they use a *discounted* characteristic function which again results in each of their ODEs having an extra term. Thirdly, Duffie et al. (2000) have not specified a covariance matrix Σ and are using a matrix notation that gives their ODEs a slightly different look. Finally, the result stated in Pelsser and Schrager (2006) accidently lacks some signs[8] which (along with our other arguments) might confuse a reader trying to compare the different settings.

6.2.5 The One-Factor CIR model

One of the most popular affine one-factor models is the CIR model named after Cox et al. (1985). In the original paper, Cox, Ingersroll, and Ross assume that the short rate follows a *square root process* so that

$$dr(t) = \psi(\theta - r(t))dt + \sigma\sqrt{r(t)}dW^{\mathbb{Q}}(t). \tag{6.14}$$

Thus, compared with the generic affine framework we have the two simple relationships

$$X(t) = r(t) = V(t), \tag{6.15}$$

which implies that $\omega_0 = 0$, $\omega_X = 1$, $\alpha_1 = 0$ and $\beta_1 = 1$. The CIR model for the short rate exhibits means reversion around a long term level θ–just as the classic model by Vasicek (1977)[9].

One of the obvious benefits of the CIR-model is that the square root factor $\sqrt{r(t)}$ guarantees that the short rate stays non-negative which is contrary to the Vasicek model. It can be shown that if $2\theta \geq \sigma^2$, the positive drift at low values of the process is so big in relation to the volatility that the process cannot reach zero and thus stays strictly positive. Another implication of the CIR model concerns the future distribution of interest rates. Given the interest rate at time t, the future value of the interest rate is non-central χ^2-distributed[10]. This makes the distribution more complex than in the Vasicek case where the corresponding distribution simply is Gaussian. One could say that the non-negative interest rates in the CIR model come at a cost of simplicity in the interest rate distribution.

As mentioned in the previous subsection, the CIR (and Vasicek) model implies that the ODEs

[8]Pelsser and Schrager (2006) have stated a minus too much in their first terminal condition regarding $\delta(T)$. Furthermore, their second ODE regarding $\gamma(t)$ lacks a minus in the very beginning.

[9]In the one-factor Vasicek model the short rate is assumed following an Ornstein-Uhlenbeck process:

$$dr(t) = \psi(\theta - r(t))dt + \sigma dW^{\mathbb{Q}}(t)$$

[10]The χ^2-distribution with a degrees of freedom and non-centrality parameter b has the density

$$f_{\chi^2(a,b)} = \sum_{i=0}^{\infty} \frac{\exp\left(\frac{-b}{2}\right)\frac{b}{2}^i\left(\frac{1}{2}\right)^{i+\frac{a}{2}}}{i!\Gamma\left(i+\frac{a}{2}\right)} y^{i-1+\frac{a}{2}} \exp\left(-\frac{y}{2}\right)$$

where Γ denotes the so-called *Gamma-function* defined as

$$\Gamma(m) = \int_0^{\infty} x^{m-1}\exp(-x)dx.$$

given by (6.6) and (6.7), with the afore mentioned conditions, have explicit solutions. For a one-factor CIR model the solutions are

$$A(t,T) = \frac{2\psi\theta}{\sigma^2}\left(\log(2\xi) + \frac{1}{2}(\psi+\xi)(T-t) - \log((\xi+\psi)(\exp(\xi(T-t))-1)+2\xi)\right) \tag{6.16}$$

$$B(t,T) = \frac{2(\exp(\xi(T-t))-1)}{(\xi+\psi)(\exp(\xi(T-t))-1)+2\xi} \tag{6.17}$$

where $\xi = \sqrt{\psi^2 + 2\sigma^2}$.

The solutions might look complex, but they are actually very easy to implement since the specification parameters under the $\mathbb{Q}$-measure can by plugged directly into the solution. The CIR model will constitute our model foundation in our empirical investigations.

Chapter 7

Pricing Counterparty Credit Risk

In the previous chapters of Part II we have neglected the fact that a given counterparty might default and thus fail to pay his obligations. This risk will now be incorporated in our framework. As we argued in the beginning of this thesis paper, counterparty credit risk has become increasingly important in recent years and should thus be taken into account. In this chapter we will examine our choice of handling counterparty credit risk; to price it!

We place ourselves in the position of an assumed default-free counterparty, entering some financial contract with another counterparty who has some positive probability of default before the contract reaches maturity. Thus, we neglect our own default risk, because if we default then we are out of the game anyway. This sort of setup is referred to as the *unilateral* case since only one of the two counterparties is assumed *defaultable*[1]. We assume absence of any form of collateral. If the contract implies a positive probability of an event in which the second counterparty must deliver some positive cash flow to the first counterparty, then it makes sense that the *value* of being at the first counterparty's side of the contract should be smaller than the value of being in a similar contract with a default-free counterparty.

We have two objectives. The first one is to derive a generic formula which states the price of *any* contract which is subject to counterparty credit risk. As we will see, the value of the risky contract is the difference between the value of the risk-free contract and some component. Since this component is indeed the difference between the risk-free and the risky contract, it is intuitive to think of this component as a measure of counterparty credit risk. Hence, the component has been given the name CVA–the Credit Value Adjustment.

The second objective is to use the generic formula on our choice of contract which is an interest rate swap. Using the properties of such a contract enables us to reform the pricing formula to a more tangible version. As we will see, the CVA of a swap equals a sum of *swaption* prices, each multiplied by the probability of a default in the time span of the swaption in question.

7.1 Pricing of a Credit Risky Contract in the General Case

Let us turn to the general pricing expression. Our derivation and presentation of the formula falls somewhere between Gregory (2010) and Brigo and Masetti (2005). The latter were the first to give a rigorous proof for the generic unilateral pricing formula.

[1]As mentioned in Chapter 1, a more complex setup would be the bilateral one.

Let us begin by denoting the price of a ZCB initiated at time t with maturity at time T by $D(t,T)$ as usual. We will denote the time of maturity of the contract by T and the default-time of the (secondary) counterparty by τ. At time t we will denote the value of the *risk-free* claim by $\pi(t,T)$ and the value of the corresponding *risky* claim by $\tilde{\pi}(t,T)$. Note that π and $\tilde{\pi}$ are *values*, which means that they consist of a sum of cashflows and discountings, e.g. $\pi(t,T) = \sum_{i\in[t;T]} D(t,i)CF(i)$, where $CF(i)$ denotes the cashflow at time i which can be either positive or negative. One could say that $\pi(t,T)$ is the Net Present Value (NPV) of future cashflows. The recovery size, which is defined as the (assumed constant) fraction of the *risk-free* contract value which is payed in case of default, is denoted R. We let $\pi(t,T)^+ = \max\left(\pi(t,T),0\right)$ and $\pi(t,T)^- = \min\left(\pi(t,T),0\right)$.

Now, if $\tau > T$ there is no default in the time span of the contract and the second part of the contract will meet its obligations. On the other hand, if $\tau < T$, the second counterparty will *not* meet (all of) its obligations and a default occurs. With this in mind, the following proposition should make sense.

Proposition 7.1. *At valuation time t, and provided that the counterparty has not defaulted before time t, the price of a contract under counterparty credit risk is*

$$\tilde{\pi}(t,T) = \pi(t,T) - (1-R)\mathbb{E}^{\mathbb{Q}}\left[1_{\{t<\tau\leq T\}} D(t,\tau)\pi(\tau,T)^+ \,\middle|\, \mathcal{F}(t)\right]$$

where $\pi(\tau,T)^+$ denotes the value of the at time τ remaining positive cashflows.

Proof. Since the no-arbitrage price of a contract can be written as the expectation (under the risk-neutral measure) of the discounted future cashflows, the price can be written as the following:

$$\tilde{\pi}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}\pi(t,T) + 1_{\{\tau\leq T\}}\pi(t,\tau) + 1_{\{\tau\leq T\}}D(t,\tau)\left(R\pi(\tau,T)^+ + \pi(\tau,T)^-\right) \,\middle|\, \mathcal{F}(t)\right] \tag{7.1}$$

The first term is trivial since the value is $\pi(t,T)$ when there is no default before maturity. The second part is equal to the cashflows paid up to the default time. The third part catches the fact that at time τ we receive only a fraction R of our expected positive cashflows, but that we still have to pay the *full amount* of our expected *negative cashflows*. Using the fact that $\pi(\cdot)^- = \pi(\cdot) - \pi(\cdot)^+$, and rearranging the parts, we obtain:

$$\begin{aligned}\tilde{\pi}(t,T) &= \mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}\pi(t,T) + 1_{\{\tau\leq T\}}\pi(t,\tau) + 1_{\{\tau\leq T\}}D(t,\tau)\left((R-1)\pi(\tau,T)^+ + \pi(\tau,T)\right) \,\middle|\, \mathcal{F}(t)\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}\pi(t,T) + 1_{\{\tau\leq T\}}\left(\pi(t,\tau) + D(t,\tau)\pi(\tau,T)\right) + 1_{\{\tau\leq T\}}D(t,\tau)(R-1)\pi(\tau,T)^+ \,\middle|\, \mathcal{F}(t)\right]\end{aligned} \tag{7.2}$$

Since

$$\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau\leq T\}}\pi(t,\tau) + 1_{\{\tau\leq T\}}D(t,\tau)\pi(\tau,T) \,\middle|\, \mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau\leq T\}}\pi(t,T) \,\middle|\, \mathcal{F}(t)\right] \tag{7.3}$$

we can again rearrange so that we get:

$$\tilde{\pi}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}\pi(t,T) + 1_{\{\tau\leq T\}}\pi(t,T) + 1_{\{\tau\leq T\}}D(t,\tau)\left((R-1)\pi(\tau,T)^{+}\right)\,\middle|\,\mathcal{F}(t)\right] \tag{7.4}$$

Finally, using that $1_{\{\tau>T\}}\pi(t,T)+1_{\{\tau\leq T\}}\pi(t,T) = \pi(t,T)$, that $\pi(t,T)$ is $\mathcal{F}(t)$-measureable, that R is assumed constant and changing sign inside and outside the third term, we obtain the result:

$$\tilde{\pi}(t,T) = \pi(t,T) - (1-R)\mathbb{E}^{\mathbb{Q}}\left[1_{\{t<\tau\leq T\}}D(t,\tau)\pi(\tau,T)^{+}\,\middle|\,\mathcal{F}(t)\right] \tag{7.5}$$

□

The proposition shows the relationship between the risky and risk-free contract price which is indeed the CVA. This CVA is also called the Expected Loss (EL) since it measures the expected loss taking the probability of default into account. Considering the formula, the risky and risk-free prices are equal in case of zero chance of default before maturity which seems natural. Note that under the assumption of independence between default probability and interest rates we obtain

$$\tilde{\pi}(t,T) = \pi(t,T) - (1-R)\mathbb{Q}(t<\tau\leq T)\mathbb{E}^{\mathbb{Q}}\left[D(t,\tau)\pi(\tau,T)^{+}\,\middle|\,\mathcal{F}(t)\right], \tag{7.6}$$

since the expectation to an indicator is the probability of the event associated with the indicator.

The character of the CVA is interesting since it has an option-like expression. Because of the factor $\left(\pi(\tau,T)\right)^{+}$, the CVA can be seen as a European call option with zero strike written on the residual NPV of future cashflows. Counterparty risk thus adds an optionality level to the original payoff, which can complicate the pricing significantly since we now have to price *an option* on the future cashflows. Option pricing usually requires a model which *might* be in contrast to the traditional pricing of the contract in question. This is the case for e.g. interest rate swaps, forwards, futures and possibly CDSs[2] since the risk-free pricing of such contracts usually do not require a model. But as we have now shown, the *risky* price does *always* require a model. In theory, the pricing of a risk free swap is model-independent, when assuming that the present forward rates are the true future interest rates, but interest rate *option* pricing is *not* model-independent. Specifically, one has to specify the interest rate dynamics in order to either simulate or achieve some sort of (maybe semi-) analytical solution to the expectation.

Finally, we must note that while we have argued that the CVA mathematically is equal to the price of an option, the similarity only holds in quantitative terms. Reason being, that a true call option is an instrument in which the owner has *the right*, but not an obligation, to buy a certain asset. This is, of course, in contrast to the CVA which does not resemble a right in any way. Hence, the CVA is *not* an option, but the pricing is.

7.2 Pricing a credit risky interest rate swap

We will now derive and describe the price of a risky swap. Our inspiration is mainly Brigo and Masetti (2005). Consider a contract where we enter a payer (receiver) swap with a risky counterparty. This means that at times $T_{a+1},\ldots,T_b$ we pay (receive) a fixed rate K and receive (pay)

[2] As we will see later on, the price of a CDS *can* be obtained without specifying a model for the default probabilities. This will however infer some assumptions regarding observed CDS quotes.

the floating rate L, which would typically be some xIBOR rate. We take a unit notional on the swap and denote the *year-fraction* between time T_{i-1} and T_i by Δ_i.

According to Proposition 7.1 we can write the value of a risky swap as

$$\tilde{\pi}(t; T_a, T_b) = \pi(t; T_a, T_b) - \mathrm{CVA}(t). \tag{7.7}$$

We can then show that the following proposition holds.

Proposition 7.2. *At valuation time t, provided that the counterparty has not defaulted before time t and assuming independence between τ and interest rates, the value of the CVA for an interest rate swap is*

$$CVA(t) = (1-R)\int_{T_a}^{T_b} \pi_{swaption}\Big(t; s, T_b, K, y_{s,b}(t)\Big) q_t(s) ds$$

where $\pi_{swaption}\Big(t; s, T_b, K, y_{s,b}(t)\Big)$ is the price at time t of a swaption with maturity s, underlying forward swap rate $y_{s,b}(t)$ which is defined from equation (3.14) *on page 16, final maturity of the underlying swap T_b and a strike $K = y_{a,b}(t)$ so that the strike corresponds to the fixed rate that renders the swap value zero. $q_t(s)$ denotes the time t density of τ and the swaption* type *(payer/receiver) corresponds to the swap in question.*

Comments. Note immediately, that the integral varies by s over the time span of the swap $[T_a; T_b]$. This means that there is an infinite number of swaptions to be calculated–one for each swaption maturity s.

Proof. Given that we are considering a swap with the above specifications, we use Proposition 7.1 to argue as follows.

$$\begin{aligned}
\mathrm{CVA}(t) &= (1-R)\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau \le T_b\}} D(t,\tau)\pi(\tau, T_b)^+ \,\middle|\, \mathcal{F}(t)\right] && (7.8)\\
&= (1-R)\mathbb{E}^{\mathbb{Q}}\left[\int_{T_a}^{T_b} q_t(s) D(t,s)\pi(s,T_b)^+ ds \,\middle|\, \mathcal{F}(t)\right] && (7.9)\\
&= (1-R)\int_{T_a}^{T_b} \mathbb{E}^{\mathbb{Q}}\left[q_t(s) D(t,s)\pi(s,T_b)^+ \,\middle|\, \mathcal{F}(t)\right] ds && (7.10)\\
&= (1-R)\int_{T_a}^{T_b} q_t(s)\mathbb{E}^{\mathbb{Q}}\left[D(t,s)\pi(s,T_b)^+ \,\middle|\, \mathcal{F}(t)\right] ds && (7.11)\\
&= (1-R)\int_{T_a}^{T_b} \pi_{swaption}\Big(t; s, T_b, K, y_{s,b}(t)\Big) q_t(s) ds && (7.12)
\end{aligned}$$

In equation (7.10) we have used Fubini's rule to change the integration order, and in equation (7.11) we have used the assumed independence between τ and interest rates. To obtain equation (7.12), we have used that $\pi(s, T_b)$ can be characterized as the discounted cashflows in a forward starting swap, starting at time s with maturity T_b, where the fixed rate is fixed at time t. Hence, the time t risk-neutral expectation of the time t discounted value of $\pi(s, T_b)^+$ can be regarded as a *swaption* with *maturity* s and a strike which equals the forward swap rate $y_{a,b}(t)$.

□

To use the proposition in its current form is, of course, very difficult without making an approximation so that only a finite number of swaption prices has to be calculated. In the following we will derive an expression which alters the integral into a sum, given some basic assumptions.

Assume that default can only occur at the same time as swap payments, i.e. at time $T_{a+1}, \dots, T_b$. We then choose[3] to *postpone* the time default, so that τ is re-defined as

$$\tau \equiv \inf\{T_i : T_i \geq \tau, i \in \mathbb{Z}\}. \tag{7.13}$$

The proposition can then be approximated as

$$\begin{aligned} \mathrm{CVA}(t) &= (1-R) \sum_{i=a+1}^{b-1} \pi_{swaption}\Big(t; T_i, T_b, K, y_{i,b}(t)\Big)\mathbb{Q}\Big(\tau \in]T_{i-1}, T_i]\Big) & (7.14) \\ &= (1-R) \sum_{i=a+1}^{b-1} \pi_{swaption}\Big(t; T_i, T_b, K, y_{i,b}(t)\Big)\bigg(\mathbb{Q}\Big(\tau > T_{i-1} | \mathcal{F}(t)\Big) - \mathbb{Q}\Big(\tau > T_i | \mathcal{F}(t)\Big)\bigg). & (7.15) \end{aligned}$$

This is the CVA formula which we will use in order to price risky swaps. Basically it requires us to do two things:

1. Calculate $b - a - 1$ swaption prices with the relevant maturities and relevant swap rates.
2. Calculate the survival probabilities $\mathbb{Q}\Big(\tau > T_i | \mathcal{F}(t)\Big)$ for $i = a+1, \dots, b$.

It is clear from these statements that the process of determining the "true" (risky) swap price is rather complicated compared to calculating the risk free price. Not only do we have to do a significantly larger amount of calculations, but we also have to specify *at least* one model; the model for swaption pricing. However, in many cases one would also choose to specify a default model since not doing so restricts availability (and probably also the quality) of default probabilities. As we will argue in Chapter 9, the model-independent framework only allows for an inversion of a generic CDS pricing formula, so that each observed CDS quote can be inverted into a corresponding default probability. Assessing default probabilities in between maturities of observed CDS premia would then require some sort of interpolation scheme.

[3] This is a choice since we could just as well push τ in the opposite direction, i.e. we could *anticipate* τ. According to Brigo and Masetti (2005), there is no notable difference between the use of postponement or anticipation.

Chapter 8

Pricing Swaptions in an Affine Framework

8.1 Introduction

The pricing of swaptions depends on the modeling of the term structure of interest rates, since swaptions are options on swaps where the underlying swap rate is interest rate dependent. So far we have established swaption pricing formulas where the price of a swaption was expressed in terms of the swap rate under the associated swap measure. This was shown in Subsection 4.4.1. Based on this discovery, we were able to derive a semi-analytical pricing formula using Fourier inversion techniques in Subsection 5.3.1 where the price of swaptions were tied to the CCF of the swap rate dynamics. Now, in order to apply the semi-analytical pricing formula to price swaptions, we have to specify the term structure of interest rates. A well known and often used method of pricing swaptions is in the affine term structure framework, which we already covered in Chapter 6. Papers by Pelsser and Schrager (2006), Munk (1999), Dufresne and Goldstein (2002), and Singleton and Umantsev (2002b) all propose swaption pricing in Affine Term Structure Models (ATSM).

Singleton and Umantsev (2002b) propose a method in which they approximate the optimal exercise boundary for coupon-bearing securities with straight line segments that closely match the part of the boundary where the density of the affine state process in concentrated. By this approximation, the exercise probability of the swaption is reduced, and the exercise probabilities needed for computing the swaption price can then be calculated using Fourier inversion methods.

Dufresne and Goldstein (2002) use the fact that a swaption can be viewed as an option on a coupon bond. This is done since the moments of the coupon bond can be calculated through the joint moments of the individual ZCBs, which are now in closed form. The probability distribution of the future coupon bond prices is then estimated using a technique called Edgeworth expansion, which in turn provides an estimation of the swaption price.

Munk (1999) first generalizes the concept of stochastic duration to multi-factor diffusion models by defining the stochastic duration of a coupon bond as the time to maturity of the zero-coupon bond having the same relative volatility as the coupon bond. He then shows that the price of a European option on a coupon bond, which is equal to a swaption, is approximately proportional to the price of an option on a ZCB with maturity equal to the stochastic duration of the coupon bond.

In this thesis paper we have chosen to follow the approach proposed in Pelsser and Schrager (2006) where an approximate swaption price is also derived based on ATSM. The idea is to derive the dynamics of the swap rate and the underlying factors determining the short rate directly under the swap measure for a general ATSM. Under the swap measure the dynamics of the factors will have a *stochastic* drift term and the dynamics of the swap rate will have a *stochastic* volatility term. In order to remain in the affine setup, a approximate dynamics is proposed where the stochastic elements in the drift and volatility term are replaced by their time zero value. This approximation is justified by assuming that the stochastic elements are *low-variance martingales* (LVM) and therefore well approximated by their time zero values. This leaves us with a *time-dependent* drift for the factors and a *time-dependent* volatility for the swap rate, and we can again represent the dynamics in the affine framework. Now using known results from ATSM, we can derive a CCF from the approximate swap rate dynamics under the swap measure and then use Fourier inversion techniques to price swaptions according to the semi-analytical formulas in Proposition 5.1 on page 49. So, the cardinal idea is that the approximation allows us to write the swap rate dynamics under the swap measure in affine form, which implies that the CCF of the swap rate is known in close form. This means that the main task is to derive the dynamics and then show that the dynamics can in fact be in affine form. According to Pelsser and Schrager (2006), their method is comparable with the other three papers in terms of computational time and superior in accuracy compared to Munk (1999).

8.2 Dynamics under the Swap Measure

8.2.1 Deriving the Factor Dynamics

We begin by deriving the dynamics of the underlying factors under the swap measure. As mentioned, we do this in a general ATSM setup. We place ourselves under the risk neutral measure and consider an N factor ATSM where the instantaneous short rate $r(t)$ is modeled as an affine function of N unobservable factors $X(t) = (X_1(t), \dots, X_N(t))$. According to Section 6.1, this implies that the short rate follows the diffusion process

$$r(t) = \omega_0 + \sum_{i=1}^{N} \omega_i X_i(t) = \omega_0 + \omega_X^{\intercal} X(t), \tag{8.1}$$

where ω_0 is scalar, ω_X is an N-dimensional vector, and the N-dimensional vector of factors evolves according to the diffusion

$$dX(t) = \Psi(\Theta - X(t))dt + \Sigma\sqrt{V(t)}dW^{\mathbb{Q}}(t). \tag{8.2}$$

Here $W^{\mathbb{Q}}(t)$ is an N-dimensional Brownian motion, Ψ and Σ are $N \times N$ matrices, Θ is an N-dimensional vector, and the matrix $V(t)$ is defined as

$$V(t) = \operatorname{diag}(\alpha + \beta X(t)) = \begin{pmatrix} \alpha_1 + \beta_1^{\intercal} X(t) & 0 & \dots & 0 \\ 0 & \alpha_2 + \beta_2^{\intercal} X(t) & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & \alpha_N + \beta_N^{\intercal} X(t) \end{pmatrix} \tag{8.3}$$

where β is an $N \times N$ matrix and α is an N-dimensional vector. Again, we assume that $M(t) = \exp(\int_0^t r(s)ds)$ denotes the money market account and $D(t,T)$ denotes the the time t value of a ZCB maturing at time T. Hence, the price of a ZCB is given by $D(t,T) = \exp(A(t,T) - B(t,T)^\intercal X(t))$ where A and B are the solutions to two ordinary differential equations. Then, according to Subsection 6.2.3, the dynamics of the ZCB under the risk-neutral measure is given by

$$dD(t,T) = r(t)D(t,T)dt - B(t,T)^\intercal D(t,T)\Sigma\sqrt{V(t)}dW^{\mathbb{Q}}(t). \tag{8.4}$$

The goal is to derive the dynamics under the swap measure $\mathbb{Q}^{n+1,N}$. Therefore we need to derive the Girsanov Kernel φ for the transition from $\mathbb{Q}$ to $\mathbb{Q}^{n+1,N}$. In Subsection 4.4.1 we showed that the money market account relative to the annuity factor is a $\mathbb{Q}^{n+1,N}$-martingale and that in general all asset price processes relative to the annuity factor are $\mathbb{Q}^{n+1,N}$-martingales. This means that if we can derive the dynamics of $\frac{M(t)}{P_{n+1,N}(t)}$ under the $\mathbb{Q}$-measure, we can use the Girsanov Theorem along with the fact that $\frac{M(t)}{P_{n+1,N}(t)}$ is a $\mathbb{Q}^{n+1,N}$-martingale, to determine the Girsanov Kernel. Since

$$\frac{M(t)}{P_{n+1,N}(t)} = \frac{1}{\frac{P_{n+1,N}(t)}{M(t)}}, \tag{8.5}$$

by Ito's formula we have that

$$d\left(\frac{M(t)}{P_{n+1,N}(t)}\right) = -\frac{1}{\left(\frac{P_{n+1,N}(t)}{M(t)}\right)^2} d\left(\frac{P_{n+1,N}(t)}{M(t)}\right) + \frac{1}{\left(\frac{P_{n+1,N}(t)}{M(t)}\right)^3}\left(d\left(\frac{P_{n+1,N}(t)}{M(t)}\right)\right)^2. \tag{8.6}$$

In order to calculate (8.6), we begin by deriving the dynamics of $\frac{P_{n+1,N}(t)}{M(t)}$. Since $dP_{n+1,N}(t) = \sum_{i=n+1}^{N} \Delta dD(t,T_i)$ and $d\left(\frac{1}{M(t)}\right) = -r(t)\frac{1}{M(t)}dt$ and by an application of Ito's product rule we can show that[1]

$$d\left(\frac{P_{n+1,N}(t)}{M(t)}\right) = -\sum_{i=n+1}^{N}\left(\Delta B(t,T_i)^\intercal \frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}\right)\frac{P_{n+1,N}(t)}{M(t)}dW^{\mathbb{Q}}(t), \tag{8.7}$$

[1]Derivation can be found in Appendix B.

which also proves the already known fact that $\frac{P_{n+1,N}(t)}{M(t)}$ is a $\mathbb{Q}$-martingale. Now, plugging (8.7) into (8.6) yields[2]

$$
\begin{aligned}
d\left(\frac{M(t)}{P_{n+1,N}(t)}\right) &= \frac{1}{\left(\frac{P_{n+1,N}(t)}{M(t)}\right)^2}\sum_{i=n+1}^{N}\left(\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}\right)\frac{P_{n+1,N}(t)}{M(t)}dW^{\mathbb{Q}}(t) \\
&\quad + \frac{1}{\left(\frac{P_{n+1,N}(t)}{M(t)}\right)^3}\left(\sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}\right)^2\left(\frac{P_{n+1,N}(t)}{M(t)}\right)^2 dt \quad (8.8)\\
&= \frac{M(t)}{P_{n+1,N}(t)}\left(\sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}dW^{\mathbb{Q}}(t)\right.\\
&\quad \left.+\left(\sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}\right)^2 dt\right) \quad (8.9)\\
&= \frac{M(t)}{P_{n+1,N}(t)}\sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}\left(dW^{\mathbb{Q}}(t)\right.\\
&\quad \left.+\sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}dt\right) \quad (8.10)\\
&= \sum_{i=n+1}^{N}\left(\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}\right)\frac{M(t)}{P_{n+1,N}(t)}dW^{\mathbb{Q}^{n+1,N}}(t), \quad (8.11)
\end{aligned}
$$

where we in the last equality have used the Girsanov Theorem such that

$$
dW^{\mathbb{Q}^{n+1,N}}(t) = dW^{\mathbb{Q}}(t) + \sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}dt. \quad (8.12)
$$

Hence, the Girsanov Kernel φ for the transition from $\mathbb{Q}$ to $\mathbb{Q}^{n+1,N}$ is given by

$$
\varphi(t) = \sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}. \quad (8.13)
$$

This implies that the dynamics for the N unobservable factors under the $\mathbb{Q}^{n+1,N}$-measure are given by

$$
\begin{aligned}
dX(t) &= \Psi(\Theta - X(t))dt + \Sigma\sqrt{V(t)}\left(dW^{\mathbb{Q}^{n+1,N}}(t) - \sum_{i=n+1}^{N}\Delta B(t,T_i)^\intercal\frac{D(t,T_i)}{P_{n+1,N}(t)}\Sigma\sqrt{V(t)}dt\right) \quad (8.14)\\
&= \left(\Psi(\Theta - X(t)) - \Sigma V(t)\Sigma^\intercal\left(\sum_{i=n+1}^{N}\Delta B(t,T_i)\frac{D(t,T_i)}{P_{n+1,N}(t)}\right)\right)dt + \Sigma\sqrt{V(t)}dW^{\mathbb{Q}^{n+1,N}}(t) \quad (8.15)
\end{aligned}
$$

[2]The result of our derivation differs slightly compared to that in Pelsser and Schrager (2006), where they end up with a negative volatility term in (8.11).

Note that the presence of $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ for $i = n+1, \dots, N$ in (8.13) makes the Girsanov Kernel stochastic, since ZCBs prices in this setup are stochastic processes. This in turn makes the drift term in the factor dynamics stochastic. Hence, changing measures from $\mathbb{Q}$ to $\mathbb{Q}^{n+1,N}$ implies a stochastic drift term for the factors due to the randomness in the Girsanov Kernel.

8.2.2 Deriving the Swap Dynamics

We now turn to the task of deriving the swap rate dynamics under the swap measure. Given the affine specification of the short rate dynamics, the swap rate can be expressed as

$$
\begin{aligned}
y_{n,N}(t) &= \frac{D(t,T_n) - D(t,T_N)}{P_{n+1,N}(t)} &(8.16)\\
&= \frac{\exp(A(t,T_n) - B(t,T_n)^\intercal X(t)) - \exp(A(t,T_N) - B(t,T_N)^\intercal X(t))}{\sum_{i=n+1}^{N} \Delta \exp(A(t,T_i) - B(t,T_i)^\intercal X(t))} &(8.17)\\
&\equiv f(X(t),t) &(8.18)
\end{aligned}
$$

Since the swap rate is a $\mathbb{Q}^{n+1,N}$-martingale, its $\mathbb{Q}^{n+1,N}$-dynamics must be "drift-less" and contain no dt-terms. Thus, by Itô's formula

$$
\begin{aligned}
dy_{n,N}(t) &= \frac{\partial f(X(t),t)}{\partial X(t)} X(t) &(8.19)\\
&= \Bigg(\frac{\left(-B(t,T_n)^\intercal D(t,T_n) + B(t,T_N)^\intercal D(t,T_N)\right) P_{n+1,N}(t)}{P^2_{n+1,N}(t)} \\
&\quad - \frac{(D(t,T_n) - D(t,T_N)) \left(-\sum_{i=n+1}^{N} \Delta B(t,T_i)^\intercal D(t,T_i)\right)}{P^2_{n+1,N}(t)} \Bigg) X(t) &(8.20)\\
&= \Bigg(- B(t,T_n)^\intercal \frac{D(t,T_n)}{P_{n+1,N}(t)} + B(t,T_N)^\intercal \frac{D(t,T_N)}{P_{n+1,N}(t)} \\
&\quad + y_{n,N}(t) \sum_{i=n+1}^{N} \Delta B(t,T_i)^\intercal \frac{D(t,T_i)}{P_{n+1,N}(t)} \Bigg) \Sigma \sqrt{V(t)} dW^{\mathbb{Q}^{n+1,N}}(t). &(8.21)
\end{aligned}
$$

If we define $q_n^y(t) = -\frac{D(t,T_n)}{P_{n+1,N}(t)}$, $q_N^y(t) = (1 + \Delta y_{n,N}(t)) \frac{D(t,T_N)}{P_{n+1,N}(t)}$ and $q_i^y(t) = \Delta y_{n,N}(t) \frac{D(t,T_i)}{P_{n+1,N}(t)}$ for $i = n+1, \dots, N-1$, we can reduce (8.21) to

$$
dy_{n,N}(t) = \left(\sum_{i=n}^{N} q_i^y(t) B(t,T_i)^\intercal \right) \Sigma \sqrt{V(t)} dW^{\mathbb{Q}^{n+1,N}}(t) \tag{8.22}
$$

which confirms that the swap rate is a $\mathbb{Q}^{n+1,N}$-martingale, since it only consists of a volatility term. As with the factor dynamics, we see that the volatility in (8.22) also contains the stochastic term $\frac{D(t,T_i)}{P_{n+1,N}(t)}$.

8.2.3 Approximation of Dynamics

Next step is to approximate the dynamics in (8.15) and (8.22) in order to express the dynamics on affine form under the swap-measure. In Pelsser and Schrager (2006) an approximation is proposed where the term $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ is replaced by its time zero value $\frac{D(0,T_i)}{P_{n+1,N}(0)}$. This substitution makes the Girsanov Kernel deterministic and implies an affine time dependent drift term for the approximate factor dynamics and an affine time dependent volatility term for the approximate swap rate dynamics. According to Pelsser and Schrager (2006) this approximation is justified since the term $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ for $i = n+1, \ldots, N$ is considered to be a LVM. First, since asset prices normalised by the annuity factor are martingales under the swap measure, we know that $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ is a $\mathbb{Q}^{n+1,N}$-martingale. Second, Pelsser and Schrager (2006) claim that $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ has low variance. This means that the probability of $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ deviating from its mean value is "small". Given this low variance assumption, these martingales can be approximated by their expectation, specifically their time zero values. This approximation is the key element in the model proposed in Pelsser and Schrager (2006) and is inspired by the method in Brace et al. (2001). In the latter article it is shown that in the log-normal version of The Libor Market Model the swap rate can be approximated by log-normal martingales by substituting terms similar to $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ with their time zero values.

So, we will approximate the random term $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ by its conditional expected value under the swap measure, $\frac{D(0,T_i)}{P_{n+1,N}(0)}$. Hence, by substituting $\frac{D(0,T_i)}{P_{n+1,N}(0)}$ for $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ in (8.15) and $q_i(0)$ for $q_i(t)$ in (8.22) we obtain the following approximate $\mathbb{Q}^{n+1,N}$-dynamics,

$$dy_{n,N}(t) = \left(\sum_{i=n}^{N} q_i^y(0) B(t,T_i)^\intercal\right) \Sigma\sqrt{V(t)}dW^{\mathbb{Q}^{n+1,N}}(t), \tag{8.23}$$

$$dX(t) = \left(\Psi(\Theta - X(t)) - \Sigma V(t)\Sigma^\intercal \left(\sum_{i=n+1}^{N} \Delta B(t,T_i)\frac{D(0,T_i)}{P_{n+1,N}(0)}\right)\right) dt + \Sigma\sqrt{V(t)}dW^{\mathbb{Q}^{n+1,N}}(t). \tag{8.24}$$

Thus, given the LVM assumption, the approximations under the swap measure result in an affine model with time-dependent coefficients for the factor dynamics and the swap rate dynamics.

If we consider the swap rate as a pseudo factor along with the other N unobservable factors, we can express the swaption prices in (4.76) and (4.77) on page 40 as a linear combination of the factors minus the strike. Letting $\tilde{X}(t)$ denote the $N+1$-dimensional vector of factors, i.e. $\tilde{X}(t) = (y_{n,N}(t)\ X^\intercal(t))^\intercal$, and e_1, the first $N+1$-dimensional basis vector, i.e. $e_1 = (1\ 0 \ldots 0)^\intercal$, we can express the swaption prices as

$$\pi_{PS}(0) = P_{n+1,N}(0)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(e_1^\intercal \tilde{X}(t) - K)^+\right] = P_{n+1,N}(0)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(y_{n,N}(t) - K)^+\right], \tag{8.25}$$

$$\pi_{RS}(0) = P_{n+1,N}(0)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(K - e_1^\intercal \tilde{X}(t))^+\right] = P_{n+1,N}(0)\mathbb{E}^{\mathbb{Q}^{n+1,N}}\left[(K - y_{n,N}(t))^+\right]. \tag{8.26}$$

Hence, a swaption can be viewed as an option on the first factor of the affine model specified by (8.23) and (8.24) and we can therefore use results related to ATSM when pricing swaptions. Specifically, the result concerning a closed form solution for the CCF in an affine setup can now be applied. However, first we need to state the joint dynamics of (8.23) and (8.24), $\tilde{X}(t)$, on affine form.

8.3 The Approximative Dynamics on Affine Form

The goal is to write the dynamics of $\tilde{X}(t)$ on affine form. To do this, we define the following two N-dimensional vectors,

$$w(t) = \left(\sum_{i=n+1}^{N} \Delta B(t,T_i) \frac{D(0,T_i)}{P_{n+1,N}(0)} \right), \tag{8.27}$$

$$k(t) = \left(\sum_{i=n}^{N} q_i^y(0) B(t,T_i) \right). \tag{8.28}$$

We denote the N-dimensional identity matrix I_N[3]. Now, following Pelsser and Schrager (2006), the dynamics of $\tilde{X}(t)$ can be written on affine form as

$$\begin{aligned} d\tilde{X}(t) = & \left(\begin{bmatrix} 0 \\ \Psi\Theta - \Sigma \mathrm{diag}(\alpha)\Sigma^\intercal w(t) \end{bmatrix} + \begin{bmatrix} 0 \\ -\Psi X(t) - \Sigma \mathrm{diag}(\beta X(t))\Sigma^\intercal w(t) \end{bmatrix} \right) dt \\ & + \begin{bmatrix} k(t)^\intercal \\ I_N \end{bmatrix} \Sigma \sqrt{V(t)} dW^{\mathbb{Q}^{n+1,N}}(t) \qquad (8.29) \\ = & \tilde{\Psi}(t)(\tilde{\Theta}(t) - \tilde{X}(t))dt + \tilde{\Sigma}(t)\sqrt{\tilde{V}(t)} dW^{\mathbb{Q}^{n+1,N}}(t), \qquad (8.30) \end{aligned}$$

where $\begin{bmatrix} k(t)^\intercal \\ I_N \end{bmatrix}$ is $(N+1 \times N)$-matrix, and $\Sigma \mathrm{diag}(\alpha)\Sigma^\intercal w(t)$ and $\Sigma \mathrm{diag}(\beta X(t))\Sigma^\intercal w(t)$ are N-dimensional vectors.

Besides, from the fact that the dimensions of the parameters increase to $N+1$, it is not quite evident how $\tilde{\Psi}, \tilde{\Theta}, \tilde{\Sigma}$ and $\tilde{V}$ are defined when we go from (8.29) to (8.30)[4]. So before we can continue to develop the swaption pricing setup, we must first specify each new time-dependent parameter in (8.30). First, we start by determining the drift term in (8.29). From (8.29) and (8.30) we can conclude that

$$\tilde{\Psi}(t)\tilde{\Theta}(t) = \begin{bmatrix} 0 \\ \Psi\Theta - \Sigma \mathrm{diag}(\alpha)\Sigma^\intercal w(t) \end{bmatrix} \tag{8.31}$$

and

$$\tilde{\Psi}(t)\tilde{X}(t) = \begin{bmatrix} 0 \\ \Psi X(t) + \Sigma \mathrm{diag}(\beta X(t))\Sigma^\intercal w(t) \end{bmatrix}. \tag{8.32}$$

[3] The Identity matrix is defined as

$$I_N = \begin{pmatrix} e_1^\intercal \\ \vdots \\ e_N^\intercal \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

[4] This is not even specified in Pelsser and Schrager (2006).

We focus on (8.32) and specify a single entry in the $\Sigma \text{diag}(\beta X(t))\Sigma^{\intercal} w(t)$ vector in order to isolate $X(t)$ from the expression:

$$\begin{aligned}
\left(\Sigma \text{diag}(\beta X(t))\Sigma^{\intercal} w(t)\right)_i &= \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\sum_{l=1}^{N}\text{diag}(\beta X(t))_{kl}\Sigma^{\intercal}_{lj}w(t)_j && (8.33)\\
&= \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\text{diag}(\beta X(t))_{kk}\Sigma^{\intercal}_{kj}w(t)_j && (8.34)\\
&= \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\beta^{\intercal}_{k}X(t)\Sigma_{jk}w(t)_j && (8.35)\\
&= \left(\sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\Sigma_{jk}w(t)_j\beta^{\intercal}_{k}\right)X(t) && (8.36)\\
&\equiv \Gamma^{\intercal}_{i}(t)X(t) && (8.37)
\end{aligned}$$

where Γ_i is an N-dimensional vector. Hence, we define the matrix

$$\Gamma(t) = \begin{pmatrix} \Gamma^{\intercal}_1(t) \\ \vdots \\ \Gamma^{\intercal}_N(t) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{1k}\Sigma_{jk}w(t)_j\beta^{\intercal}_k \\ \vdots \\ \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{Nk}\Sigma_{jk}w(t)_j\beta^{\intercal}_k \end{pmatrix}. \tag{8.38}$$

Now, defining the entries in Ψ as $\Psi = (\Psi_1 \dots \Psi_N)^{\intercal}$ where $\Psi^{\intercal}_i = (\Psi_{i1} \dots \Psi_{iN})$, we can conclude that in order for (8.32) to hold, we must have that

$$\tilde{\Psi}(t) = \begin{pmatrix} 0 & 0 \\ 0 & \Gamma(t) + \Psi \end{pmatrix}, \tag{8.39}$$

where $\tilde{\Psi}(t)$ is an $(N+1 \times N+1)$ matrix.

Turning to (8.31), we take the same approach and begin by specifying a single entry in the $\Sigma \text{diag}(\alpha)\Sigma^{\intercal} w(t)$ vector:

$$\begin{aligned}
(\Sigma \text{diag}(\alpha)\Sigma^{\intercal} w(t))_i &= \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\sum_{l=1}^{N}\text{diag}(\alpha)_{kl}\Sigma^{\intercal}_{lj}w(t)_j && (8.40)\\
&= \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\text{diag}(\alpha)_{kk}\Sigma^{\intercal}_{kj}w(t)_j && (8.41)\\
&= \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{ik}\Sigma_{jk}w(t)_j\alpha_k && (8.42)\\
&\equiv \Omega_i(t). && (8.43)
\end{aligned}$$

By letting $\Omega_i(t)$ denote the i^{th} entry in the N-dimensional vector, such that

$$\Omega(t) = \begin{pmatrix} \Omega_1(t) \\ \vdots \\ \Omega_N(t) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{1k}\Sigma_{jk}w(t)_j\alpha_k \\ \vdots \\ \sum_{j=1}^{N}\sum_{k=1}^{N}\Sigma_{Nk}\Sigma_{jk}w(t)_j\alpha_k \end{pmatrix}, \tag{8.44}$$

we can establish (8.31) as a system of $N+1$ equations with $\tilde{\Theta}(t)$ as the unknown vector,

$$\tilde{\Psi}(t)\tilde{\Theta}(t) = \begin{pmatrix} 0 \\ \Psi\Theta - \Omega(t) \end{pmatrix} \Leftrightarrow \begin{pmatrix} 0 & 0 \\ 0 & \Gamma(t) + \Psi \end{pmatrix} \begin{pmatrix} \tilde{\Theta}_1(t) \\ \vdots \\ \tilde{\Theta}_{N+1}(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \Psi\Theta - \Omega(t) \end{pmatrix}. \tag{8.45}$$

From (8.45) it is obvious that $\tilde{\Theta}_1(t)$ becomes zero and therefore finding $\tilde{\Theta}_2(t)\dots\tilde{\Theta}_{N+1}(t)$ is reduced to solving N equations in the system

$$\begin{pmatrix} \Gamma_1^\intercal(t) + \Psi_1^\intercal \\ \vdots \\ \Gamma_N^\intercal(t) + \Psi_N^\intercal \end{pmatrix} \begin{pmatrix} \tilde{\Theta}_2(t) \\ \vdots \\ \tilde{\Theta}_{N+1}(t) \end{pmatrix} = \begin{pmatrix} \Psi_1^\intercal \\ \vdots \\ \Psi_N^\intercal \end{pmatrix} \begin{pmatrix} \Theta_1 \\ \vdots \\ \Theta_N \end{pmatrix} - \begin{pmatrix} \Omega_1(t) \\ \vdots \\ \Omega_N(t) \end{pmatrix}, \tag{8.46}$$

which can be solved for $\tilde{\Theta}_2(t)\dots\tilde{\Theta}_{N+1}(t)$ by inverting the matrix $\Gamma(t) + \Psi$, given that $\Gamma(t) + \Psi$ is nonsingular.

We now turn to specifying the volatility term in (8.30). Note that since the Brownian motion in (8.30) is $(N+1)$-dimensional, the matrixproduct $\tilde{\Sigma}(t)\tilde{V}(t)$ must be a matrix with $(N+1)$ columns. First, let us focus on the covariance matrix. From (8.29) we see that

$$\begin{pmatrix} k(t)^\intercal \\ I_N \end{pmatrix} \Sigma = \begin{pmatrix} k(t)^\intercal \\ e_1^\intercal \\ \vdots \\ e_N^\intercal \end{pmatrix} (\sigma_1 \dots \sigma_N) = \begin{pmatrix} k(t)^\intercal\sigma_1 & k(t)^\intercal\sigma_2 & \dots & k(t)^\intercal\sigma_N \\ e_1^\intercal\sigma_1 & e_1^\intercal\sigma_2 & \dots & e_1^\intercal\sigma_N \\ \vdots & \vdots & \ddots & \vdots \\ e_N^\intercal\sigma_1 & e_N^\intercal\sigma_2 & \dots & e_N^\intercal\sigma_N \end{pmatrix} \tag{8.47}$$

leaving us with an $(N+1 \times N)$ matrix. We expand the matrix in (8.47) by a column of zeroes and define the $(N+1 \times N+1)$ covariance matrix in (8.30) as

$$\tilde{\Sigma}(t) = \begin{pmatrix} 0 & k(t)^\intercal\sigma_1 & k(t)^\intercal\sigma_2 & \dots & k(t)^\intercal\sigma_N \\ 0 & e_1^\intercal\sigma_1 & e_1^\intercal\sigma_2 & \dots & e_1^\intercal\sigma_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & e_N^\intercal\sigma_1 & e_N^\intercal\sigma_2 & \dots & e_N^\intercal\sigma_N \end{pmatrix} \tag{8.48}$$

Since the square root operator is defined for squared matrices only, we define $\tilde{V}(t)$ as

$$\tilde{V}(t) = \mathrm{diag}(\tilde{\alpha} + \tilde{\beta}\tilde{X}(t)) \tag{8.49}$$

where

$$\tilde{\beta} = \begin{pmatrix} \tilde{\beta}_1^\intercal \\ \vdots \\ \tilde{\beta}_{N+1}^\intercal \end{pmatrix} \text{ and } \tilde{\alpha} = \begin{pmatrix} \tilde{\alpha}_1 \\ \vdots \\ \tilde{\alpha}_{N+1} \end{pmatrix} \tag{8.50}$$

so that $\tilde{\beta}_i^\intercal = \left(0 \; \beta_{(i-1)1} \ldots \beta_{(i-1)N}\right)$ and $\tilde{\alpha}_i = \alpha_{i-1}$ for $i = 2, \ldots, N+1$. Hence, the volatility term in (8.30) arises from straightforward calculations

$$\tilde{\Sigma}(t)\sqrt{\tilde{V}(t)} = \begin{pmatrix} 0 & k(t)^\intercal\sigma_1 & k(t)^\intercal\sigma_2 & \ldots & k(t)^\intercal\sigma_N \\ 0 & e_1^\intercal\sigma_1 & e_1^\intercal\sigma_2 & \ldots & e_1^\intercal\sigma_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & e_N^\intercal\sigma_1 & e_N^\intercal\sigma_2 & \ldots & e_N^\intercal\sigma_N \end{pmatrix} \begin{pmatrix} \sqrt{\tilde{\alpha}_1 + \tilde{\beta}_1^\intercal\tilde{X}(t)} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sqrt{\tilde{\alpha}_{N+1} + \tilde{\beta}_{N+1}^\intercal\tilde{X}(t)} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & k(t)^\intercal\sigma_1\sqrt{\tilde{\alpha}_2 + \tilde{\beta}_2^\intercal\tilde{X}(t)} & \ldots & k(t)^\intercal\sigma_N\sqrt{\tilde{\alpha}_{N+1} + \tilde{\beta}_{N+1}^\intercal\tilde{X}(t)} \\ 0 & e_1^\intercal\sigma_1\sqrt{\tilde{\alpha}_2 + \tilde{\beta}_2^\intercal\tilde{X}(t)} & \ldots & e_1^\intercal\sigma_N\sqrt{\tilde{\alpha}_{N+1} + \tilde{\beta}_{N+1}^\intercal\tilde{X}(t)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & e_N^\intercal\sigma_1\sqrt{\tilde{\alpha}_2 + \tilde{\beta}_2^\intercal\tilde{X}(t)} & \ldots & e_N^\intercal\sigma_N\sqrt{\tilde{\alpha}_{N+1} + \tilde{\beta}_{N+1}^\intercal\tilde{X}(t)} \end{pmatrix}. \tag{8.51}$$

So, through tedious and rather technical calculations we have specified the parameters in (8.30) and therefore justified that the joint approximative dynamics of the swap rate and the factors under the swap measure can be stated in affine form. We can now continue developing the swaption pricing setup.

8.4 The Swaption Pricing Formula

We have now specified how the approximate factor and swap rate dynamics can be expressed on affine form under the swap measure. We have also shown that given this affine specification, we can now express swaptions as options on the first factor of the affine model, in this case the swap rate, which is implied by the the pricing formulas in (8.25) and (8.26). However, it is not immediately obvious how the approximate dynamics of the swap rate is distributed and therefore using (8.25) and (8.26) to derive an analytical formula can turn out to be very complicated, even impossible.

The goal is therefore to derive a semi-analytical formula, as proposed in Pelsser and Schrager (2006), using the Fourier inversion setup we presented in Subsection 5.3.1. This implies that we need to express an CCF for the approximate swap rate. This is where the affine specification becomes especially useful, since it allows us to use the results for CCFs stated in Proposition 6.2 on page 55. Here we saw that in ATSM the CCF is given on exponential affine form. Thus, we can express the CCF for the affine dynamics of $\tilde{X}$ as

$$\phi_{\tilde{X}}(v, t, T_n) = \exp(\gamma(t) + \delta(t) \cdot \tilde{X}(t)) \tag{8.52}$$

where $\gamma(t)$ and $\delta(t)$ are solutions to the following set of complex ODEs

$$\frac{d\delta(t)}{dt} = \tilde{\Psi}^\intercal(t)\delta(t) - \frac{1}{2}\sum_{i=1}^{N+1}\left[\tilde{\Sigma}^\intercal(t)\delta(t)\right]_i^2 \tilde{\beta}_i \tag{8.53}$$

$$\frac{d\gamma(t)}{dt} = -\tilde{\Theta}^\intercal(t)\tilde{\Psi}(t)\delta(t) - \frac{1}{2}\sum_{i=1}^{N+1}\left[\tilde{\Sigma}^\intercal(t)\delta(t)\right]_i^2 \tilde{\alpha}_i \tag{8.54}$$

with the terminal conditions

$$\delta(T_n) = ive_1 \tag{8.55}$$
$$\gamma(T_n) = 0. \tag{8.56}$$

The presence of the first basis vector e_1 ensures that at time T_n we end up with the CCF for just the swap rate. Notice that δ and γ are of dimensions $N+1$ and scalar, respectively. The complex ODEs in (8.53) and (8.54), along with the terminal conditions (8.55) and (8.56), will have to be solved numerically in most cases. This also applies for the one-factor CIR model which will be our focal point in our application of the model.

Now given the CCF in (8.52), we have everything needed in order to price both payer and receiver swaptions according to Proposition 5.1 on page 49. Hence, the time t value of a payer swaption is given as

$$\pi_{PS}(t,K) = \frac{\exp(-\alpha K)}{\pi} \int_0^\infty \Re\Big[\exp(-iuK)\psi_P(u,t,T_n)\Big] du, \tag{8.57}$$
$$\psi_P(u,t,T_n) = P_{n+1,N}(t) \frac{\phi_{\tilde{X}}(u - i\alpha, t, T_n)}{(iu+\alpha)^2}, \tag{8.58}$$

along with the corresponding receiver swaption

$$\pi_{RS}(t,K) = \frac{\exp(\alpha K)}{\pi} \int_0^\infty \Re\Big[\exp(-iuK)\psi_R(u,t,T_n)\Big] du, \tag{8.59}$$
$$\psi_R(u,t,T_n) = P_{n+1,N}(t) \frac{\phi_{\tilde{X}}(u + i\alpha, t, T_n)}{(iu-\alpha)^2} \tag{8.60}$$

where $\phi_{\tilde{X}}$ is given by (8.52). These formulas will constitute the foundation in our application of the swaption pricing model.

Chapter 9

Deriving Default Probabilities: Intensity Models

9.1 Concepts of Deriving Default Probabilities

We have already seen that we need the default probabilities for a given firm in order to price a risky swap with the firm in question as counterparty. Our goal in this chapter is to answer the following question: "What is the probability of default for company x in a given time frame?". A qualified answer to this question will typically stem from a framework that uses one of the following approaches/models:

- A so-called *structural model* that views debt, equity and/or other claims issued by a firm as contingent claims on the firm's asset value. The perhaps most popular structural model is Merton's (1974), which basically (in analogy to the classic Black-Scholes model) models the asset value of the firm as a geometric Brownian Motion. Default is then defined as the first time the process hits zero.
- Historical default frequencies obtained by rating agencies, e.g. Moody's (2008). Moody's frequently publish default probabilities for different entities and time horizons.
- *Bootstrapping* from observable market data, typically Credit Default Swaps (CDS's), to obtain *implied* default probabilities.
- An intensity model which models the probability of default through some intensity process (much more on this approach below).

The last two approaches can sometimes be merged in the sense that intensity models can be calibrated using e.g. CDS data. Both approaches rely on the assumption that default probabilities are known by the market. That is to say that default probabilities are reflected in the market price of firm related traded assets. Since e.g. a corporate bond is cheaper than a government bond, it makes sense that such an asset contains a default component. In fact Longstaff et al. (2004) show that the *spread* between such two bonds *primarily* consists of credit risk[1]. While we will derive building blocks for the third and fourth mentioned approach, we will only implement the last one–the intensity modeling. The reason for this is that while bootstrapping default probabilities is straightforward, the method is restricted to stipulate default probabilities in the time frame of the given data. Furthermore, in order to justify the results of the implied approach, the construction of intensity based models is required.

[1] They also show that the second largest component consists of liquidity risk.

9.2 The Intensity Models

Our presentation and derivation of intensity models follow Lando (1998) and Lando (2004). We will begin by defining some basic probabilistic terms and functions, and then focus on the basic driver of the intensity models; the Cox process. Strictly speaking, intensity models are a framework which, in the most general setting, is not necessarily tied to Cox processes. The loss of generality is however very small, since Cox processes–as we shall see–describe a broader class of processes. After the introduction to the framework, we will present some central results that will afterwards be used to price a risky ZCB and a Credit Default Swap (CDS). The pricing of the latter is particularly important, since it allows for an *inversion* of the CDS pricing formula so that default probabilities may be found implicitly. The chapter will be completed by describing how an intensity setup works under affine assumptions. Developing the framework further, we will describe how a one factor CIR model can be used as an intensity driver.

According to our definition of an intensity model, default is defined as the first jump of some stochastic jump process[2]. So when in the following section we describe the probability of the process being zero, it is equivalent to the *survival* of the firm. Likewise, the probability of the process being greater than zero is linked to the event of a default.

9.2.1 Survival Functions and Hazard Rates

Let us define *time of default* as a positive continuous random variable τ. The properties of the variable are given by the distribution function F and the density function f under the measure $\mathbb{Q}$. The survival function S is thus given by

$$S(t) = \mathbb{Q}(\tau > t) = 1 - F(t) = 1 - \int_0^t f(s)ds.$$

The *hazard rate* can then be defined as

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} = -\frac{d \log S(t)}{dt}. \tag{9.1}$$

Using the first and last part of this expression, the survival function can be derived in terms of the hazard rate:

$$\begin{aligned} h(t) &= -\frac{d \log S(t)}{dt} \\ \Leftrightarrow -h(t)dt &= d \log S(t) \\ \Leftrightarrow \log S(t) - logS(0) &= -\int_0^t h(t)dt \\ \Leftrightarrow S(t) &= \exp\left(-\int_0^t h(t)dt\right) \end{aligned} \tag{9.2}$$

Expressions (9.1) and (9.2) do not provide much insight into the interpretation of the hazard rate. However, using the definition given by equation (9.1) we can show how the hazard rate is linked

[2] An intensity model could also be used so that a jump defines a shift *in rating class*. We will however not incorporate ratings in this thesis paper.

to conditional probabilities:

$$
\begin{aligned}
P(\tau \leq t + \Delta t | \tau > t) &= \frac{P(\tau \leq t + \Delta t, \tau > t)}{P(\tau > t)} \\
&= \frac{P(\tau \leq t + \Delta t) - P(\tau < t)}{P(\tau > t)} \\
&= 1 - \frac{P(\tau > t + \Delta t)}{P(\tau > t)} \\
&= 1 - \frac{\exp\left(-\int_0^{t+\Delta t} h(s)ds\right)}{\exp\left(-\int_0^t h(s)ds\right)} \\
&= 1 - \exp\left(-\int_0^{t+\Delta t} h(s)ds + \int_0^t h(s)ds\right) \\
&= 1 - \exp\left(-\int_t^{t+\Delta t} h(s)ds\right)
\end{aligned}
\tag{9.3}
$$

Furthermore, it can be shown that

$$
h(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(\tau \leq t + \Delta t | \tau > t) \tag{9.4}
$$

which tells us that $h(t)\Delta t$ is *approximately* the conditional probability of a default in a small interval after t, given survival up to (and including) time t. It is thus stated that the hazard rate is a *deterministic* function that only depends on the passing of time.

9.2.2 Inhomogeneous Poisson Processes and Cox-Processes

An inhomogeneous Poisson process N is a Poisson process with a deterministic intensity function $l(u)$, so that the increments are described as

$$
P(N(t) - N(s) = k) = \frac{\left(\int_s^t l(u)du\right)^k}{k!}, \; k = 0, 1, \ldots. \tag{9.5}
$$

Letting the process start in $N_0 = 0$, the probability of no arrivals (equivalent to a survival function in the given framework) is

$$
P(N(t) = 0) = \exp\left(-\int_0^t l(u)du\right). \tag{9.6}
$$

Note the strong and vital comparison to equation (9.2): in the current framework the deterministic intensity $l(u)$ *is* the hazard rate.

A Cox process is a generalization of the inhomogeneous Poisson process where the intensity is a *stochastic process* itself. It is important to realize that if we condition on a particular realization $l(\cdot, \omega)$, the Cox process becomes an inhomogeneous Poisson process so that $l(\cdot, \omega)$ is the effective hazard rate. Thus, by letting $\lambda\Big(X(s)\Big)$ denote the stochastic intensity at time s, we have that

$$
l(s, \omega) = \lambda\Big(X(s)\Big) \tag{9.7}
$$

where X is an $\mathbb{R}^d$-valued stochastic process and $\lambda : \mathbb{R}^d \to [0, \infty[$ is a non-negative right-continuous function. If additionally we assume that λ is integrable, i.e.

$$\Lambda(t) \equiv \int_0^t \lambda(s)ds < \infty \text{ for } t \in [0,\mathrm{T}] \tag{9.8}$$

then we can formally define a Cox Process as

$$\tilde{N}(t) \equiv N\Big(\Lambda(t)\Big). \tag{9.9}$$

The intuition of the *state-variable* X is that it contains all information regarding the riskiness of the firm and thus the firm's probability of default. This could e.g. be credit ratings, market value, stock prices, interest rates, etc.

We now introduce a unit exponential random variable E_1 independent of X. The time of default can then be defined as

$$\tau = \inf\left\{t : \int_0^t \lambda\Big(X(s)\Big)ds \geq E_1\right\}. \tag{9.10}$$

To be precise about X and E_1, and to be precise in the remainder of this section, we will fix a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ that supports E_1 and $X = \{X(t) : 0 \leq t \leq T\}$ so that X is right-continuous with left limits. The variable $\mathcal{F}(t)$ containing all the information can then be divided into two sub σ-algebras; $\mathcal{G}(t)$ containing information about X, and $\mathcal{H}(t)$ containing information regarding the occurrence of default. Mathematically speaking, the informational setup may be stated as follows:

$$\begin{aligned}
\mathcal{G}(t) &= \sigma\{X(s) : 0 \leq s \leq t\} \\
\mathcal{H}(t) &= \sigma\{1_{\{\tau \leq s\}} : 0 \leq s \leq t\} \\
\mathcal{F}(t) &= \mathcal{G}(t) \vee \mathcal{H}(t)
\end{aligned} \tag{9.11}$$

Since the intensity is a function of X, and $\mathcal{G}(t)$ stores all the information regarding X, we can say that the hazard rate is adapted to $\mathcal{G}(t)$. If we further assume a risk free spot rate process $r(t)$, with a corresponding money market account

$$M(t) = \exp\left(\int_0^t r(s)ds\right), \tag{9.12}$$

then the price of a risk free ZCB can be expressed as

$$D(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\frac{M(t)}{M(T)}\,\middle|\,\mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{M(t)}{M(T)}\,\middle|\,\mathcal{G}(t)\right]. \tag{9.13}$$

The last equation in the above follows the fact that all market variables (and in particular the risk free spot rate) are stored in X.

9.2.3 Three Pivotal Results

Before we turn our attention to the price implications of choosing an intensity setup, we will derive three technical results that are central in order to accomplish any true pricing formulas. The two first Lemmas are of general importance whereas the third is used solely as a building block in CDS pricing. Note that in the original paper by Lando (1998) a couple of other building blocks are proposed. In our case however, these are negligible since they are used to price other types of claims than risky ZCBs and CDS's[3].

The following results all assume a probability space and a Cox process consistent with the above framework.

Lemma 9.1.

$$\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}|\mathcal{G}(T)\right] = \exp\left(-\int_0^T \lambda\Big(X(s)\Big)ds\right)$$

Proof.

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}|\mathcal{G}(T)\right] &= \mathbb{Q}\Big(\tau > T|\mathcal{G}(T)\Big) \\
&= \mathbb{Q}\left(\int_0^T \lambda\Big(X(s)\Big)ds < E_1 \,\middle|\, \mathcal{G}(T)\right) && (9.14) \\
&= 1 - \mathbb{Q}\left(\int_0^T \lambda\Big(X(s)\Big)ds \geq E_1 \,\middle|\, \mathcal{G}(T)\right) \\
&= 1 - \left(1 - \exp\left(-\int_0^T \lambda\Big(X(s)\Big)ds\right)\right) && (9.15) \\
&= \exp\left(-\int_0^T \lambda\Big(X(s)\Big)ds\right)
\end{aligned}$$

In equation (9.14) we have used equation (9.10) to realise that survival is the probability of λ *not* reaching E_1. In equation (9.15) we have used that $\int_0^T \lambda\Big(X(s)\Big)ds$ is known, given $\mathcal{G}(T)$, so that the probability equals a simple distribution function of E_1 in the point $\int_0^T \lambda\Big(X(s)\Big)ds$. □

Lemma 9.2.

$$\mathbb{Q}(\tau > T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T \lambda\Big(X(s)\Big)ds\right)\right]$$

[3] In particular, Lando also provides results that enable pricing of a continuous stream of payments until default.

Proof.

$$\begin{aligned}\mathbb{Q}(\tau > T) &= \mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\Big[\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}|\mathcal{G}(T)\right]\Big] &(9.16)\\ &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T \lambda\Big(X(s)\Big)ds\right)\right] &(9.17)\end{aligned}$$

Getting to equation (9.16), we have used the so called *law of iterative expectations*[4], which states that if $t < t^{'} \leq T$, it holds that

$$\mathbb{E}\left[X(T)|\mathcal{F}(t)\right] = \mathbb{E}\left[\mathbb{E}\left[X(T)|\mathcal{F}(t^{'})\right]|\mathcal{F}(t)\right], \tag{9.18}$$

given that X is a $\mathcal{F}(t)$-measurable integrable random variable. The proof is fairly straightforward and will not be included in this thesis paper. Finally, equation (9.17) holds by using Lemma 9.1. □

Note how the two above Lemmas both state the survival probability. The difference simply is that Lemma 9.1 conditions the probability of the information at time T regarding X, while Lemma 9.2 is purely unconditional.

The next result is different since it is more directly linked to pricing methods. The idea is to price a cash flow $Z(\tau)$ which is payed *only* at default time, given that default happens before some final maturity T. The link to CDS pricing should then be obvious since a CDS exactly pays the investor some amount in case of a (relevant) default before maturity. Z is generally a stochastic process adapted to X, but for the CDS pricing Z will simply be constant.

Lemma 9.3.

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^{\tau} r(s)ds\right)Z(\tau)\,\middle|\,\mathcal{F}(t)\right]\\ &\quad= 1_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left[\int_t^T Z(s)\lambda(s)\exp\left(-\int_t^s\Big(r(u)+\lambda(u)\Big)du\right)ds\,\middle|\,\mathcal{G}(t)\right]\end{aligned}$$

Proof. From Lemma 9.1 we know that the conditional distribution function for τ is given as

$$\mathbb{Q}\Big(\tau \leq s|\tau > t, \mathcal{G}(t)\Big) = 1 - \exp\left(-\int_t^s \lambda(u)du\right)$$

for $s > t$. Differentiating, one obtains the conditional density function:

$$\begin{aligned}\frac{\partial}{\partial s}\mathbb{Q}\Big(\tau \leq s|\tau > t, \mathcal{G}(t)\Big) &= -\frac{\partial}{\partial s}\exp\left(-\int_t^s \lambda(u)du\right)\\ &= \lambda(s)\exp\left(-\int_t^s \lambda(u)du\right) &(9.19)\end{aligned}$$

[4]The law is also known as the *tower rule*, the *law of total expectation*, or the *smoothing theorem*.

With the density given by expression (9.19) in mind, we can prove the result:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^{\tau} r(s)ds\right) Z(\tau) \,\middle|\, \mathcal{F}(t)\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^{\tau} r(s)ds\right) Z(\tau) \,\middle|\, \mathcal{G}(T)\vee\mathcal{H}(t)\right] \,\middle|\, \mathcal{F}(t)\right] && (9.20) \\
&= 1_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left[\int_t^T \underbrace{\exp\left(-\int_t^s r(u)du\right) Z(s)}_{\text{transformation of } \tau} \underbrace{\lambda(s)\exp\left(-\int_t^s \lambda(u)du\right) ds}_{\text{density of } \tau} \,\middle|\, \mathcal{F}(t)\right] && (9.21) \\
&= 1_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left[\int_t^T Z(s)\lambda(s)\exp\left(-\int_t^s \Big(r(u)+\lambda(u)\Big)du\right) ds \,\middle|\, \mathcal{F}(t)\right] \\
&= 1_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left[\int_t^T Z(s)\lambda(s)\exp\left(-\int_t^s \Big(r(u)+\lambda(u)\Big)du\right) ds \,\middle|\, \mathcal{G}(t)\right] && (9.22)
\end{aligned}$$

In equation (9.20) the law of iterative expectations is used. The inner expectation in equation (9.20) is conditional by $\mathcal{G}(T)$ which results in $r(s)$, $Z(\tau)$ and the density being known. Equation (9.21) then arises by calculating the inner expectation as a simple transformation of τ and by "pulling" $\mathcal{H}(t)$ out of the expectation as an indicator[5]. The calculation of the expectation is (as usual) done by integrating over the product of the transformation and the density function. The last equation (9.22) holds because everything inside the expectation is driven solely by the state variable X and is thus independent of $\mathcal{H}(t)$. □

Note how Lemma 9.3 eliminates the stochastic time of default τ so that the limits in the integral become deterministic. Note also how the expectation can be interpreted as the *risk weighted* expected cash flow discounted by a *risk adjusted* interest rate, since the risk source (intensity) λ is used *both* in and out of the discounting factor.

9.3 Pricing

We will now show how the technical results can be used to obtain true pricing formulas for ZCBs and CDS's. Note how the relationship between λ and X deliberately has not been specified explicitly. This will also be the case in this section, which improves the strength of the pricing formulas.

9.3.1 Pricing of Zero Coupon Bonds

A good example of pricing in the intensity framework can be shown by pricing a ZCB since this is the simplest contract possible. We can thus turn our attention to the implications of the chosen framework.

[5]See Lando (2004) p. 114-115.

Proposition 9.1. *Assume that* no *default has occurred before time t and a short rate process $r\Big(X(t)\Big)$ such that the price of a ZCB is prescribed as*

$$D(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r\Big(X(s)\Big)ds\right)\middle|\,\mathcal{G}(t)\right].$$

A ZCB with maturity T, zero recovery, promised *payment of one, issued by a credit risky firm which under the $\mathbb{Q}$-measure has an intensity $\lambda\Big(X(s)\Big)$ according to the framework described in Subsection 9.2.2 then has the time t price*

$$\tilde{D}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \left(r\Big(X(s)\Big)+\lambda\Big(X(s)\Big)\right)ds\right)\middle|\,\mathcal{F}(t)\right].$$

Proof.

$$\begin{aligned}
\tilde{D}(t,T) &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r\Big(X(s)\Big)ds\right)1_{\{\tau>T\}}\middle|\,\mathcal{F}(t)\right]\\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r\Big(X(s)\Big)ds\right)1_{\{\tau>T\}}\middle|\,\mathcal{G}(T)\right]\middle|\,\mathcal{F}(t)\right]\\
&= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r\Big(X(s)\Big)ds\right)\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau>T\}}\middle|\,\mathcal{G}(T)\right]\middle|\,\mathcal{F}(t)\right] &(9.23)\\
&= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r\Big(X(s)\Big)ds\right)\exp\left(-\int_t^T \lambda\Big(X(s)\Big)ds\right)\middle|\,\mathcal{F}(t)\right] &(9.24)\\
&= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \left(r\Big(X(s)\Big)+\lambda\Big(X(s)\Big)\right)ds\right)\middle|\,\mathcal{F}(t)\right]
\end{aligned}$$

Equation (9.23) is valid since $r(X(t))$ is adapted to the filtration $\mathcal{G}(T)$ and thus deterministic. In equation (9.24) Lemma 9.1 has been used directly. □

The proposition tells us that the price of a risky ZCB can be calculated simply by changing the discounting rate $r\Big(X(s)\Big)$ to the *intensity adjusted* discounting rate $r\Big(X(s)\Big)+\lambda\Big(X(s)\Big)$. By doing this we get rid of the indicator function which makes the calculation much easier.

The result can be modified to fit the pricing of a generic claim, which has some *promised* payment $f(X(T))$ at maturity and an *actual* payment $f(X(T))1_{\{\tau>T\}}$. To prove this, the same calculation steps can be used starting with

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r\Big(X(s)\Big)ds\right)f\Big(X(T)\Big)1_{\{\tau>T\}}\middle|\,\mathcal{F}(t)\right], \tag{9.25}$$

and ending up with

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T\left(r\Big(X(s)\Big)+\lambda\Big(X(s)\Big)\right)ds\right)f(X(T)\,\Big|\,\mathcal{F}(t)\right]. \tag{9.26}$$

9.3.2 Pricing of a Credit Default Swap

In order to price a CDS we follow an approach that lies somewhere between Mortensen (2006), Lando (2004), and Felthütter (2008). As mentioned in subsection 3.4.1 on page 19, the standard market practice in the CDS market (and in other swap markets) is to find the premium that gives the swap zero value upon inception. The premium is computed using information on the default intensity (assuming an underlying Cox process), the recovery in default, as well as some assumptions in order to achieve more tangible results.

As usual, we will set the notional amount of the CDS to one so that the results may be scaled as required. We will price at time 0 to simplify things. Let us denote *the value* of the premium leg payed by the *protection buyer* by π^{prem} and *the value* of the protection leg paid (in case of default) by the *protection seller* by π^{prot}. The premium payments $c^{ds}(T)$ are paid in arrears at a frequency f until maturity T or default τ, whichever comes first. The frequency f is defined so that there are $c^{ds}(T)/f$ payments each year. If a default happens, an accrual premium payment is made. This size is naturally calculated using the timespan between the last premium payment date and the time of default. Since we price at time 0, the value of the premium leg is given by

$$\pi^{prem}=\mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{Tf}\exp\left(-\int_0^{t_j}r(s)ds\right)1_{\{\tau>t_j\}}\frac{c^{ds}(T)}{f}+\exp\left(-\int_0^{\tau}r(s)ds\right)1_{\{t_{j-1}<\tau\leq t_j\}}c^{ds}(T)(\tau-t_{j-1})\right]$$

where $t_j=j/f$ for $j=1,\dots,fT$.

Likewise, the value of the protection leg is

$$\pi^{prot}=\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^{\tau}r(s)ds\right)1_{\{\tau\leq T\}}(1-R)\right] \tag{9.27}$$

where R as usual denotes the recovery fraction.

The two expressions above are the most general cases where no assumptions are made. Deriving the fair premium from these equations will however give results that are not directly computational. We thus proceed by making our first assumption, that premium payments are calculated without accrual. Using Proposition 9.1 we can then price the premium leg as a sum of risky ZCBs:

$$\pi^{prem}=\mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{Tf}\exp\left(-\int_0^{t_j}r(s)ds\right)1_{\{\tau>t_j\}}\frac{c^{ds}(T)}{f}\right] \tag{9.28}$$

$$=\frac{c^{ds}(T)}{f}\mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{Tf}\exp\left(-\int_0^{t_j}r(s)ds\right)1_{\{\tau>t_j\}}\right] \tag{9.29}$$

$$=\frac{c^{ds}(T)}{f}\sum_{j=1}^{Tf}\tilde{D}(0,t_j) \tag{9.30}$$

Assuming that a Cox process is driving the default probability, the value of the protection leg given by equation (9.27) can be developed further using Lemma 9.3:

$$
\begin{aligned}
\pi^{prot} &= (1-R)\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^{\tau} r(s)ds\right) 1_{\{\tau \le T\}}\right] \\
&= (1-R)\mathbb{E}^{\mathbb{Q}}\left[\int_0^T \lambda(t)\exp\left(-\int_0^t \Big(r(s)+\lambda(s)\Big)ds\right)dt\right] \\
&= (1-R)\int_0^T \mathbb{E}^{\mathbb{Q}}\left[\lambda(t)\exp\left(-\int_0^t \Big(r(s)+\lambda(s)\Big)ds\right)\right]dt \qquad (9.31)
\end{aligned}
$$

In equation (9.31) we have used Fubini's theorem to change the integral order.

To achieve further simplification we make our second fundamental assumption, that there is independence between the default intensity and the short rate process under the martingale measure $\mathbb{Q}$. The value of the protection leg then becomes

$$
\begin{aligned}
\pi^{prot} &= (1-R)\int_0^T \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^t r(s)ds\right)\lambda(t)\exp\left(-\int_0^t \lambda(s)ds\right)\right]dt \\
&= (1-R)\int_0^T \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^t r(s)ds\right)\right]\mathbb{E}^{\mathbb{Q}}\left[\lambda(t)\exp\left(-\int_0^t \lambda(s)ds\right)\right]dt \\
&= (1-R)\int_0^T \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^t r(s)ds\right)\right]\mathbb{E}^{\mathbb{Q}}\left[-\frac{\partial}{\partial t}\exp\left(-\int_0^t \lambda(s)ds\right)\right]dt \qquad (9.32) \\
&= (1-R)\int_0^T D(0,t)\left(-\frac{\partial}{\partial t}S(t)\right)dt \qquad (9.33) \\
&= (1-R)\int_0^T D(0,t)h(t)S(t)dt \qquad (9.34)
\end{aligned}
$$

where $h(t)$ as usual denotes the expectation to the stochastic intensity, i.e. the hazard rate. In equation (9.32) we have used equation (9.19), and in equation (9.33) we have used Lemma 9.2. In the last equation (9.34) we have used the representation of the survival probability given by equation (9.2) on page 75.

We can now equate the values π^{prem} and π^{prot} in order to obtain the fair premium,

$$
\begin{aligned}
c^{ds}(T) &= f\frac{(1-R)\int_0^T D(0,t)h(t)S(t)dt}{\sum_{j=1}^{Tf}\tilde{D}(0,t_j)} \\
&= f\frac{(1-R)\int_0^T D(0,t)h(t)S(t)dt}{\sum_{j=1}^{Tf} D(0,t_j)S(t_j)} \qquad (9.35) \\
&= f\frac{(1-R)\int_0^T D(0,t)f(t)dt}{\sum_{j=1}^{Tf} D(0,t_j)S(t_j)} \qquad (9.36)
\end{aligned}
$$

where the denominator in equation (9.35) stems from applying independence to the price of a risky ZCB given by Proposition 9.1. The density $f(t)$ in equation (9.36) stems from the basic definition of the hazard rate given by equation (9.1) on page 75. In order to provide the most tangible result possible, the third and final assumption can be made; we calculate as if settlements only take place simultaneously as premium payments. A default taking place between time t_{j-1}

and t_j will then be "pushed" forward to time t_j, i.e.

$$\widehat{\mathbb{Q}}(\tau = t_j) \equiv \mathbb{Q}\Big(\tau \in (t_{j-1}, t_j]\Big) = S(t_{j-1}) - S(t_j). \tag{9.37}$$

Note that this assumption includes the first assumption regarding zero premium accrual. We can then discretized equation (9.36):

$$c^{ds}(T) = f \frac{(1-R)\sum_{j=1}^{Tf} D(0,t_j)\widehat{\mathbb{Q}}(\tau = T_j)}{\sum_{j=1}^{Tf} D(0,t_j)S(t_j)} \tag{9.38}$$

$$= f \frac{(1-R)\sum_{j=1}^{Tf} D(0,t_j)(S(t_{j-1}) - S(t_j))}{\sum_{j=1}^{Tf} D(0,t_j)S(t_j)} \tag{9.39}$$

Finally, let us summarize the above assumptions in the following result for the pricing of a CDS.

Proposition 9.2. *Assume that the default probability of an entity is given by a Cox Process so that the survival probability is given by Lemma 9.2. Further assume that i) the recovery fraction R is a constant, ii) default is calculated as if it occurred simultaneously with the forthcoming premium payment, iii) there is independence between the default intensity and the short rate process and iv) the premium payments are payed with a frequency/year denominator f.*

The fair constant premium $c^{ds}(T)$ on a CDS written on this entity with maturity T is then given by

$$c^{ds}(T) = f \frac{(1-R)\sum_{j=1}^{Tf} D(0,t_j)\Big(S(t_{j-1}) - S(t_j)\Big)}{\sum_{j=1}^{Tf} D(0,t_j)S(t_j)}.$$

The fair premium under the mentioned assumptions basically requires two types of inputs: survival probabilities and ZCB prices. Equivalently, we can find the survival probabilities (and thus the default probabilities) by using CDS premiums as input along with ZCB prices. E.g. with yearly premium payments and setting $T = 1$ we get

$$c^{ds}(1) = \frac{(1-R)\widehat{\mathbb{Q}}(\tau = 1)}{1 - \widehat{\mathbb{Q}}(\tau = 1)} \tag{9.40}$$

which can be explicitly solved to find $\widehat{\mathbb{Q}}(\tau = 1)$. We can then set $T = 2$ and solve

$$c^{ds}(2) = \frac{(1-R)(D(0,1)\widehat{\mathbb{Q}}(\tau = 1) + D(0,2)\widehat{\mathbb{Q}}(\tau = 2))}{D(0,1)(1 - \widehat{\mathbb{Q}}(\tau = 1)) + D(0,2)(1 - \widehat{\mathbb{Q}}(\tau = 1) - \widehat{\mathbb{Q}}(\tau = 2))} \tag{9.41}$$

for $\widehat{\mathbb{Q}}(T = 2)$ and so on.

By inferring the different default probabilities one is deriving the so called *default term structure*. Note how this is done by assuming that a credit event is equivalent to a default event, which typically is not true (see Section 3.4 on page 18). Note also how this *can* be done without specifying a model. In our upcoming empirical investigations we will nevertheless explicitly establish a model in order to obtain default probabilities in any point of time independently of data availability.

9.4 Affine Modeling

In order to truly apply the intensity framework, one has to specify an intensity process. By choosing the process consistent with the affine framework described in Chapter 6, some key properties become easier to obtain.

Proposition 6.1 on page 53 showed how affine assumptions made it possible to compute semi-analytical solutions to the cardinal expression

$$D(0,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T r(s)ds\right)\right].$$

Now, considering an intensity setup where from Lemma 9.2 we have the following relationship

$$\mathbb{Q}(\tau > T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T \lambda\Big(X(s)\Big)ds\right)\right]$$

we see that the interest rate has been swapped for the intensity. It is thus clear that we can find the risk neutral survival probabilities in closed form as given by Proposition 6.1:

$$\mathbb{Q}(\tau > T) = \exp\Big(A(0,T) - B(0,T)^{\intercal}\lambda_0\Big) \tag{9.42}$$

where λ_0 is a vector.

9.4.1 The One-Factor CIR Intensity Model

Since the CIR model will also be our model foundation in modeling actual default probabilities, we will end this chapter by a short comment on that assumption.

By assuming that the intensity process is of the one-factor CIR type, it becomes apparent that the above mentioned ODE solutions are given in closed form. The solutions are then given by equation (6.16) and (6.17) on page 57. So by choosing a CIR intensity model, analytical solutions to the survival probabilities exist. The model implied default term structure is thus very easy to compute, given a full CIR parameter set, including the latent variable λ_0.

Part III

Applications

Chapter 10

Empirical Remarks

We now turn to the question of how one could estimate the parameters in our two models; the swaption pricing model and the default model. The CIR model will constitute the foundation in both, but the estimation procedures will differ. The difference is pronounced due to one major reason: while a true risk free asset might not exist, a proxy for the asset does, and the interest rate is as such an *observable* size. Consequently, interest rate time series exist and fitting to these is deemed possible. This is done by using the maximum likelihood method described in Kladivko (2004) and Brigo et al. (2007). Interest rate parameters are thus fitted to interest data rather than swaption prices since we are looking for *global* estimates which are applicable in our default model as well.

Contrary to the interest rate case, a default intensity time series *does not* exist and another source of market data must be incorporated. Given the estimated interest rate parameters, we choose to fit the intensity model to observed CDS quotes. As discussed previously, this has the disadvantage of assuming that a credit event is the very same as a default event, which is not necessarily true. Nevertheless, the approach will display a different (more classical) estimation routine which is close to what is proposed in Mortensen (2006).

Apart from the two estimation procedures a thorough examination of the swaption pricing procedure will be conducted. The authors have chosen to assign a decent amount of weight here due to the complexity of the model and the many possible sources of (approximation) errors.

The part will be completed by a presentation and a discussion of the final results, the CVA prices. All our investigations are done assuming two different counterparties; the bank HSBC and the automobile manufacturer Fiat. For each counterparty we place ourselves at two different points in time; 1 September 2008 and the 1 June 2011. Naturally, we will only use historical observations to estimate our models. This applies to both dates. The amount of historical data used in each estimation routine might differ according to data availability.

Chapter 11

Interest Rate Estimation

The goal of this chapter is to find a proxy for the risk free interest rate and then estimate the parameters for our one-factor CIR short rate model based on a time series of this proxy. The choice of a proxy for the risk free interest rate is essential for this thesis paper, since both our swaption pricing model as well as the default probability estimation rely on the short rate parameters.

We begin by introducing the estimation strategy for the one-factor CIR model. Here we will choose a maximum likelihood estimation scheme, which works especially well for our specific short rate model. Then we will move on to briefly discuss the implementation of the estimation scheme, which will be carried out in *Matlab*[1]. The estimation technique and the implementation in Matlab along with Matlab code will follow Brigo et al. (2007) and Kladivko (2004). Afterwards, we will discuss various proxies for a risk free interest rate and then try to justify our specific choice. We end this chapter by presenting the estimates and comment on the results.

11.1 Estimation Strategy

The idea is to find estimates for the parameter vector $\varpi = (\psi, \theta, \sigma)$ using maximum likelihood. Therefore, we first need to define the transition densities for the CIR process. It turns out that for the CIR process, the transition densities are known in closed form. Now, in order to estimate the parameters we use a discretized version of the CIR process, such that

$$r(t + \Delta t) = r(t) + \psi(\theta - r(t))\Delta t + \sigma\sqrt{r(t)}\sqrt{\Delta t}\epsilon(t), \tag{11.1}$$

where $\epsilon(t)$ is standard normally distributed. Then, given $r(t)$ at time t the conditional density of $r(t + \Delta t)$ at time $t + \Delta t$ is given by

$$f_{CIR}(r(t + \Delta t)|r(t); \varpi, \Delta t) = c\exp(-u - v)\left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}) \tag{11.2}$$

[1] See `http://www.mathworks.com/products/matlab/`

where

$$\begin{aligned} c &= \frac{2\psi}{\sigma(1-\exp(-\psi\Delta t))}, \\ u &= cr(t)\exp(-\psi\Delta t), \\ v &= cr(t+\Delta t), \\ q &= \frac{2\psi\theta}{\sigma^2} - 1. \end{aligned}$$

Here I_q denotes the modified Bessel function of the first kind of order q.[2] As mentioned in section 6.2.5, the one-factor CIR process follows a non-central chi-squared distribution with $2q+2$ degrees of freedom and non-centrality parameter $2u$, i.e.

$$r(t+\Delta t)|r(t) \sim \chi^2(2cr(t), 2q+2, 2u).$$

We base our parameter estimation on a time series consisting of N observations, which are equally spaced with time step Δt. Hence, we can establish the log-likelihood function as

$$\ln(L(\varpi)) = \prod_{i=1}^{N-1} \ln\left(f_{CIR}(r(t+\Delta t)|r(t); \varpi, \Delta t)\right). \tag{11.3}$$

This means that we can find the maximum likelihood parameters by maximizing the log-likelihood function, such that

$$\hat{\varpi} = (\hat{\psi}, \hat{\theta}, \hat{\sigma}) = \arg\max_{\varpi} \ln(L(\varpi)), \tag{11.4}$$

where $\varpi = (\psi, \theta, \sigma) \in \mathbb{R}_+$.

We choose the initial estimates by using the Ordinary Least Squares method (OLS) on (11.1). Hence, by dividing (11.1) by $\sqrt{r(t)}$, we find the initial estimates for ψ and θ by minimizing the following OLS objective function

$$(\hat{\psi}, \hat{\theta}) = \arg\min_{\psi,\theta} \sum_{i=1}^{N-1} \left(\frac{r(t_{i+1}) - r(t_i)}{\sqrt{r(t_i)}} - \frac{\psi\theta\Delta t}{\sqrt{r(t_i)}} + \psi\sqrt{r(t_i)}\Delta t \right)^2. \tag{11.5}$$

Then, $\hat{\sigma}$ is found as the standard deviation of the residuals from $(\hat{\psi}, \hat{\theta})$. The solutions for (11.5) can be found in Appendix C.

11.1.1 Implementation

As mentioned in the introduction to this chapter, we will use Matlab to implement the estimation procedure. Matlab is a well known and often used programming language in the financial industry. It contains a lot of built-in mathematical functions and algorithms, which makes it extremely useful and easy to use.

[2]For any real number q, the modified Bessel function of the first kind can be expressed as

$$I_q(z) = \left(\frac{1}{2}z\right)^q \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(q+k+1)},$$

where $\Gamma(z)$ is the gamma function.

When calculating the transition densities, we have to use the modified Bessel function of the first kind. Luckily, Matlab contains this function in the command *besseli*. In order to maximize the log-likelihood and thereby solve the optimization problem in (11.4), we will use the Matlab function *fminsearch*. This Matlab function finds the minimum of a scalar function of several variables given initial estimates using a simplex search method known as the Nelder-Mead Simplex Method.[3] Selected Matlab code can be found in Appendix F.

11.2 Choice of Risk Free Rate

We now turn to the task of deciding which interest rate we will use as a proxy for the risk free rate. First, a risk free interest rate is a term which is purely theoretical and does not exist in real life. Nevertheless, the role of a risk free rate plays a fundamental role in financial theory, and therefore we need to find a reasonable proxy in order to estimate the short rate parameters.
Standard market practice is to use one of the following two instruments

1. Swap rates.
2. Government issued bonds.

Since this thesis paper concerns the pricing of credit risky swaps and since (by definition) the swap rate is the rate that renders the value of the swap contract zero upon inception, it seems somewhat counterintuitive to use swap rates as proxies for the risk free interest rate. Instead, we will use the yields on government bonds as a proxy. The reason for choosing government issued bonds is that the credit risk associated with these bonds is usually close to zero. Although this is not always the case for certain sovereign countries issuing government bonds, the chances of historically stable western countries defaulting on their obligations have usually been considered insignificant. In turn, this has resulted in a very liquid market where strong demand has driven prices up and yields down. However, there are other effects than credit risk that determine the level of the yields. We will now try to name some of these effects:

- Investors usually prefer bonds with high liquidity, which applies for certain government bonds. Especially during financial crises such as the one we experienced in 2008, where markets are in distress and bond liquidity in general is scarce, investors usually place their money in government bonds as a sort of safe haven.
- Because of regulatory requirements, financial institutions hold a certain amount of government bonds.
- The capital requirements that financial institutions face when investing in bonds are usually lower for government bonds compared to other bonds.
- Some Central Banks invest in their own government bonds in order to stimulate the economy in times of economic downturn. The latest example of this is in the United States where the Federal Reserve since late 2008 has been actively buying both long and short term US government bonds.

[3]For more information on the Matlab function fminsearch go to `http://www.mathworks.com/help/techdoc/ref/fminsearch.html`.

All these effects increase the demand for government bonds, which in turn drive yields down. Therefore, one could argue that the yield on these bonds is artificially low and not just a consequence of low credit risk. Hence, a risk free interest rate should probably be higher. However, we still believe the yield of government bonds is the best proxy for a risk free interest rate and we will therefore choose a specific government bond time series in order to estimate our parameters for the short rate.

11.3 Data Description

Having decided on which type of instruments to use as our risk free rate, we now need to decide exactly which government bonds to use and with what maturity. Here we follow Nowman (1997) and choose U.S. Treasury bills. More precisely, we choose three-month benchmark bond yield data.[4] The term benchmark bond implies the latest issue of the U.S. Treasury bills with a maturity of three months. The time series consists of daily yield data from 11 November 2005 to 1 June 2011 for each business day. The time series is shown in Figure 11.1. The data are retrieved from *Nordea Analytics*.[5]

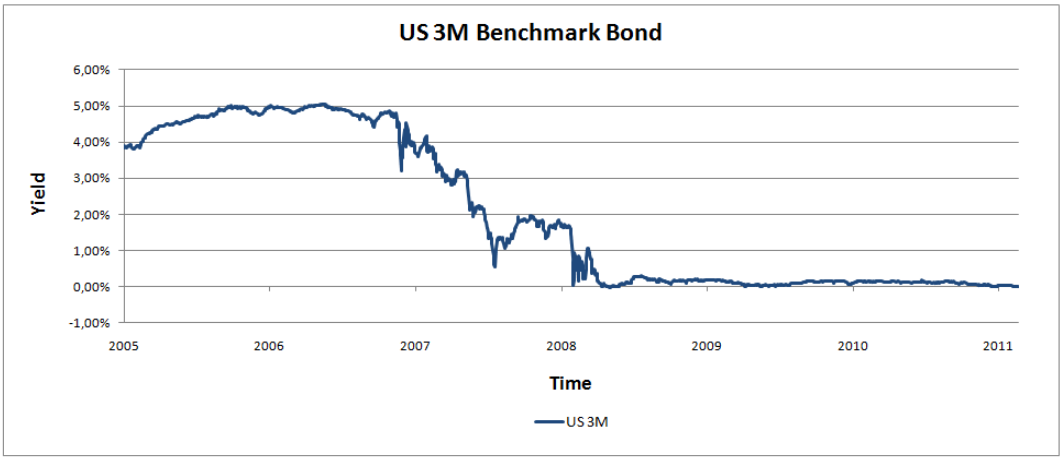


Figure 11.1: *Time series of three-month US Treasury Benchmark Bonds yield data from 21-11-2005 to 01-06-2011.*

Looking at Figure 11.1, we can conclude that during the last 5 years yields on three-months U.S. Treasury bill have been very volatile. The time series can roughly be divided into three periods. The first period is before the financial crisis where yields where relatively high and stable. The second period began in 2007 where the financial markets started experiencing the first signs of distress and ended in late 2008 where the financial crisis peaked. In this period yields where very volatile and generally downward sloping. The third period is where we are now, where yields are stable at an historically low level.

As already mentioned, it is quite evident from Figure 11.1 that the arrival of the financial crisis in 2007-2008 drove the yield down towards zero. This was probably a consequence of the general opinion that the United States government would not default on their debt. Hence, investors started buying US Treasury bills because of their high liquidity and low credit risk. Two very attractive attributes in a distressed market. Also, around that time the Federal Reserve began

[4] Nowman (1997) uses the U.S. Treasury bill one-month yield data.

[5] See `http://nordea.eu/Vores+serviceydelser/Internationale+produkter+og+serviceydelser/Markets/Nordea+Analytics/906592.html`.

buying short term government bonds in order to drive yields down and thereby increase access to cheap credit. These two factors are probably the main reasons for the very steep decline in the yield. The extraordinarily strong demand for treasuries even meant that the yield became negative a couple of times during December 2008! Since negative yields are not consistent with our one-factor CIR process, we choose to change those observations to 0.00001 in order to estimate the parameters. Given the two dates specified in Chapter 10, we divide the time series into two overlapping periods:

- First period covers the yield from 21-11-2005 to 02-09-2008, which is a total of 1106 observations.
- Second period covers the yield from 21-11-2005 to 01-06-2011, which is a total of 1414 observations.

Our goal is to estimate two sets of parameters, one set for each period.

11.4 Results

The maximum likelihood estimates are shown in Table 11.1.

Estimation period	ψ	θ	σ
21-11-2005 – 02-09-2008	0.2592	0.0063	0.0840
21-11-2005 – 01-06-2011	0.6957	0.0097	0.1448

Table 11.1: *The maximum likelihood estimates for the short rate parameters for each period.*

If we compare the results in Table 11.1 with the time series in Figure 11.1, we can conclude that the maximum likelihood estimates seem very dependent on the amount of observations that the estimation is based on. This is especially evident for the σ estimates, which are significantly different for the two periods. However, it seems reasonable that σ is higher for the longest period, since it captures all three periods described in the previous section. The θ estimates also seem to make sense at their low levels, since they express the mean-reversion level for the short rate. The fact that ψ is higher for the longest period is justified by the fact that the yield experiences a longer interval around its mean reversion level.

In Figure 11.2 we have shown simulated paths for both sets of parameters initiated at $r(0) = 0.0389$, which was the yield observed on the 21th of November 2005. The simulation is performed using an Euler approximation. Normally, the square root operator in the volatility term ensures that the short rate cannot become negative. However, since we have to approximate the process in order to simulate it, the short rate can go below zero and thereby stop the simulation. This is likely to happen since the mean-reversion parameter for both processes is close to zero. Therefore, we use the following modified Euler approximation:

$$r(t_{i+1}) = r(t_i) + \psi(\theta - \max(r(t_i), 0))(t_{i+1} - t_i) + \sigma\sqrt{\max(r(t_i), 0)}\sqrt{t_{i+1} - t_i}\epsilon(t_i). \tag{11.6}$$

The max operator in (11.6) will not necessarily ensure positive short rates, it will, however, ensure that the simulation works even if the short rate drops below zero.
As we can see in Figure 11.2, both short rates tend towards zero which is in line with the low mean-reversion parameters and the observed time series in Figure 11.1. The simulation based on the parameters estimated for the longest period experiences higher volatility, which corresponds well with the higher level of σ.

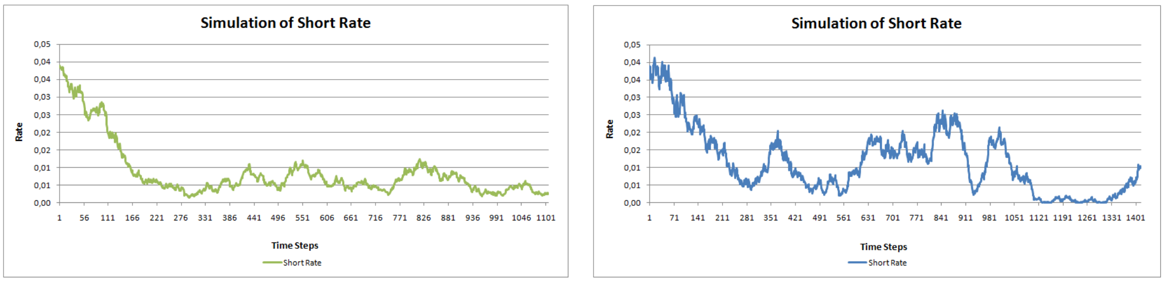


Figure 11.2: *Simulations of the short rate for the two sets of parameters. The panel to the left is for the period 21-11-2005 to 02-09-2008 with 1106 time steps. The panel on the right is for the period 21-11-2005 to 01-06-2011 with 1414 time steps. Both simulations are initiated at* $r(0) = 0.0389$.

Chapter 12

Default Estimation

This chapter concerns our implementation of the intensity framework as described in Chapter 9. After the strategy is introduced we will turn to a brief description of the companies that constitute our fictional swap counterparties, and subsequently their CDS data will be presented. Afterwards, we will comment on our programming routine in R^1 and the results will be displayed in terms of goodness of fit and the model entailed default term structure. The latter is obviously the main product of this chapter.

12.1 Estimation Strategy

Our goal of the estimation process is to use a decent amount of market information under the precondition that the routine is intuitive and that the interpretation of the results is somewhat straightforward.

Slightly inspired by Mortensen (2006), we use a rather simple and intuitive routine which goes as follows.

For a given entity,

- Select a historical time period consisting of N days and collect CDS premia for n different maturities for each day. Consequently, the *CDS term structure* is used for each day.
- Choose an initial guess of the parameter vector $\varphi = (\kappa, \theta, \sigma, \lambda_0)$.
- Compute the vector of CDS premiums c_{fit}^{ds}, as described in Chapter 9, making use of the fact that the survival probabilities are given in closed form.
- Obtain the optimal parameter vector φ_i for day i by minimizing the Root Mean Squared Error (RMSE) between the observed CDS quotes c_{obs}^{ds} and fitted CDS quotes c_{fit}^{ds}:

$$\varphi_i = \arg\min \text{RMSE}_i \tag{12.1}$$

$$\text{RMSE}_i = \sqrt{\frac{1}{n}\sum_{k=1}^{n} c_{obs}^{ds} - c_{fit}^{ds}} \tag{12.2}$$

- Compute the final parameter vector $\varphi = \frac{1}{N}\sum_{i=1}^{N} \varphi_i$.

[1] See `http://www.r-project.org/`.

Comments. First, note that contrary to our interest rate estimation, we choose to estimate all four parameters including the unobservable parameter λ_0. An initial guess of λ_0 is still necessary, while it will often be rather arbitrary since it is unobservable (unlike an initial interest rate), but choosing it positive and close to zero should generally provide a smooth minimization process. Second, note that the final parameter vector is a simple average of each day's estimated parameters. The reasonableness of this approach is, of course, most sensible when the variation of the different parameter vectors is fairly small.

12.2 Implementation

Implementing the above process in *R* involves minimizing an object function (the RMSE) which is done using the built-in function *nlminb*. In doing this, we choose to invoke some parameter restrictions. For our purpose we choose to restrict each parameter to the range $[0, 1]$, which ensures that one extreme estimation does not affect the average too much. The code for this exercise can be found in Appendix E.

12.3 Data Description

Since the CVA is greatly dependent on the default probabilities, we choose to use premia from CDS' written on entities with CDS premia that differ greatly. As mentioned, our choice falls on the two large international companies HSBC and Fiat.

On June 2, 2011, HSBC was the world's second largest bank according to Forbes[2] with around 0.3m employees and a market value of USD 186.5b. With such strong key figures it is not surprising that the CDS premia on HSBC are usually rather low. However, during late 2007 HSBC's CDS premia rose significantly, which becomes evident by considering Figure 12.1.

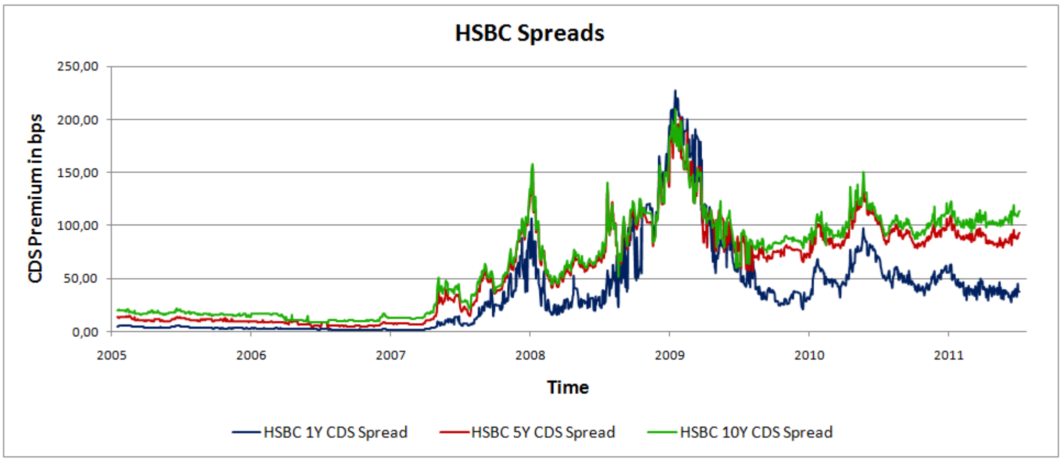


Figure 12.1: *HSBC CDS Spreads for three different maturities with the mod-mod convention in use.*

It is also notable how most of the time premia seem to be higher for longer maturities. This is however not always the case as exhibited in the HSBC data in the beginning of 2009. Reason probably being that the market was unsure of how HSBC would handle the crisis now that Lehman Brothers had defaulted[3]. But if HSBC would indeed survive the crisis, then they would

[2]See `http://www.forbes.com/companies/hsbc-holdings/`
[3]The default happened September 15, 2008.

probably survive for a longer period, i.e. the *conditional* probability of a future default, given survival of the crisis was probably small. Consequently, spreads for longer maturities went down and spreads for shorter maturities went up.

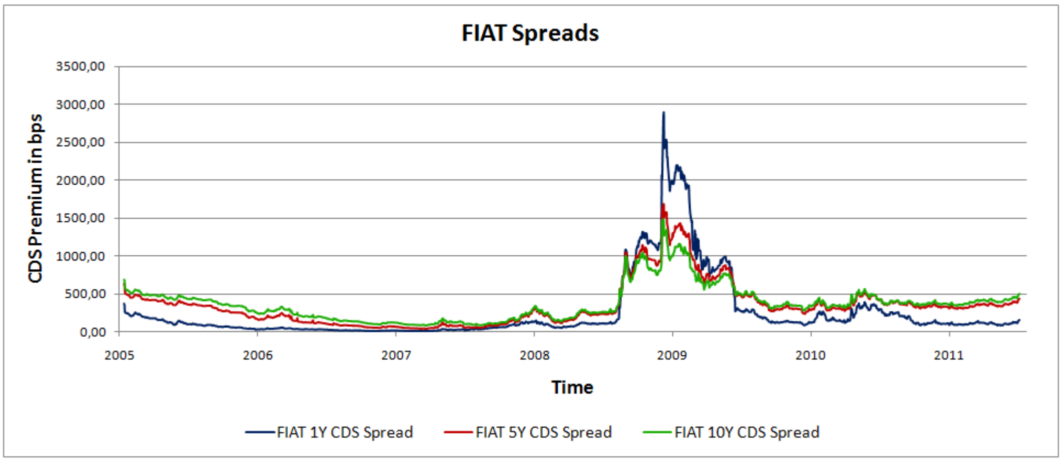


Figure 12.2: *Fiat CDS Spreads for three different maturities with the mod-mod convention in use.*

Turning to our other firm of interest, Fiat, we keep in mind that the car industry was very exposed during the crisis. On 2 June, the company had approximately 0.14m employees and a market value of USD 73.4 billion. Considering Figure 12.2, we observe a much higher overall level of CDS premia. The premia rose to nearly 3,000 bps for the one year maturity in January 2009, which corresponds to a premium of 30% of the notional on the CDS. Again, we see how shorter maturities might demand a higher spread than longer maturities. It is interesting that HSBC quotes reacted earlier than Fiat to the generic financial instabilities. It is a small piece of evidence showing how initially the crisis only concerned banks, but later became a trans-industrial and global event.

12.4 Results

We have chosen to monitor CDS quotes at the first business day of each month for a total of 12 months. This is done for each of our two firms in each of the two periods of interest. Hence, for each company we estimate using 12 CDS term structures, each consisting of the maturities 1Y, 2Y, 3Y, 4Y, 5Y, 7Y and 10Y. We will always assume a constant recovery rate R of 40%.

12.4.1 CIR Estimates

Our estimated parameters for HSBC are shown in Table 12.1.

Estimation period	Estimate/Standard dev.	κ	θ	σ	$\lambda(0)$
01-10-2007 – 02-09-2008	Estimate	0.287764	0.042305	0.478656	0.002425
	Standard deviation	0.315353	0.028069	0.306680	0.002952
01-07-2010 – 01-06-2011	Estimate	0.067957	0.093736	0.359710	0.004633
	Standard deviation	0.014547	0.015525	0.065581	0.002236

Table 12.1: *Estimated CIR intensity parameters for HSBC over two disjoint time periods.*

It is not surprising that we estimate a higher σ and get higher parameter standard deviations in the first period. Consequently, parameters in the first period are less statistically significant[4] for HSBC.

Estimation period	Estimate/Standard dev.	κ	θ	σ	$\lambda(0)$
01-10-2007 – 02-09-2008	Estimate	0.120936	0.128589	0.431823	0.005630
	Standard deviation	0.066518	0.048137	0.281103	0.002729
01-07-2010 – 01-06-2011	Estimate	0.305948	0.497977	0.767089	0.004233
	Standard deviation	0.419834	0.296556	0.181186	0.009897

Table 12.2: *Estimated CIR intensity parameters for Fiat over two disjoint time periods.*

Considering Table 12.2 we observe that, in comparison with HSBC, the significance of parameters varies more. In general we also estimate higher mean reversion levels θ in the Fiat dataset which corresponds to the higher overall level of CDS premia. The θ in the second period seems extremely high, however. Estimates generally increase going from the first period to the second one.

12.4.2 Goodness of Fit

Table 12.3 shows how well the model fits to each month in average.

Estimation period	HSBC	Fiat
01-10-2007 – 02-09-2008	0.000157	0.000295
	(0.000084)	(0.000126)
01-07-2010 – 01-06-2011	0.000261	0.000625
	(0.000107)	(0.000124)

Table 12.3: *The average of the twelve RMSE values in each period. Numbers in parentheses indicate the respective estimated standard deviations of mean RMSEs.*

Clearly, parameters in the HSBC dataset are estimated better. We generally obtain the lowest RMSE values in our earliest period, which might come as a surprise. It is albeit important to remember that the first part of the former period was relatively steady for Fiat and thus easier to estimate upon.

We remind ourselves that the reported RMSE-values are only relevant when considering parameters *before* the average is computed. After the average is calculated, it would not make sense to evaluate the goodness of fit in each month, since we would be prone to use future information, e.g. an RMSE at month 3 *would* be using information from month 12.

[4]By assuming that each estimate is normally distributed with a mean that equals the estimate and a standard deviation as reported, we have that with 95 % probability the estimate lies in the interval

$$\text{estimate} \pm 1.96 \times \text{std. dev.}$$

so that $\frac{\text{estimate}}{\text{std. dev.}} < 1.96$ means that the estimate may be zero according to the normal distribution assumption, i.e. it can be regarded insignificant.

12.4.3 Default Estimates

Given the above parameters we can obtain the default term structures as shown in Figure 12.3.

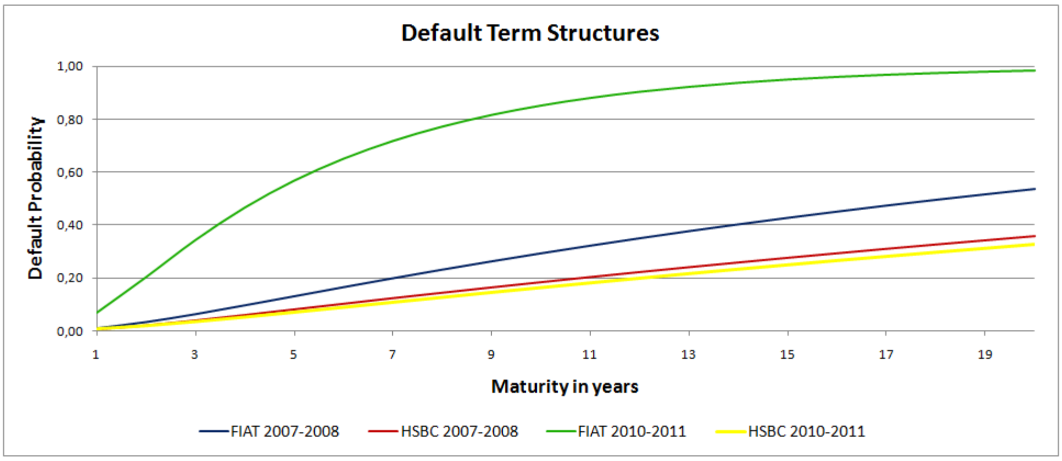


Figure 12.3: *Accumulated default probabilities according to four estimated CIR models. Each point on a graph is thus the estimated probability of a default event in less than x years for the company in question standing at the latest point of time in the estimation period.*

The first thing one might notice is the high probability of Fiat defaulting if we position ourselves in June 2011. These probabilities might not seem reasonable when using one's intuition but they are nevertheless a product of the high estimates of the volatility σ and the mean reversion level θ. Another important comment is that our estimation routine is very sensitive to each observation since we estimate over a relatively small time series. As with all historical estimations, there is no right or wrong length of time period to use. It all depends on how much emphasis one wants to put on the past. There are tons of ways to calibrate or estimate models. Our above example is solely a demonstration of how one *might* estimate intensity parameters in a one-factor CIR model using market data in an intuitive manner without consuming to much CPU time.

As reported in Table 12.4, we will finalize this chapter by displaying the one-year model implied default probabilities that are to be used in the CVA pricing.

Year	Fiat 2007-2008	HSBC 2007-2008	Fiat 2010-2011	HSBC 2010-2011
1	0.012439	0.007464	0.067222	0.007442
2	0.023212	0.014792	0.134977	0.012111
3	0.029742	0.018710	0.140018	0.015318
4	0.032984	0.020477	0.122510	0.017288
5	0.034125	0.021082	0.101603	0.018364
6	0.034069	0.021107	0.082822	0.018852
7	0.033381	0.020857	0.067138	0.018976
8	0.032383	0.020483	0.054325	0.018880
9	0.031247	0.020056	0.043932	0.018658
10	0.030066	0.019610	0.035521	0.018363

Table 12.4: *Estimated probabilities of a default in the year in question. The table is essentially based on the same information as Figure 12.3.*

Chapter 13

Swaption Model Implementation

We now turn to the implementation of the swaption pricing setup described in Chapter 8 based on a one-factor CIR short rate model using the parameters estimated in Chapter 11.

We begin by briefly describing the consequences for the affine swaption setup when the underlying short rate follows a one-factor CIR process. Then we will discuss the implementation of the model and touch upon certain computational aspects. The implementation is done in *R*. Next, we will focus on testing key parts of our implementation and the LVM assumption, which is the foundation behind the model. The tests will be based on Monte Carlo simulations. We end this chapter by presenting swaption prices for swaptions with specific expiries and underlying tenors based on the two parameter sets found in Chapter 11. Selected *R* code can be found in Appendix E.

13.1 Model Setup

In Chapter 8 we presented the affine swaption setup for an N-factor short rate model. Based on these results, we can derive the one-factor setup rather straightforwardly.

We begin by specifying the short rate and the swap rate dynamics. From Section 8.3, we have that in a one-factor setup, the dynamics of $d\tilde{X}(t) = (y_{n,N}(t)\ X(t))^\intercal$ is given by

$$\begin{aligned} d\tilde{X}(t) &= \tilde{\Psi}(t)(\tilde{\Theta}(t) - \tilde{X}(t))dt + \tilde{\Sigma}(t)\sqrt{\tilde{V}(t)}dW^{\mathbb{Q}^{n+1,N}}(t) \\ &= \begin{pmatrix} 0 & 0 \\ 0 & \Psi + \Sigma^2 w(t) \end{pmatrix} \left(\begin{pmatrix} 0 \\ \frac{\Psi\Theta}{\Psi + \Sigma^2 w(t)} \end{pmatrix} - \tilde{X}(t) \right) dt + \begin{pmatrix} 0 & k(t)\Sigma \\ 0 & \Sigma \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\tilde{X}(t)} \end{pmatrix} dW^{\mathbb{Q}^{n+1,N}}(t). \end{aligned} \tag{13.1}$$

Letting $r(t) = X(t)$, and using the notation from Chapter 11 we get

$$dr(t) = \left(\psi\theta - (\psi + \sigma^2 w(t))r(t)\right) dt + \sigma\sqrt{r(t)}dW^{\mathbb{Q}^{n+1,N}}(t), \tag{13.2}$$

$$dy_{n,N}(t) = k(t)\sigma\sqrt{r(t)}dW^{\mathbb{Q}^{n+1,N}}(t) \tag{13.3}$$

where, as usual, $w(t) = \sum_{i=n+1}^{N} \Delta B(t,T_i)\frac{D(0,T_i)}{P_{n+1,N}(0)}$ and $k(t) = \sum_{i=n}^{N} q_i^y(0)B(t,T_i)$. From the dynamics in (13.1), we can derive the following two complex ODEs which need to be solved in order to obtain the CCF:

$$\frac{d\delta(t)}{dt} = \begin{pmatrix} 0 & 0 \\ 0 & \Psi + \Sigma^2 w(t) \end{pmatrix} \delta(t) - \frac{1}{2} \begin{pmatrix} 0 \\ \Big(k(t)\Sigma\delta_1(t) + \Sigma\delta_2(t)\Big)^2 \end{pmatrix} \tag{13.4}$$

$$\frac{d\gamma(t)}{dt} = -\Phi\Theta\delta_2(t) \tag{13.5}$$

along with the boundary conditions

$$\delta(T_n) = \begin{pmatrix} iv \\ 0 \end{pmatrix} \tag{13.6}$$

$$\gamma(T_n) = 0. \tag{13.7}$$

Hence, applying the notation from (13.2)-(13.3), (13.4)-(13.5) translates into

$$\frac{d\delta_1(t)}{dt} = iv \tag{13.8}$$

$$\frac{d\delta_2(t)}{dt} = \Big(\psi + \sigma^2 w(t)\Big)\,\delta_2(t) - \frac{1}{2}\Big(k(t)\sigma\delta_1 + \sigma\delta_2(t)\Big)^2 \tag{13.9}$$

$$\frac{d\gamma(t)}{dt} = -\psi\theta\delta_2(t). \tag{13.10}$$

Solving these equations gives us the CCF for the swap rate. Prices can then be obtained using (8.57) along with (8.58) on page 73 for payer swaptions and (8.59) along with (8.60) on page 73 for receiver swaptions.

13.2 Implementation and Computational Aspects

There are two main computational aspects we need to address when implementing the swaption model:

1. Solving a system of complex ODEs given known boundary conditions.
2. Numerically evaluate the integral required for the computation of swaption prices.

Both issues can be solved using functions in *R*.

In order to obtain the CCF for the swap rate we have to solve the two complex ODEs in (13.4) and (13.5) given their boundary conditions. To do this, we choose the fourth-order Runge-Kutta method, which is an iterative method for approximating solutions of ODEs. This method is known to be a very stable and accurate approximation method used extensively in financial literature. In *R*, the fourth-order Runge-Kutta method is contained in the function *bvptwp*. Now given boundary conditions, we work backwards from time T_n in order to approximate the function values at time zero. Time steps in between time zero and T_n are set to 0.5. To ensure correct implementation, we will compare values of the CCF obtained from solving the ODE equations with CCF values obtained from Monte Carlo simulation. We will return to this comparison in the next section.

When numerically evaluating the integrals required to price swaptions in (8.57) and (8.59) on page 73 and 73, we use the *R*-function *integrate*. This function uses an algorithm called an adaptive quadrature to approximate the integral. Generally, a quadrature rule is of the following form

$$\sum_{i=0}^{n} w_i f(x_i) \approx \int_a^b f(x)dx$$

where the nodes x_i and the weights w_i are pre-computed. So, for an adaptive quadrature the integral is approximated using a quadrature rule on adaptively refined subintervals of the integral domain. There are many different types of quadrature rules, and the function *integrate* chooses the optimal quadrature rule depending on the shape of the integrand. When using the function, we have to specify the lower and the upper integration limit along with the number of subdivisions. The lower integration limit will always be set to zero according to 73 and 73, whereas the upper limit will depend on the speed of convergence to zero for the specific integrands. The number of subdivisions expresses a dynamic quantity that, given a certain error bound, determines the amount of subintervals included in the given quadrature rule. The maximum number of subdivisions will be set to 1000.

Evaluating the integrand requires that we solve the ODEs for each u we integrate over, which is costly in computational time when the integral domain is large. Therefore, in order to reduce the computational time, we will use a spline interpolation scheme for the solutions to the ODEs. A spline interpolation uses a sufficiently smooth piecewise-polynomial function called a spline to interpolate between given points. In *R* we can use the function *spline* to perform the interpolation. This function uses a cubic spline where the polynomial function is of order three. So, we solve the ODEs for a given (large) number of integers and then create a function based on an interpolation between the solutions. This makes the numerical integration significantly faster and means that we only have to solve the ODEs once for a given number of integers and swaption structure. The interpolation scheme works especially well in our case since the integrands are continuous functions. This should be evident later on when we test the spline interpolation by comparing it with analytical CCF values.

13.3 Model Verification

We now turn to the task of verifying the different methods discussed above thereby ensuring that each step in our implementation is performed correctly. We also want to investigate the LVM assumption, since the exactness of this assumption is essential in order for any practical use of the model to be justified.

All investigations will be carried out by comparing CCFs. Most of the verification procedures will consist of comparing the method implied CCF to a simulated CCF based on swap rate values achieved from simulating the $\mathbb{Q}^{n+1,N}$-dynamics of the short rate and swap rate using the Euler scheme. The simulation scheme is as follows

$$\begin{aligned} r(t_{i+1}) &= r(t_i) + \Big(\psi\theta - (\psi + \sigma^2 w(t_i)) r(t_i)\Big)(t_{i+1} - t_i) \\ &\quad + \sigma\sqrt{r(t_i)}\sqrt{t_{i+1} - t_i}\epsilon(t_i), \end{aligned} \tag{13.11}$$

$$y_{n,N}(t_{i+1}) = y_{n,N}(t_i) + k(t_i)\sigma\sqrt{r(t_i)}\sqrt{t_{i+1} - t_i}\epsilon(t_i), \tag{13.12}$$

for $t_i \in [0, T_n]$. We simulate 1000 paths for both dynamics where each path is divided into

50 time steps. The CCF is then obtained as the mean value of $\exp\left(ivy_{n,N}(T_n)\right)$ where $y_{n,N}$ denotes a vector of 1000 simulated swap rate values at time T_n. We base the verification of the implementation methods and LVM assumption on a specific swaption structure and specific short rate parameters. We choose a 5-year swaption on a 10-year swap, so that $T_n = 5$ and $T_N = 15$. The short rate parameters are $\psi = 0.69$, $\theta = 0.08$ and $\sigma = 0.03$ along with an initial short rate $r(0) = 0.08$.

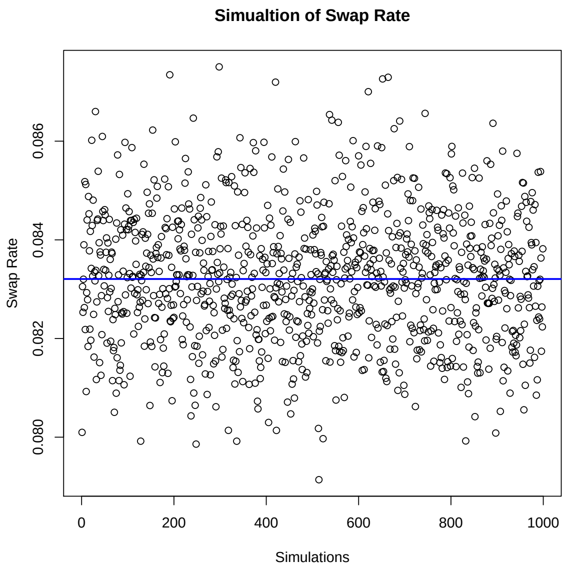


Figure 13.1: *Swap rate values at time T_n based on 1000 simulations where $\psi = 0.69$, $\theta = 0.08$, $\sigma = 0.03$ and $r(0) = 0.08$. The blue line indicates the value of the forward swap rate at time zero.*

In Figure 13.1 we have depicted 1000 simulated values of the swap rate at time T_n along with the forward swap rate at time zero. As the figure confirms, the swap rate is a $\mathbb{Q}^{n+1,N}$-martingale, so on average the swap rates at time T_n will equal the initial forward swap rate.

13.3.1 The Low Variance Martingale Assumption

Since the LVM assumption is the foundation for the entire swaption model setup, verifying that it actually holds, is obviously very important. As explained in Subsection 8.2.3, the LVM assumption implies that the terms $\frac{D(t,T_i)}{P_{n+1,N}(t)}$ in $k(t)$ and $w(t)$ can be replaced by their time zero values $\frac{D(0,T_i)}{P_{n+1,N}(0)}$. So in order to verify the assumption, we simulate (13.11) and (13.12) twice. First, we simulate where $k(t)$ and $w(t)$ include the time zero values. Second, we simulate where they do not. The verification procedure is performed by comparing CFF values based on both simulations.

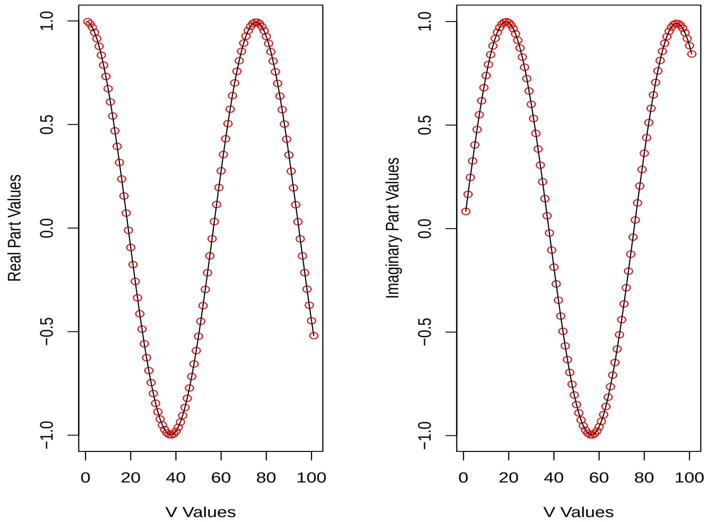


Figure 13.2: *CCF values with and without the LVM assumption based on 1000 simulations. The panel to the left illustrates the real part. The panel to the right illustrates the imaginary part. With LVM: the black line. Without LVM: the red points.*

The simulations indicate that the LVM assumption works extremely well. This is made evident in Figure 13.2, where we have depicted the real and imaginary parts of both simulated CCFs. As can be seen in the figure, the two CCFs are almost identical. The approximation errors depicted in Figure 13.3 are clearly very small, and therefore insignificant for all practical purposes. Hence, accepting the assumption seems reasonable and the implications for the rest of the model can therefore be justified.

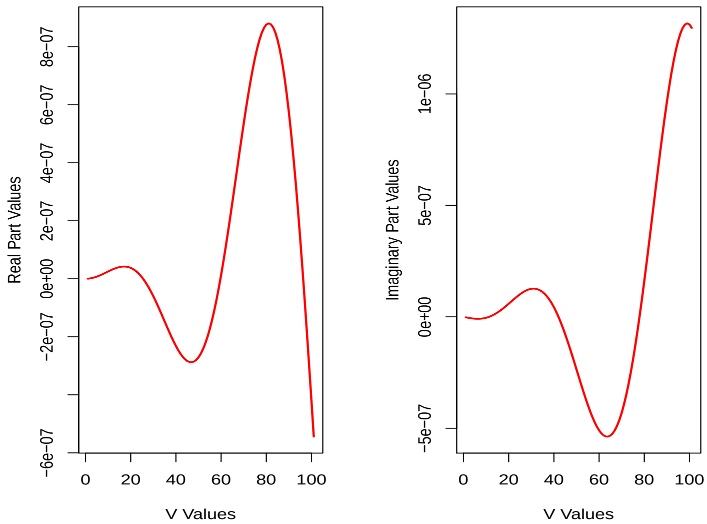


Figure 13.3: *Approximation errors of the CCF with and without the LVM assumption based on 1000 simulations. The panel to the left illustrates approximation errors for the real part. The panel to the right illustrates approximation errors for the imaginary part.*

13.3.2 The ODE Solutions

The next step is to verify the implementation of the two complex ODEs and make sure that the solutions are correct. Since the short rate and the swap rate are on affine form, we know that the CCF for the swap rate is found as

$$\phi(v, t, T_n) = \exp\left(\gamma(t) + \delta(t)\tilde{X}(t)\right) \tag{13.13}$$

where $\tilde{X}(t) = (y_{n,N}(t)\ r(t))$. Again, we verify the ODE solutions by simulating (13.11) and (13.12) to obtain a CCF and then compare the results with the analytical CCF from (13.13). The analytical and simulated CCFs are shown in Figure 13.4 divided into real and imaginary parts. The results clearly verify the solutions to the ODEs and show that (13.13) in fact does generate the CCF. This is also evident from the approximation errors depicted in Figure 13.5. Even though these errors are small, they are still noticeable. This could probably be improved by simulating a large number of paths for the short rate and swap rate. Overall, the results indicate that our implementation works and that the solutions generated by the fourth-order Runge-Kutta method can be applied.

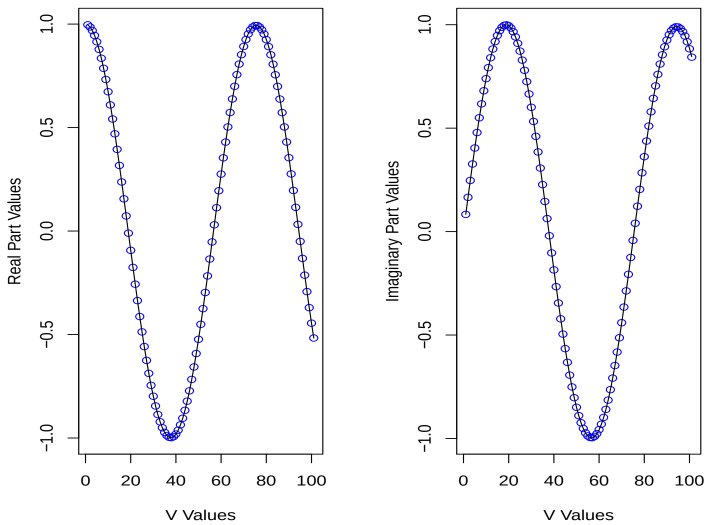


Figure 13.4: *Analytical and simulated CCF based on 1000 simulations. The panel to the left illustrates the real parts of both the analytical and simulated CCF. The panel to the right illustrates the imaginary parts. Analytical CCF: the blue points. Simulated CCF: the black line.*

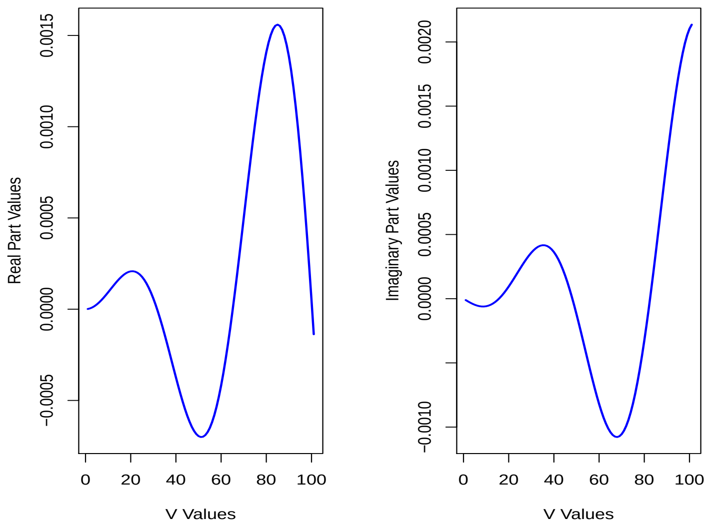


Figure 13.5: *Approximation errors for the analytical and simulated CCF based on 1000 simulations. The panel to the left illustrates approximation errors for the real part. The panel to the right illustrates approximation errors for the imaginary part.*

13.3.3 Spline Interpolation

When evaluating the integral in order to generate swaption prices, solving the complex ODEs for each u in the integrals domain is necessary. This can take considerable computational time, since the domain of the integral is usually chosen according to the shape and convergence of the integrand and can therefore potentially require a large domain. Thus, we will, as mentioned, incorporate a cubic spline interpolation scheme for the solutions to the ODEs, and thereby reduce the computational time greatly. In order to verify the accuracy of spline interpolation we will compare analytical CCF values with CCF values calculated from spline interpolating the ODE solutions. The results are depicted in Figure 13.6 along with the approximation errors in Figure 13.7. As the results indicate, the spline interpolation is very accurate and applying the interpolation scheme therefore seems like an obvious choice. The accuracy of the spline interpolation is due to the CCF being a continuous function.

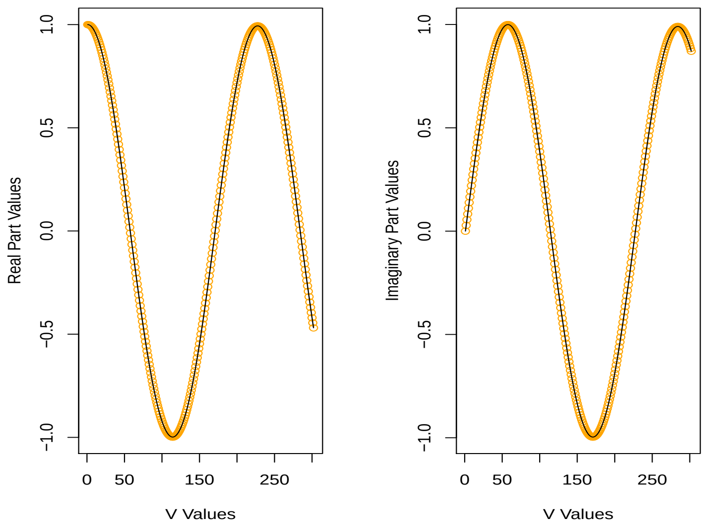


Figure 13.6: *Analytical and spline interpolated CCF. The panel to the left illustrates the real parts of both the analytical and interpolated CCF. The panel to the right illustrates the imaginary parts. Spline CCF: the blue points. Analytical CCF: the black line.*

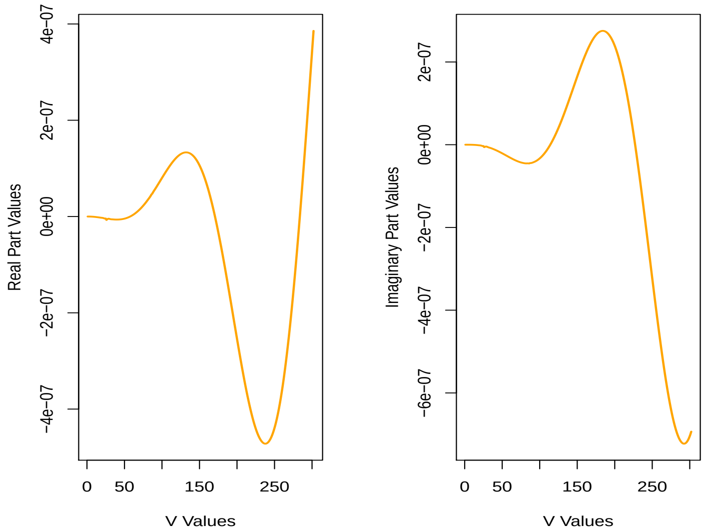


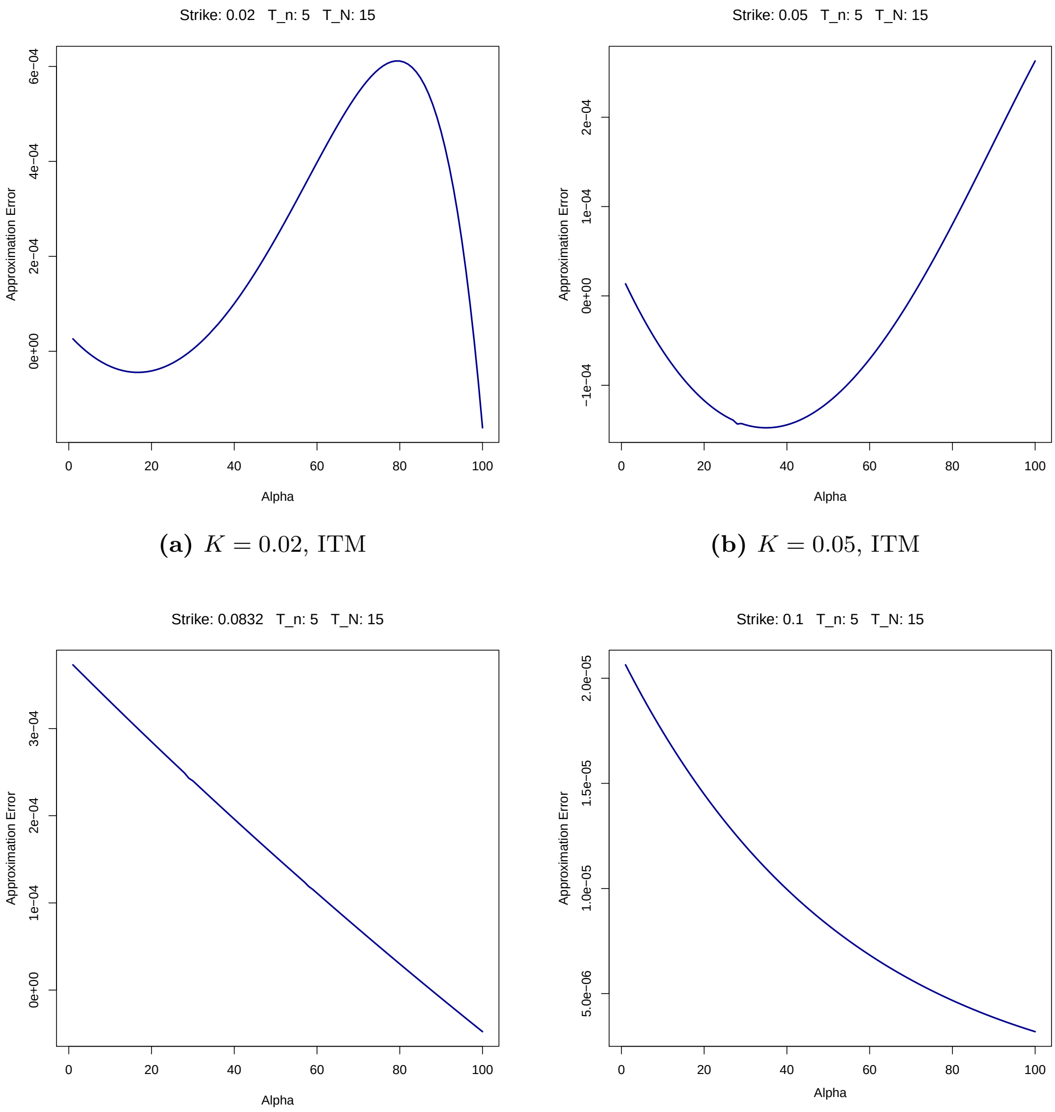
Figure 13.7: *Approximation errors for the analytical and spline interpolated CCF. The panel to the left illustrates approximation errors for the real part. The panel to the right illustrates approximation errors for the imaginary part.*

13.3.4 Choice of Alpha

Pricing using Fourier inversion implies choosing an α in order to ensure integrability when establishing the Fourier transform of the swaption price. In addition to a strictly positive restriction on α, it is not immediately obvious which α to choose. Lee (2004) proposes a theoretical approach in order to decide which α to apply when pricing options in general. We choose a different approach based on comparing prices in our setup with prices based on Monte Carlo simulation.

More precisely, we simulate 10,000 paths for (13.11) and (13.12) and then obtain prices for payer and receiver swaptions by calculating (8.25) and (8.26) on page 68 based on 10,000 swap rate values at time T_n. These prices will be independent of α and can therefore be used as a benchmark for the semi-analytical prices obtained from Fourier inversion. The semi-analytical prices are computed for levels of α ranging from 0.1 to 100. The optimal choice of α is the level that produces the smallest price difference. The approximation errors are depicted in Figure 13.8 for payer swaptions and in Figure 13.9 for receiver swaptions.

Generally, the approximation errors are small which confirms that our pricing setup works. However, choosing one optimal level for α across different strikes for both payers and receivers seems almost impossible based on the results. The approximation errors for payer and receiver swaptions seem to show the same pattern. For ITM swaptions, two choices for α generate an approximation error of zero (this is not the case for ITM receiver swaption in panel (d) in Figure 13.9). For OTM and ATM only one α gives a perfect fit. So, we can conclude that although the price differences are small the swaption prices achieved from Fourier inversion for this specific swaption structure seem very dependent on the level of α. This can be considered a model weakness, since α would have to be adjusted to strike level in order to maximize precision. All in all, the optimal choice of α would demand a more thorough investigation involving the relevant swaption structures and levels of moneyness. Nevertheless, for our purpose, we will only choose one alpha to price both payer and receiver swaptions regardless of strike level. This level will be set at $\alpha = 1$.



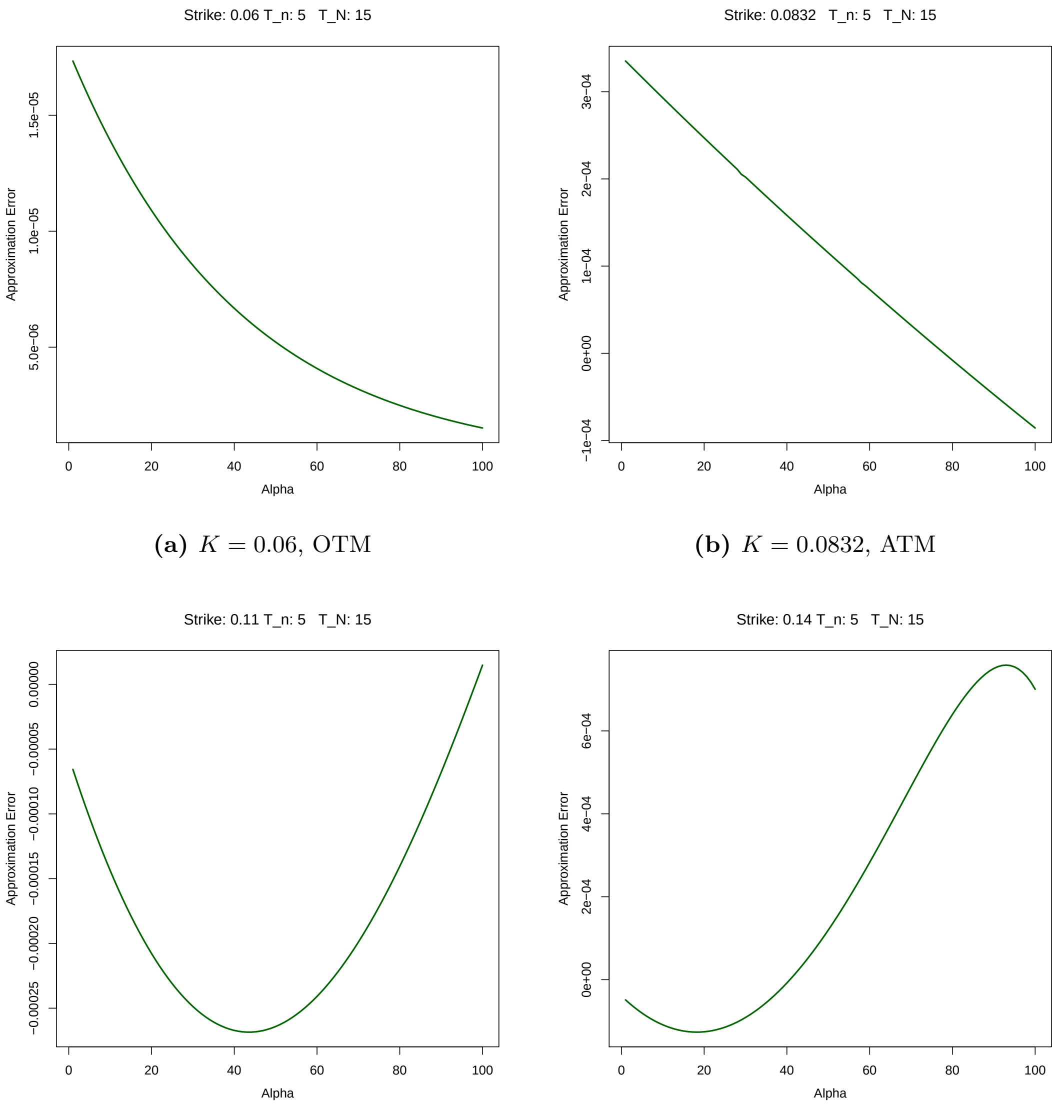
(a) $K = 0.02$, ITM

(b) $K = 0.05$, ITM

(c) $K = 0.0832$, ATM

(d) $K = 0.1$, OTM

Figure 13.8: *Approximation errors for payer swaption prices obtained from a Monte Carlo simulation and Fourier inversion. Monte Carlo prices are based on 10,000 simulations. α runs from 0.1 to 100. ATMF level at $K = 0.0832$.*



(a) $K = 0.06$, OTM

(b) $K = 0.0832$, ATM

(c) $K = 0.11$, ITM

(d) $K = 0.14$, ITM

Figure 13.9: *Approximation errors for receiver swaption prices obtained from a Monte Carlo simulation and Fourier inversion. Monte Carlo prices are based on 10,000 simulations.* α *runs from 0.1 to 100. ATMF level at* $K = 0.0832$*.*

13.4 Results

We now move on to present the swaption prices which, together with the default probabilities found in Chapter 12, will lay the foundation for CVA pricing.

The idea is to price swaptions using the two sets of short rate parameters estimated in Chapter 11 on various swaption structures with different strikes. The initial short rate level will be the observed three-month US Treasury Benchmark Bonds yield on the last day in each estimation period for the two sets of parameters. Dates, parameters and initial short rates are found in Table 13.1.

Date for $t = 0$	ψ	θ	σ	$r(0)$
02-09-2008	0.2592	0.0063	0.0840	0.0165
01-06-2011	0.6957	0.0097	0.1448	0.00048

Table 13.1: *Date of pricing along with the corresponding short rate parameters and initial short rate.*

The swaption structures and prices presented in Table 13.2 and 13.3 will be the ones used to price the CVA on a 5-year credit risky swap in the next chapter. Hence, according to Proposition 7.2 on page 61, the swaption prices are based on the appropriate forward swap rate and a strike corresponding to the fixed rate that renders a 5-year swap value zero at the time of pricing. Tables 13.2 and 13.3 contain prices evaluated using parameters and initial short rate according to 02-09-2008 and 01-06-2011 in Table 13.1, respectively. Notional is set to 1 in both tables. Prices for the same structures but with different strikes can be found in Appendix D.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	4	0.01108391	0.01196713	0.00599427
Payer	2	3	0.01034756	0.01196713	0.00508406
Payer	3	2	0.00973504	0.01196713	0.00277271
Payer	4	1	0.00922584	0.01196713	0.00196096
Receiver	1	4	0.01108391	0.01196713	0.00936047
Receiver	2	3	0.01034756	0.01196713	0.00968399
Receiver	3	2	0.00973504	0.01196713	0.00697943
Receiver	4	1	0.00922584	0.01196713	0.00452829

Table 13.2: *Swaption prices for different expiries and tenors using a strike corresponding to the fixed rate that renders a 5-year swap value zero at the time of pricing. Parameters and initial short rate according to 02-09-2008.*
Parameters: $\psi = 0.2592$, $\theta = 0.0063$, $\sigma = 0.0840$ *and* $r(0) = 0.0165$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	4	0.00807759	0.00705745	0.01145685
Payer	2	3	0.00865566	0.00705745	0.00774370
Payer	3	2	0.00898723	0.00705745	0.00565201
Payer	4	1	0.00918096	0.00705745	0.00356429
Receiver	1	4	0.00807759	0.00705745	0.00748159
Receiver	2	3	0.00865566	0.00705745	0.00308635
Receiver	3	2	0.00898723	0.00705745	0.00192073
Receiver	4	1	0.00918096	0.00705745	0.00152306

Table 13.3: *Swaption prices for different expiries and tenors using a strike corresponding to the fixed rate that renders a 5-year swap value zero at the time of pricing. Parameters and initial short rate according to 01-06-2011.*
Parameters: $\psi = 0.6957$, $\theta = 0.0097$, $\sigma = 0.1448$ *and* $r(0) = 0.00048$.

In Table 13.2 we see that, given the short rate parameters, the specific structures imply that the forward swap rate is decreasing in increasing expiries and decreasing tenors. Consequently, payer swaptions are OTM and receiver swaptions are ITM and receiver prices are higher than the corresponding payer prices. In Table 13.3, the almost exact opposite tendency is observed. Here the forward swap rate is increasing in increasing expiries and decreasing tenors making payer swaptions ITM and receiver swaptions OTM. This results in higher payer prices compared to the corresponding receiver prices. Overall, it seems that both payer and receiver swaption prices decline when expiry increases and tenor decreases regardless of moneyness.

In Table 13.4 and 13.5 we have listed the swaption structures along with the prices that we will use to price the CVA on a 10-year credit risky swap. Again, notional is set to 1 and the strike level corresponds to the fixed rate that renders a 10-year swap value zero at the time of pricing. Prices in Table 13.4 and 13.5 are evaluated using parameters and initial short rate according to 02-09-2008 and 01-06-2011 in Table 13.1, respectively. Prices for the same structures but with different strikes can be found in Appendix D.

Not surprisingly, each set of parameters implies the same forward rate structure for increasing expiries and decreasing tenors as in Table 13.2 and 13.3. Also, the same moneyness applies to payer and receiver swaptions given short rate parameters and strike. Thus, receiver prices are higher than the corresponding payer prices in Table 13.4 and lower in Table 13.5. However, prices seem a bit less predictable when comparing different swaption structures, especially when it comes to receiver swaptions. In Table 13.5 receiver swaption prices initially decline for increasing swaption maturities and decreasing swap maturities. However, this tendency is reversed as the structures have longer expiry than tenor. This ultimately means that a 1-year swaption on a 9-year swap has almost the same value as a 9-year swaption on a 1-year swap, even though both swaptions have different levels of moneyness. It is difficult to pinpoint exactly what causes this specific price pattern, since many different factors such as expiry, tenor, moneyness and the underlying parameters are in play. For payer swaptions we see the same pattern as in Table 13.2 and 13.3, where prices decline with increasing expiries and decreasing tenors.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	9	0.00913943	0.00979415	0.02317616
Payer	2	8	0.00860422	0.00979415	0.00644471
Payer	3	7	0.00816791	0.00979415	0.00423117
Payer	4	6	0.00781265	0.00979415	0.00578964
Payer	5	5	0.00752337	0.00979415	0.00669477
Payer	6	4	0.00728757	0.00979415	0.00424049
Payer	7	3	0.00709502	0.00979415	0.00343792
Payer	8	2	0.00693740	0.00979415	0.00237650
Payer	9	1	0.00680797	0.00979415	0.00132579
Receiver	1	9	0.00913943	0.00979415	0.02865414
Receiver	2	8	0.00860422	0.00979415	0.01529766
Receiver	3	7	0.00816791	0.00979415	0.01477525
Receiver	4	6	0.00781265	0.00979415	0.01674595
Receiver	5	5	0.00752337	0.00979415	0.01710038
Receiver	6	4	0.00728757	0.00979415	0.01340605
Receiver	7	3	0.00709502	0.00979415	0.01081317
Receiver	8	2	0.00693740	0.00979415	0.00755866
Receiver	9	1	0.00680797	0.00979415	0.00402387

Table 13.4: *Swaption prices for different expiries and tenors using a strike corresponding to the fixed rate that renders a 10-year swap value zero at the time of pricing. Parameters and initial short rate according to 02-09-2008.*
Parameters: $\psi = 0.2592$, $\theta = 0.0063$, $\sigma = 0.0840$ *and* $r(0) = 0.0165$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	9	0.00884053	0.00823972	0.02959479
Payer	2	8	0.00916161	0.00823972	0.02324693
Payer	3	7	0.00933355	0.00823972	0.02133001
Payer	4	6	0.00942623	0.00823972	0.01465500
Payer	5	5	0.00947668	0.00823972	0.01285245
Payer	6	4	0.00950449	0.00823972	0.01089490
Payer	7	3	0.00952007	0.00823972	0.00881937
Payer	8	2	0.00952895	0.00823972	0.00650328
Payer	9	1	0.00953412	0.00823972	0.00374704
Receiver	1	9	0.00884053	0.00823972	0.01431507
Receiver	2	8	0.00916161	0.00823972	0.00727240
Receiver	3	7	0.00933355	0.00823972	0.00627567
Receiver	4	6	0.00942623	0.00823972	0.00126930
Receiver	5	5	0.00947668	0.00823972	0.00151801
Receiver	6	4	0.00950449	0.00823972	0.00177036
Receiver	7	3	0.00952007	0.00823972	0.00196877
Receiver	8	2	0.00952895	0.00823972	0.00194512
Receiver	9	1	0.00953412	0.00823972	0.00147652

Table 13.5: *Swaption prices for different expiries and tenors using a strike corresponding to the fixed rate that renders a 10-year swap value zero at the time of pricing. Parameters and initial short rate according to 01-06-2011.*
Parameters: $\psi = 0.6957$, $\theta = 0.0097$, $\sigma = 0.1448$ *and* $r(0) = 0.00048$.

In Figure 13.10, we have depicted two integrands for different swaption structures based on the two sets of parameters. As the figure indicates, both integrands are similar in shape and are asymptotically converging to zero around some value of u in the interval 40-50. Similar convergence levels arise for the other swaption structures presented in the four tables above.

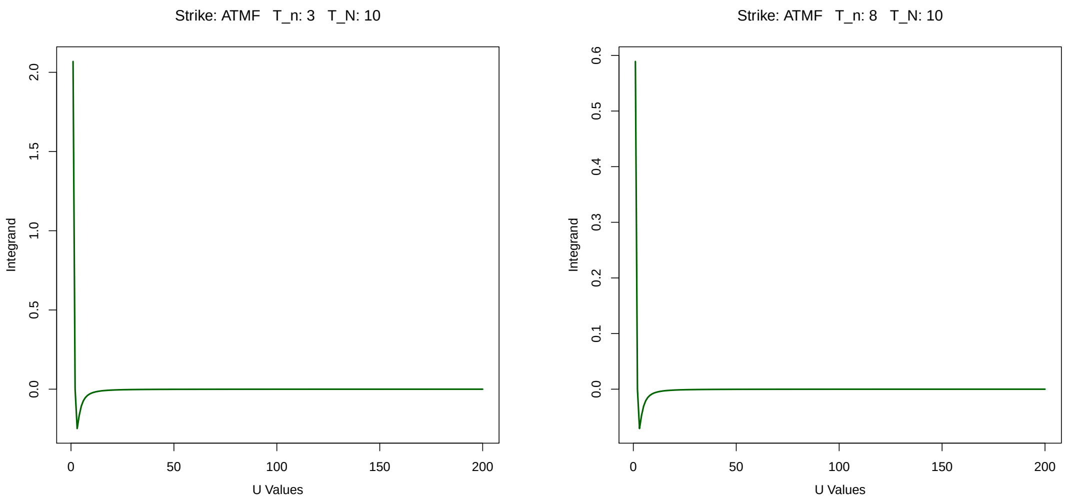


Figure 13.10: *Integrands for different ATM swaption structures.*
The left panel: $\psi = 0.2592$, $\theta = 0.0063$, $\sigma = 0.0840$ *and* $r(0) = 0.0165$.
The right panel: $\psi = 0.6957$, $\theta = 0.0097$, $\sigma = 0.1448$ *and* $r(0) = 0.00048$.

Chapter 14

CVA Results

Having derived implied default probabilities for our two fictional counterparties HSBC and Fiat and obtained swaption prices for various swaption structures, we are now ready to calculate CVAs on two credit risky swaps with maturity of five and ten years. The CVA calculations will be performed using the approximation of the analytical CVA formula stated in (7.15) on page 62 using the estimated 1-year implied modeled default probabilities presented in Table 12.4 on page 98 and the swaption prices in Table 13.2, 13.3, 13.4 and 13.5 on pages 110, 110, 111 and 112, respectively. Recovery is assumed constant and set to 40%. Both payer and receiver swaps will be considered.

Consequently, we place ourselves on the two different dates; the 2 September 2008 and 1 June 2011. The CVAs calculated on 2 September 2008 will be performed using 1-year implied default probabilities estimated in the period between 2007 and 2008 along with swaption prices based on the short rate parameters presented in Table 13.1 on page 110. Similarly, the CVAs calculated on 1 June 2011 will be performed using 1-year implied default probabilities estimated in the period between 2010 and 2011 along with swaption prices derived from the short rate parameters also presented in Table 13.1.

We consider a fictional scenario where a default free entity engages in swap contracts with the two defaultable counterparties HSBC and Fiat. These swap contracts will have maturities of five and ten years and we will consider the case where the default free entity takes both the payer and receiver side of the swap contracts. The goal is therefore to calculate the CVA that both counterparties will have to pay in order to compensate the default free entity for bearing all the default risk. The CVA can be represented as an upfront payment or be incorporated in the swap rate.

14.1 Preliminary Thoughts

Before presenting the results of our CVA calculations, some thoughts on what affects the size of the CVA might be insightful. As discussed in Section 7.2 and as (7.15) indicates, the CVA on swaps depends on the default term structure of the counterparty and the price of the swaptions that quantifies the exposure within the swap contract. Therefore, understanding the effect of these two components is essential in order to compare CVAs for different counterparties and swap contracts.

One seemingly straightforward consequence is that a higher probability of default must imply a higher CVA. This seems very intuitive and can also be confirmed quite easily by studying the approximative formula in (7.15). However as the formula indicates, this depends on the shape of the default term structure and the swap maturity. If a counterparty is considered risky in the sense that the probability of defaulting within, for instance, the next 20 years is large but small within the next five years, the CVA on a swap contract with a short maturity might be surprisingly small all else equal. Hence, the shape of the default term structure must be carefully examined in connection with the swap maturity when interpreting CVA charges.

The effect of swaption prices on CVA charges is less obvious. Of course, higher swaption prices imply higher CVAs and vice versa, but it is the term structure of the forward swap rates that determines the impact of the swaptions, since it determines if the swap is a liability or an asset. However, the effect of increasing expiries and decreasing tenors for the swaptions might reduce the impact they have.

14.2 Results

We are now ready to present the CVA calculations based on the setup described in the introduction to this chapter. The results are given in Table 14.1. The results for each counterparty are divided into date of pricing, maturity of the swap contract, and position in the swap.

Counterparty	Date	Swap Maturity	Swap Type	CVA
HSBC	02-09-2008	5Y	Payer	0.00012718
HSBC	02-09-2008	5Y	Receiver	0.00026186
HSBC	02-09-2008	10Y	Payer	0.00050620
HSBC	02-09-2008	10Y	Receiver	0.00129843
FIAT	02-09-2008	5Y	Payer	0.00020383
FIAT	02-09-2008	5Y	Receiver	0.00041890
FIAT	02-09-2008	10Y	Payer	0.00081646
FIAT	02-09-2008	10Y	Receiver	0.00208504
HSBC	01-06-2011	5Y	Payer	0.00019634
HSBC	01-06-2011	5Y	Receiver	0.00008928
HSBC	01-06-2011	10Y	Payer	0.00112999
HSBC	01-06-2011	10Y	Receiver	0.00028534
FIAT	01-06-2011	5Y	Payer	0.00182605
FIAT	01-06-2011	5Y	Receiver	0.00082502
FIAT	01-06-2011	10Y	Payer	0.00793644
FIAT	01-06-2011	10Y	Receiver	0.00214900

Table 14.1: *CVAs on 5-year and 10-year payer and receiver swaps with HSBC and Fiat as counterparties computed on 2 September 2008 and 1 June 2011.*

Considering the results in Table 14.1, we can make the following observations

- The CVA for Fiat is higher than for HSBC when comparing swaps with equal maturity computed on the same date. This makes perfect sense since the 1-year implied default probabilities in Table 12.4 on page 98 are higher for Fiat than for HSBC in both periods, whereas the swaption prices used are the same for both CVAs.
- 10-year swaps imply a higher CVA than the corresponding CVA on a 5-year swap for both HSBC and for Fiat. Again, this seems like a reasonable consequence since the CVA on

a 10-year swap contains the same first five 1-year implied default probabilities as the 5-year swap plus additional implied default probabilities associated with the longer maturity. Furthermore, the longer maturity implies more swaption prices.

- CVAs on receiver swaps are higher than CVAs on payer swaps when priced on 2 September 2008. Since payer and receiver swaps are priced using the same default probabilities, the difference in the size of CVAs lies in the swaption prices. By looking at Table 13.2 on page 110 and Table 13.4 on page 111 we see that prices for receivers are higher than for payers for given structures. Hence, for a given swap maturity we observe a higher CVA for receiver swaps.
  The exact opposite situation applies to CVA prices on 1 June 2011. Here CVAs on payer swaps are higher than on receiver swaps. Again, the reason for this consistent difference can be found by observing that prices for payers are higher than for receivers in Table 13.3 on page 110 and Table 13.5 on page 112.
- For Fiat, the CVAs computed on 1 June 2011 are higher compared to the CVAs computed on 2 September 2008. This is mainly due to the higher implied default probabilities derived from the CDS term structure observed between 2010-2011.

These observations are made simply by comparing the CVAs in Table 14.1. Fortunately, the CVAs seem to make sense in that they can be accounted for in terms of default probabilities and swaption prices.

Instead of considering the CVA as an upfront payment, we will now try to incorporate it in the swap rate in order to get a better sense of what the CVA actually entails for a risky swap. To do this, we can use the formula in Proposition 7.1 on page 59, which states the connection between the price of a risky swap and a risk free swap through the CVA. Thus, applying the formula in Proposition 7.1 combined with the price of a risk free payer swap derived in (3.13) on page 16, we can express the price of a risky payer swap denoted $\tilde{\pi}^{pay}(t,T)$ at time t as

$$\tilde{\pi}^{pay}(t,T) = D(t,T_n) - D(t,T_N) - \kappa P_{n+1,N}(t) - \text{CVA}(t). \tag{14.1}$$

Assuming that the forward swap rate for a risky payer swap, $\tilde{\kappa}^{pay}$, is chosen so the value of the risky payer swap has value zero at the time of inception, we can obtain the forward swap rate by rearranging (14.1) such that

$$\tilde{\kappa}^{pay} = \frac{D(0,T_n) - D(0,T_N) - \text{CVA}(0)}{P_{n+1,N}(0)} = y_{n,N}(t) - \frac{\text{CVA}(0)}{P_{n+1,N}(0)}. \tag{14.2}$$

By performing a similar derivation for the equivalent receiver swap we end up with the following expression for the forward swap rate

$$\tilde{\kappa}^{rec} = \frac{D(0,T_n) - D(0,T_N) + \text{CVA}(0)}{P_{n+1,N}(0)} = y_{n,N}(t) + \frac{\text{CVA}(0)}{P_{n+1,N}(0)}. \tag{14.3}$$

Consequently, the forward swap rate derived at inception differs for risky payer and receiver swaps unlike in the default free case. Compared to the default free case, we now recover a forward swap rate which is lower for payer swaps and higher for receiver swaps. This seems reasonable seeing that the default free entity will require a higher (lower) fixed rate as a premium (discount) to bear the default risk when engaging in a receiver (payer) swap with a defaultable counterparty.

As an example, let us consider the 10-year receiver swap with Fiat as counterparty using the

CVA computed on 1 June 2011. The risk free swap rate was found to be 0.00823972, which means that

$$\tilde{\kappa}^{rec} = 0.00823972 + \frac{0.00214900}{P_{1,10}(0)} = 0.008463506. \tag{14.4}$$

Hence, we can confirm an increase in the swap rate for the risky receiver swap. Including the CVA in the swap rate appears to be a more intuitive way to illustrate the effect of counterparty credit risk in swap valuation, since it preserves the original swap characteristics such as zero value at inception.

Chapter 15

Discussion

The CVA prices presented in the last chapter were derived from a swaption model framework and an intensity framework. Both models were combined in one single formula that quantified the counterparty risk on a swap contract in terms of the CVA. In order to recover the CVA, we had to assume that the short rate and the intensity were independent. The assumption of independence is one of the topics we wish to discuss in this chapter.

In our implementation of the swaption and intensity model we employed two different estimation procedures. The parameters for the underlying short rate in the swaption framework were found using maximum likelihood whereas the parameters for the intensity model were found by minimizing the RMSE. Both estimation procedures were justified, yet still subject to estimation errors and have therefore room for improvement. Hence, alternative estimation schemes will also be a topic of this chapter.

15.1 Independence and Wrong-way Risk

The key step in the derivation of the CVA formula for swaps was to assume independence between default probabilities, interest rates and the recovery rate, which we assumed constant. This assumption meant that we could compute CVAs for swaps as a sum of independent terms, where each term consisted of a swaption price multiplied by a default probability. The recovery rate could then be multiplied as a factor to the sum. Allowing this simplification implied that we could focus on pricing swaptions and deriving default probabilities separately while not having to worry about the interaction between the two, thus reducing the complexity of the computation greatly.

The interpretation of independence between interest rates and default probabilities basically means that the level of interest rates has no effect on the credit quality of the counterparty. Whether this is actually a realistic assumption obviously depends on the counterparty in question. But one could argue that the financial health of firms usually relies on interest rates, since they determine the price of borrowing money, which all institutions depend on. For instance, the credit quality of highly leveraged institutions may deteriorate dramatically when interest rates rise making loans more expensive. This could translate into a higher probability of default. Also, declining interest rate levels usually imply a downturn in the economy, which could lead to a large number of corporate defaults.

This discussion leads us to the concept of *Wrong-way risk*. Wrong-way risk is a term used to describe an unfavourable dependence between exposure and the credit quality of the counterparty. Unfavorable meaning that there is a positive relationship between exposure and the default probability of the counterparty. The presence of wrong-way risk in derivatives can potentially increase the counterparty credit risk substantially even though the dependence seems negligible. Connected to the concept of wrong-way risk is the concept of *right-way risk*. Right-way risk is the opposite of wrong-way risk in that it indicates a beneficial relationship between exposure and default probability, which leads to a lower counterparty risk. So, whereas wrong-way risk is best avoided when dealing with derivatives, right-way risk is a positive feature.

Both wrong-way and right-way risks can be present in swap contracts. Consider the scenario where the default free entity engages in a receiver swap contract with a defaultable counterparty. The economy goes into a recession which leads to lower interest rates and a higher level of corporate defaults due to falling profits from lower consumption. In this scenario, the exposure in the receiver swap contract rises due to the falling interest rates. However, the higher number of defaults due to the recession increases the default risk of the counterparty. This suggests the presence of wrong-way risk in the receiver swap. On the other hand, a payer swap would, in the same scenario, be exposed to right-way risk, since the swap exposure declines along with the falling interest rates.

The fact that swaps could be exposed to wrong-way and right-way risk underlines the fact that CVAs on swaps should take this into account in order to accurately price the counterparty risk. Wrong-way risk will increase the CVA, while right-way risk will decrease the CVA. The most important of these two risk concepts seem to be wrong-way risk due to the negative consequences associated with this kind of risk. Generally, there are two ways of incorporating wrong-way risk into the computation of CVAs on swaps:

- Quantify the economic relationship between interest rates and the default of the counterparty. This may, however, be very difficult and might also increase the complexity of the CVA computation significantly.
- Take a simple and more ad hoc approach, where one for example adjusts the default probabilities upwards in order to account for the wrong-way risk.

Both methods could be applied in the CVA formula for swaps we derived in Section 7.2, however, the first approach would require some additional mathematical modeling. In the next section we will make a suggestion as to how this can be done.

15.1.1 Modelling Wrong-way Risk by Correlation

One way we could model wrong-way risk would be to abandon the independence assumption in the CVA formula in Proposition 7.2 on page 61. Instead, we could compute the CVA using the expectation in Proposition 7.1 on page 59. As a consequence, we would have to model the correlation between the short rate and the default intensity. In Brigo et al. (2010) the short rate and the intensity processes are correlated since they assume that the driving Brownian motions in both process are instantaneously correlated. In our case, this would apply for the two one-factor CIR processes. This would introduce a correlation parameter in the estimation process resulting in a more complex estimation. Pricing the CVA would then have to be performed using numerical techniques such as Monte Carlo simulation.

Using correlation as a measurement for wrong-way risk can turn out to be dangerous since correlation only measures the linear relationship between the short rate and the intensity. Consider the case where a small movement in interest rates has no real affect on the credit quality of the counterparty, but where a large movement has a drastic affect. This relationship would not be captured by the correlation. Hence, correlation is not necessarily the best proxy for wrong-way risk. However, it would probably result in a more realistic CVA size but at the same time also make the computational aspects more complex and thus more time consuming.

15.2 Intensity Estimation

Generally, choosing a one-factor CIR model for the intensity to stipulate the default dynamics seems to work well when minimizing the RMSE between the observed CDS quotes and fitted CDS quotes. This can be confirmed by looking at Table 12.3 on page 97. This suggests that expanding the intensity dynamics to a multi factor setup is unnecessary. The implied default term structure derived from taking the mean of the parameter estimates also generally seems to correspond well to the observed CDS quotes. However, in the case of Fiat the implied default term structure, given the mean of the parameter estimates, turned out to be less convincing. In Figure 12.3 on page 98 we saw that when positioning ourselves in June 2011 the accumulated default probabilities for Fiat were extremely high, predicting a probability of default of up to 80% within the next nine years and a probability of almost 100% within the next 20 years. This seems rather unrealistic when comparing the default probabilities with the observed CDS quotes in that period. However, the shape of the term structure was a consequence of the high estimates for the volatility and mean reversion parameter. Thus, one could argue that a better estimation procedure might result in parameter estimates that allow for an implied default term structure more in line with the observed CDS quotes.

One way of obtaining a more realistic default term structure could be to base the estimation on a larger set of historically observed CDS term structures. However, expanding the set of observations could affect the parameter estimates negatively, since averaging over a large set of data might increase the standard deviation, thereby making the parameters less statistically significant. This would probably be the case for HSBC since the volatility on its CDS term structures during the last four years has been relatively high cf. Figure 12.1 on page 95. A natural extension to an estimation procedure based on a large set of CDS term structures would be to place greater emphasis on observations that are considered more important for the estimation. For instance, placing greater emphasis on recent observations and less on observations in the past. This could be achieved by applying an *Exponential Weighted Moving Average* approach where the weighting decreases exponentially, so that for each older observation the weight decreases exponentially never reaching zero. This would be a more intuitive way to handle a large set of data.

Another approach could be to fix one global set of parameters to fit all the observed term structures instead of fitting a set of parameters for each observed term structure. Thus leaving the initial intensity $\lambda(0)$ to act as the latent variable to be fitted for each observed term structure. In this way, we could avoid the possibility of averaging over a set of parameters with high variance. This estimation procedure would in principle result in a perfect fit for the term structures, since $\lambda(0)$ would always equal the residual. The variation in the λs would then give an indication of the performance of the fit given the set of parameters. It is difficult to say whether this would result in a more realistic implied default structure.

Instead of deriving default probabilities using an intensity framework, we could have chosen a more straightforward approach where implied default probabilities are derived directly from observed CDS quotes by a simple bootstrapping procedure. Since this method does not require any model specification, one could argue that these probabilities come closer to the ”true” default probabilities. This method does, however, have certain limitations. The most important of these being that the method, unlike the intensity approach, lacks the ability to predict future default probabilities beyond the longest CDS maturity observed in the market. Another limitation is that one has to incorporate an interpolation scheme in order to obtain default probabilities in between yearly observations. However, a bootstrapping procedure would probably have resulted in a more realistic 10-year implied default term structure for Fiat.

15.3 Swaption Estimation

The estimation procedure we used for the swaption model framework was to fit parameters for the one-factor CIR short rate model based on the observed yield of the three-month US Treasury Benchmark Bond using maximum likelihood. This procedure turned out to be very stable and gave parameter estimates that seemed reasonable. A more natural approach would have been to estimate the parameters by minimizing the RMSE between model prices and observed swaption prices obtained from the market. This would ensure that the model prices are consistent with the market. In this way, the short rate parameters would be estimated implicitly in the model, thereby avoiding the decision of choosing a relevant risk free interest rate proxy.

One problem with this procedure could turn out to be our choice of short rate model. As we argued in the beginning of this thesis paper, the choice of modeling the short rate according to a simple one-factor CIR model was decided upon to illustrate how pricing swaptions based on affine models could be done. On the other hand, we also showed that the setup was flexible and could easily be expanded allowing for a more complicated short rate model with several factors. One could imagine that the one-factor CIR model would lack the flexibility needed to fit swaption prices due to the relatively small amount of parameters. Therefore, a multi-factor short rate model might be more appropriate even though this would result in a more comprehensive and complicated estimation process. An argument against this estimation procedure is that we would obtain parameters fitted specifically to the swaption market. Hence, using the same parameters for deriving implied default probabilities from the CDS market might not be appropriate, since the parameters would have a bias towards the swaption market.

Another approach, more in line with the one we chose, would be to extract ZCB prices from the market to create a sort of empirical ZCB term structure. We could then choose the short rate parameters so that the theoretical ZCB term structure fits the empirical term structure. In this case, we would still need to specify a risk free term structure. There is, however, no reason why this method should result in more realistic swaption prices.

Chapter 16

Conclusions

In this master's thesis we have treated the subject of how to price counterparty credit risk in interest rate swaps. For a generic risky claim, we proved that it could be divided into a risk free price minus a CVA. Furthermore, we proved that the price of a credit risky swap, under independence assumptions, equaled an infinite sum of swaption prices multiplied by marginal default probabilities. Our goal then became to price swaptions and derive default probabilities. In doing so, we established a general asset pricing framework and showed how to expand it to include Fourier inversion techniques so that the swaption pricing model of Pelsser and Schrager (2006) could be applied. Furthermore, it was necessary to investigate the class of affine term structure models that the swaption pricing model was built within. The swaption model had its immediate advantages of stipulating the dynamics of possibly any affine term structure model. Moreover, the affine framework ensured semi-analytical solutions to zero coupon bonds and the characteristic function of the future interest rate distribution.

In the pursuit of default probabilities, we examined the intensity models proposed in Lando (1998). By adopting comparable affine assumptions in terms of the intensity process, we were able to reuse the tractable properties of the affine class so that closed form solutions were obtainable to the survival and default probabilities.

By specifying a one-factor CIR model for both the short interest rate and the default intensity, we were able to explore the models in detail. For the swaption pricing model, we had to solve a set of complex differential equations in order to obtain the mentioned characteristic function. In order to conduct this operation, we found it necessary to solve only for integer values of the characteristic function and then incorporate a cubic spline. By using Monte Carlo methods, we were able to show that this approach had practically no effect on the characteristic function. Concerning the LVM approximation of the swap rate dynamics under the swap measure, we tested against the non-approximated dynamics by using Monte Carlo methods and found that the LVM assumption worked out very satisfyingly. In terms of actual swaption prices, we found that they were somewhat dependent on the dampening-factor α, and which α to choose seemed ambiguous when comparing to the equivalent Monte Carlo prices, which of course is also subject to estimation errors. Lastly, for the swaption model we computed both payer and receiver swaption prices for different swaption structures.

Considering the implementation of the intensity model, the affine assumptions made usage more straightforward compared to the swaption pricing model. Estimations naturally seemed better given less fluctuation in observed CDS spreads. Standard deviations of estimates clearly became undesirably high in certain periods of time. In particular, default estimates of Fiat using 2010-

2011 data seemed overestimated. Perhaps a different estimation routine or a more complex model would be able to achieve a more realistic forecast.

Studying the interplay between the two models, we presented CVA prices for different swap specifics using models estimated over different time periods for two different companies acting as counterparties in a swap contract. In spite of the mentioned drawbacks, the interplay of our proposed models still exhibited CVA prices that seemed reasonable in terms of size, given that the overall default risk level has increased since 1994. Furthermore, the CVAs exhibited a reasonable quantitative relationship between each other. As expected, we were prone to find CVAs that were characterized by a small risk premium that became significant due to the typically large size of swap notional in a trade and due to the immense size of the swap market in general. Ultimately, we showed explicitly how the CVA could be directly incorporated in the swap rate, allowing the swap value to remain zero upon inception so that an upfront payment could be avoided.

References

Andersen, L. and Piterbarg, V. (2007). *Interest Rate Modeling. Volume 1: Foundations and Vanilla Models.* Atlantic Financial Press, first edition.

Berndt, A., Jarrow, R., and Kang C. (2006). Restructuring risk in credit default swaps: An empirical analysis. *Carnegie Mellon University.*

Björk, T. (2009). *Arbitrage Theory in Continuous Time.* Oxford University Press, third edition.

Boettcher, B. (2010). Towards active management of counterparty credit risk with cva. *Algorithmics.*

Brace, A., Dunn, T., and Barton, G. (2001). Towards a central interest rate model. *Option Pricing, Interest Rates and Risk Management: Handbooks in Mathematical Finance, Cambridge University Press.*

Brigo, D. and Capponi, A. (2009). Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps. *Risk Books, London.*

Brigo, D., Dalessandro, A., Neugebauer, M., and Triki, F. (2007). A stochastic processes toolkit for risk management. *Fitch Solutions.*

Brigo, D. and Masetti, M. (2005). Risk neutral pricing of counterparty risk. *Risk Books, London. (Forthcoming).*

Brigo, D., Pallavicini, A., and Papatheodorou, V. (2010). Bilateral counterparty risk valuation for interest-rate products: impact of volatilities and correlations. *Risk Books, London.*

Canabarro, E. and Duffie, D. (2003). Measuring and marking counterparty risk. *Euro Money Books.*

Carr, P. and Madan, D. (1998). Option valuation using the fast Fourier transform. *Journal of Computational Finance.*

Cooper, I., A. and Mello, S., A. (1991). Counterpart valuation adjustment. *THE JOURNAL OF FINANCE, VOL. XLVI.*

Cox, J., Ingersoll, J., and Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica, 53(2) (1985):385-407.*

Dai, Q. and Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance, Vol. LV, No. 5.*

Dalskov, L. M. (2007). Fixed income derivatives lecture notes. *Independent Working Paper.*

Duffee, G. (1998). Forecasting future interest rates: Are affine models failures? *Working Paper, Federal Reserve Board.*

Duffie, D. and Huang, M. (1996). Swap rates and credit quality. *THE JOURNAL OF FINANCE, VOL. LI, NO. 3.*

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance, Vol. 6. No. 4 (October l996), 379-406.*

Duffie, D., Pan J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometria, Vol 68, No. 6, 1343-1376.*

Dufresne, P. and Goldstein, R. (2002). Pricing Swaptions in an affine framework. *Journal of Derivatives 10, 9-26.*

Felthütter, P. (2008). An empirical investigation of an intensity-based model for pricing cdo tranches. *Department of Finance.*

Glasserman, P. (2003). *Monte Carlo Methods in Financial Enigineering.* Springer.

Gregory, J. (2010). *Counterparty Credit Risk.* Wiley Finance, first edition.

Grimmett, G. and Stirzaker, D. (2001). *Probability Theory and Random Processes.* Oxford, third edition.

Haase, J., Ilg, M., and Werner, R. (2010). Model-free bounds on bilateral counterparty valuation. *MPRA Paper No. 24796.*

Heston, S. (1993). A closed-form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies 6(2), 327-343.*

Huge, B. and Lando, D. (1999). Swap pricing with two-sided default risk in a rating-based model. *European Finance Review 3: 239–268, Kluwer Academic Publishers.*

Hunt, P. and Kennedy, J. (2000). *Financial Derivatives In Theory and Practice.* Wiley Series In Probability and Statistics, first edition.

ISDA (2010). Isda margin survey 2010. *See ISDA's homepage.*

Kladivko, K. (2004). Maximum likelihood estimation of the cox-ingersoll-ross process: The matlab implementation. *Department of Statistics and Probability Calculus, University of Economics, Prague and Debt Management Department, Ministry of Finance of the Czech Republic.*

Øksendal, B. (1998). *Stochastic Differential Equations.* Springer-Verlag, fifth edition.

Lando, D. (1998). On cox processes and credit risky securities. *Kluwer Academic Publishers, Boston.*

Lando, D. (2004). *Credit Risk Modeling: Theory and applications.* Princeton University Press, first edition.

Lee, R. (2004). Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance.*

Longstaff, F., Mithal, D., and Neis, E. (2004). Corporate yield spreads: Default risk or liquidity? New evidence from the credit-default swap market. *NBER Working Paper No. w10418.*

Moody's (2008). Corporate default and recovery rates, 1920-2007. *Moodys.com*.

Mortensen, A. (2006). Semi-analytical valuation of basket credit derivatives in intensity-based models. *Independent Working Paper*.

Mozumdar, A. (1999). Default risk of interest rate swaps: Theory and evidence. *Dept of Finance, Pamplin College of Business*.

Munk, C. (1999). Stochastic duration and fast coupon bond option pricing in multi-factor models. *Review of Derivatives Research, 3, 157-181*.

Munk, C. (2005). *Fixed Income*. Independent Working Paper.

Nowman, K. (1997). Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *American Finance Association*.

Pelsser, A. and Schrager, D. (2006). Pricing swaptions and coupon bond options in affine term structure models. *Mathematical Finance, Vol. 16, No.4, 673-694*.

Poulsen, R. (1999). Numeraireskift. *Independent Working Paper*.

Pykhtin, M. and Zhu, S. (2007). A guide to modelling counterparty credit risk. *GLOBAL ASSOCIATION OF RISK PROFESSIONALS*.

Schmelzle, M. (2010). Option pricing formulae using fourier transform: Theory and application. *Unknown*.

Shahram, A., Ding, J., Whitehead, P., and Laudicina, L. (2010). Counterparty valuation adjustment. *Unknown publisher*.

Shreve, S. (2004). *Stochastic Calculus for Finance II: Continuous-Time models*. Springer Finance, first edition.

Singleton, K. and Umantsev, L. (2002a). Pricing coupon bond options and swaptions in affine term structure models. *Math. Finance 12, 427–446*.

Singleton, K. and Umantsev, L. (2002b). Pricing coupon bond options and swaptions in affine term structure models. *Mathematical Finance 12, 427-446*.

Sorensen, E., H. and Bollier, T., F. (1994). Default risk of interest rate swaps: Theory and evidence. *Financial Analysts Journal, Vol. 50, No. 3 (May - Jun., 1994), pp. 23-33*.

Stein, H. and Lee, K. P. (2010). Counterparty valuation adjustments. *The Handbook of Credit Derivatives*.

Vasicek, O. (1977). An equilibrium characterisation of the term structure. *J. Financ. Econ. 5, 177–188*.

Whetten, M. (2004). Credit default swap primer. *Nomura Fixed Income Research*.

Yi, C. (2010). Dangerous knowledge: Credit value adjustment with credit triggers. *RBC Capital Market*.

Appendices

Appendix A

Dynamics of $\frac{1}{M(t)}$

The dynamics of the money market account is given by

$$dM(t) = r(t)M(t)dt.$$

By an application of Itô's formula, the dynamics of $\frac{1}{M(t)}$ is found as

$$\begin{aligned} d\left(\frac{1}{M(t)}\right) &= -\frac{1}{M(t)^2}dM(t) + \frac{1}{M(t)^3}(dM(t))^2 \\ &= -\frac{1}{M(t)^2}r(t)M(t)dt + \frac{1}{M(t)^3}r(t)^2M(t)^2(dt)^2 \\ &= -r(t)\frac{1}{M(t)}dt. \end{aligned}$$

Appendix B

Derivation of Equation (8.7) on page 65

Since $dP_{n+1,N}(t) = \sum_{i=n+1}^{N} \Delta dD(t,T_i)$ and $d\left(\frac{1}{M(t)}\right) = -r(t)\frac{1}{M(t)}dt$[1], an application of Itô's product rule yields

$$
\begin{aligned}
d\left(\frac{P_{n+1,N}(t)}{M(t)}\right) &= -\sum_{i=n+1}^{N} \Delta D(t,T_i) r(t) \frac{1}{M(t)} dt + \frac{1}{M(t)} \sum_{i=n+1}^{N} \Delta dD(t,T_i) \\
&\quad - r(t)\frac{1}{M(t)} dt \sum_{i=n+1}^{N} \Delta dD(t,T_i) \\
&= \frac{1}{M(t)} \Bigg(-\sum_{i=n+1}^{N} \Delta D(t,T_i) r(t) dt \\
&\quad + \sum_{i=n+1}^{N} \Delta \left(r(t) D(t,T_i) dt - B(t,T_i)^\intercal D(t,T_i) \Sigma \sqrt{V(t)} dW^{\mathbb{Q}}(t) \right) \\
&\quad - r(t) dt \sum_{i=n+1}^{N} \Delta \left(r(t) D(t,T_i) dt - B(t,T_i)^\intercal D(t,T_i) \Sigma \sqrt{V(t)} dW^{\mathbb{Q}}(t) \right) \Bigg) \\
&= \frac{1}{M(t)} \Bigg(-\sum_{i=n+1}^{N} \Delta D(t,T_i) r(t) dt + \sum_{i=n+1}^{N} \Delta r(t) D(t,T_i) dt \\
&\quad - \sum_{i=n+1}^{N} \Delta B(t,T_i)^\intercal D(t,T_i) \Sigma \sqrt{V(t)} dW^{\mathbb{Q}}(t) \Bigg) \\
&= -\sum_{i=n+1}^{N} \Delta B(t,T_i)^\intercal \frac{D(t,T_i)}{M(t)} \Sigma \sqrt{V(t)} dW^{\mathbb{Q}}(t) \\
&= -\sum_{i=n+1}^{N} \left(\Delta B(t,T_i)^\intercal \frac{D(t,T_i)}{P_{n+1,N}(t)} \Sigma \sqrt{V(t)} \right) \frac{P_{n+1,N}(t)}{M(t)} dW^{\mathbb{Q}}.
\end{aligned}
$$

[1]See Appendix A for a derivation.

Appendix C

Initial OLS Estimates for ψ and θ

The initial OLS estimates achieved by solving (11.5) on page 89 are given as

$$\hat{\psi} = \frac{N^2 - 2N + 1 + \sum_{i=1}^{N-1} r(t_{i+1}) \sum_{i=1}^{N-1} \frac{1}{r(t_i)} - \sum_{i=1}^{N-1} r(t_i) \sum_{i=1}^{N-1} \frac{1}{r(t_i)} - (N-1) \sum_{i=1}^{N-1} \frac{r(t_{i+1})}{r(t_i)}}{\left(N^2 - 2N + 1 - \sum_{i=1}^{N-1} r(t_i) \sum_{i=1}^{N-1} \frac{1}{r(t_i)}\right) \Delta t},$$

$$\hat{\theta} = \frac{(N-1) \sum_{i=1}^{N-1} r(t_{i+1}) - \sum_{i=1}^{N-1} \frac{r(t_{i+1})}{r(t_i)} \sum_{i=1}^{N-1} r(t_i)}{N^2 - 2N + 1 + \sum_{i=1}^{N-1} r(t_{i+1}) \sum_{i=1}^{N-1} \frac{1}{r(t_i)} - \sum_{i=1}^{N-1} r(t_i) \sum_{i=1}^{N-1} \frac{1}{r(t_i)} - (N-1) \sum_{i=1}^{N-1} \frac{r(t_{i+1})}{r(t_i)}}.$$

Appendix D

Swaption Prices

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	4	0.01108391	0.008	0.01572655
Payer	1	4	0.01108391	0.010	0.00986714
Payer	1	4	0.01108391	0.012	0.00595074
Payer	1	4	0.01108391	0.014	0.00408964
Payer	2	3	0.01034756	0.008	0.01115864
Payer	2	3	0.01034756	0.010	0.00746831
Payer	2	3	0.01034756	0.012	0.00505427
Payer	2	3	0.01034756	0.014	0.00369796
Payer	3	2	0.00973504	0.008	0.00631092
Payer	3	2	0.00973504	0.010	0.00410456
Payer	3	2	0.00973504	0.012	0.00275635
Payer	3	2	0.00973504	0.014	0.00198192
Payer	4	1	0.00922584	0.008	0.00353003
Payer	4	1	0.00922584	0.010	0.00254452
Payer	4	1	0.00922584	0.012	0.00195343
Payer	4	1	0.00922584	0.014	0.00155515
Receiver	1	4	0.01108391	0.008	0.00396856
Receiver	1	4	0.01108391	0.010	0.00573530
Receiver	1	4	0.01108391	0.012	0.00944223
Receiver	1	4	0.01108391	0.014	0.01520836
Receiver	2	3	0.01034756	0.008	0.00448857
Receiver	2	3	0.01034756	0.010	0.00648033
Receiver	2	3	0.01034756	0.012	0.00974760
Receiver	2	3	0.01034756	0.014	0.01407580
Receiver	3	2	0.00973504	0.008	0.00303934
Receiver	3	2	0.00973504	0.010	0.00460292
Receiver	3	2	0.00973504	0.012	0.00702505
Receiver	3	2	0.00973504	0.014	0.01002328
Receiver	4	1	0.00922584	0.008	0.00238093
Receiver	4	1	0.00922584	0.010	0.00326877
Receiver	4	1	0.00922584	0.012	0.00455157
Receiver	4	1	0.00922584	0.014	0.00602809

Table D.1: *Swaption prices for different expiries, tenors and strikes. Parameters and initial short rate according to 02-09-2008.*
Parameters: $\psi = 0.2592$, $\theta = 0.0063$, $\sigma = 0.0840$ *and* $r(0) = 0.0165$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	4	0.00807759	0.006	0.01428714
Payer	1	4	0.00807759	0.008	0.00936138
Payer	1	4	0.00807759	0.010	0.00635672
Payer	1	4	0.00807759	0.012	0.00474148
Payer	2	3	0.00865566	0.006	0.01040495
Payer	2	3	0.00865566	0.008	0.00580401
Payer	2	3	0.00865566	0.010	0.00326116
Payer	2	3	0.00865566	0.012	0.00260198
Payer	3	2	0.00898723	0.006	0.00731194
Payer	3	2	0.00898723	0.008	0.00436170
Payer	3	2	0.00898723	0.010	0.00243594
Payer	3	2	0.00898723	0.012	0.00152823
Payer	4	1	0.00918096	0.006	0.00436950
Payer	4	1	0.00918096	0.008	0.00291530
Payer	4	1	0.00918096	0.010	0.00190495
Payer	4	1	0.00918096	0.012	0.00137854
Receiver	1	4	0.00807759	0.006	0.00619253
Receiver	1	4	0.00807759	0.008	0.00905704
Receiver	1	4	0.00807759	0.010	0.01384238
Receiver	1	4	0.00807759	0.012	0.02002184
Receiver	2	3	0.00865566	0.006	0.00266517
Receiver	2	3	0.00865566	0.008	0.00389281
Receiver	2	3	0.00865566	0.010	0.00717607
Receiver	2	3	0.00865566	0.012	0.01234763
Receiver	3	2	0.00898723	0.006	0.00153519
Receiver	3	2	0.00898723	0.008	0.00245292
Receiver	3	2	0.00898723	0.010	0.00439336
Receiver	3	2	0.00898723	0.012	0.00735366
Receiver	4	1	0.00918096	0.006	0.00131160
Receiver	4	1	0.00918096	0.008	0.00178000
Receiver	4	1	0.00918096	0.010	0.00269139
Receiver	4	1	0.00918096	0.012	0.00408738

Table D.2: *Swaption prices for different expiries, tenors and strikes. Parameters and initial short rate according to 01-06-2011.*
Parameters: $\psi = 0.6957$, $\theta = 0.0097$, $\sigma = 0.1448$ *and* $r(0) = 0.00048$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	9	0,00913943	0.007	0.03679131
Payer	1	9	0,00913943	0.009	0.02629927
Payer	1	9	0,00913943	0.011	0.01954531
Payer	1	9	0,00913943	0.013	0.01544135
Payer	2	8	0.00860422	0.007	0.01789134
Payer	2	8	0.00860422	0.009	0.00816675
Payer	2	8	0.00860422	0.011	0.00554047
Payer	2	8	0.00860422	0.013	0.00486095
Payer	3	7	0.00816791	0.007	0.01227909
Payer	3	7	0.00816791	0.009	0.00523414
Payer	3	7	0.00816791	0.011	0.00382243
Payer	3	7	0.00816791	0.013	0.00324209
Payer	4	6	0.00781265	0.007	0.01165067
Payer	4	6	0.00781265	0.009	0.00658539
Payer	4	6	0.00781265	0.011	0.00511319
Payer	4	6	0.00781265	0.013	0.00351797
Payer	5	5	0.00752337	0.007	0.01184249
Payer	5	5	0.00752337	0.009	0.00777077
Payer	5	5	0.00752337	0.011	0.00545813
Payer	5	5	0.00752337	0.013	0.00400300
Payer	6	4	0.00728757	0.007	0.00729544
Payer	6	4	0.00728757	0.009	0.00468087
Payer	6	4	0.00728757	0.011	0.00369795
Payer	6	4	0.00728757	0.013	0.00247859
Payer	7	3	0.00709502	0.007	0.00529525
Payer	7	3	0.00709502	0.009	0.00364793
Payer	7	3	0.00709502	0.011	0.00312946
Payer	7	3	0.00709502	0.013	0.00216330
Payer	8	2	0.00693740	0.007	0.00390761
Payer	8	2	0.00693740	0.009	0.00265519
Payer	8	2	0.00693740	0.011	0.00205220
Payer	8	2	0.00693740	0.013	0.00157291
Payer	9	1	0.00680797	0.007	0.00206548
Payer	9	1	0.00680797	0.009	0.00146415
Payer	9	1	0.00680797	0.011	0.00115831
Payer	9	1	0.00680797	0.013	0.00090618

Table D.3: *Payer swaption prices for different expiries, tenors and strikes. Parameters and initial short rate according to 02-09-2008.*
Parameters: $\psi = 0.2592$, $\theta = 0.0063$, $\sigma = 0.0840$ *and* $r(0) = 0.0165$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Receiver	1	9	0,00913943	0.007	0.01885548
Receiver	1	9	0,00913943	0.009	0.02512333
Receiver	1	9	0,00913943	0.011	0.03512834
Receiver	1	9	0,00913943	0.013	0.04779187
Receiver	2	8	0.00860422	0.007	0.00595418
Receiver	2	8	0.00860422	0.009	0.01111071
Receiver	2	8	0.00860422	0.011	0.02336973
Receiver	2	8	0.00860422	0.013	0.03758080
Receiver	3	7	0.00816791	0.007	0.00470627
Receiver	3	7	0.00816791	0.009	0.01062803
Receiver	3	7	0.00816791	0.011	0.02219064
Receiver	3	7	0.00816791	0.013	0.03458510
Receiver	4	6	0.00781265	0.007	0.00715696
Receiver	4	6	0.00781265	0.009	0.01314926
Receiver	4	6	0.00781265	0.011	0.02274086
Receiver	4	6	0.00781265	0.013	0.03220683
Receiver	5	5	0.00752337	0.007	0.00944155
Receiver	5	5	0.00752337	0.009	0.01453571
Receiver	5	5	0.00752337	0.011	0.02139360
Receiver	5	5	0.00752337	0.013	0.02911447
Receiver	6	4	0.00728757	0.007	0.00624297
Receiver	6	4	0.00728757	0.009	0.01094130
Receiver	6	4	0.00728757	0.011	0.01727522
Receiver	6	4	0.00728757	0.013	0.02337046
Receiver	7	3	0.00709502	0.007	0.00503452
Receiver	7	3	0.00709502	0.009	0.00885208
Receiver	7	3	0.00709502	0.011	0.01380121
Receiver	7	3	0.00709502	0.013	0.01829888
Receiver	8	2	0.00693740	0.007	0.00402013
Receiver	8	2	0.00693740	0.009	0.00639598
Receiver	8	2	0.00693740	0.011	0.00942350
Receiver	8	2	0.00693740	0.013	0.01257555
Receiver	9	1	0.00680797	0.007	0.00223825
Receiver	9	1	0.00680797	0.009	0.00344425
Receiver	9	1	0.00680797	0.011	0.00494683
Receiver	9	1	0.00680797	0.013	0.00650348

Table D.4: *Receiver swaption prices for different expiries, tenors and strikes. Parameters and initial short rate according to 02-09-2008.*
Parameters: $\psi = 0.2592$, $\theta = 0.0063$, $\sigma = 0.0840$ *and* $r(0) = 0.0165$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Payer	1	9	0.00884053	0.007	0.02995469
Payer	1	9	0.00884053	0.009	0.01942417
Payer	1	9	0.00884053	0.011	0.01349789
Payer	1	9	0.00884053	0.013	0.01022182
Payer	2	8	0.00916161	0.007	0.02359293
Payer	2	8	0.00916161	0.009	0.01337047
Payer	2	8	0.00916161	0.011	0.00772328
Payer	2	8	0.00916161	0.013	0.00526652
Payer	3	7	0.00933355	0.007	0.02166806
Payer	3	7	0.00933355	0.009	0.01190996
Payer	3	7	0.00933355	0.011	0.00691360
Payer	3	7	0.00933355	0.013	0.00592841
Payer	4	6	0.00942623	0.007	0.01495732
Payer	4	6	0.00942623	0.009	0.00597763
Payer	4	6	0.00942623	0.011	0.00178326
Payer	4	6	0.00942623	0.013	0.00099775
Payer	5	5	0.00947668	0.007	0.01309419
Payer	5	5	0.00947668	0.009	0.00587023
Payer	5	5	0.00947668	0.011	0.00215999
Payer	5	5	0.00947668	0.013	0.00115113
Payer	6	4	0.00950449	0.007	0.01108152
Payer	6	4	0.00950449	0.009	0.00547159
Payer	6	4	0.00950449	0.011	0.00241275
Payer	6	4	0.00950449	0.013	0.00138460
Payer	7	3	0.00952007	0.007	0.00895462
Payer	7	3	0.00952007	0.009	0.00484914
Payer	7	3	0.00952007	0.011	0.00252230
Payer	7	3	0.00952007	0.013	0.00160543
Payer	8	2	0.00952895	0.007	0.00658903
Payer	8	2	0.00952895	0.009	0.00397084
Payer	8	2	0.00952895	0.011	0.00238389
Payer	8	2	0.00952895	0.013	0.00164340
Payer	9	1	0.00953412	0.007	0.00378775
Payer	9	1	0.00953412	0.009	0.00253223
Payer	9	1	0.00953412	0.011	0.00172653
Payer	9	1	0.00953412	0.013	0.00129346

Table D.5: *Payer swaption prices for different expiries, tenors and strikes. Parameters and initial short rate according to 01-06-2011.*
Parameters: $\psi = 0.6957$, $\theta = 0.0097$, $\sigma = 0.1448$ *and* $r(0) = 0.00048$.

Type	Expiry	Tenor	Forward swap rate	Strike	Price
Receiver	1	9	0.00884053	0.007	0.01418273
Receiver	1	9	0.00884053	0.009	0.02078534
Receiver	1	9	0.00884053	0.011	0.03199341
Receiver	1	9	0.00884053	0.013	0.04586124
Receiver	2	8	0.00916161	0.007	0.00718216
Receiver	2	8	0.00916161	0.009	0.01214245
Receiver	2	8	0.00916161	0.011	0.02167675
Receiver	2	8	0.00916161	0.013	0.03441145
Receiver	3	7	0.00933355	0.007	0.00623364
Receiver	3	7	0.00933355	0.009	0.00970188
Receiver	3	7	0.00933355	0.011	0.01792896
Receiver	3	7	0.00933355	0.013	0.03017940
Receiver	4	6	0.00942623	0.007	0.00124684
Receiver	4	6	0.00942623	0.009	0.00356980
Receiver	4	6	0.00942623	0.011	0.01067480
Receiver	4	6	0.00942623	0.013	0.02119713
Receiver	5	5	0.00947668	0.007	0.00149048
Receiver	5	5	0.00947668	0.009	0.00363749
Receiver	5	5	0.00947668	0.011	0.00929545
Receiver	5	5	0.00947668	0.013	0.01766155
Receiver	6	4	0.00950449	0.007	0.00174268
Receiver	6	4	0.00950449	0.009	0.00359079
Receiver	6	4	0.00950449	0.011	0.00798783
Receiver	6	4	0.00950449	0.013	0.01442065
Receiver	7	3	0.00952007	0.007	0.00194415
Receiver	7	3	0.00952007	0.009	0.00340252
Receiver	7	3	0.00952007	0.011	0.00663799
Receiver	7	3	0.00952007	0.013	0.01128689
Receiver	8	2	0.00952895	0.007	0.00192490
Receiver	8	2	0.00952895	0.009	0.00299516
Receiver	8	2	0.00952895	0.011	0.00509573
Receiver	8	2	0.00952895	0.013	0.00804488
Receiver	9	1	0.00953412	0.007	0.00146457
Receiver	9	1	0.00953412	0.009	0.00204221
Receiver	9	1	0.00953412	0.011	0.00306925
Receiver	9	1	0.00953412	0.013	0.00446986

Table D.6: *Receiver swaption prices for different expiries, tenors and strikes. Parameters and initial short rate according to 01-06-2011.*
Parameters: $\psi = 0.6957$, $\theta = 0.0097$, $\sigma = 0.1448$ *and* $r(0) = 0.00048$.

Appendix E

Selected R code

E.1 RMSE Estimation for CDS prices

```
NumMinCDS <- function(Obs, X){  # 'X' og 'Obs' vectors

# Define object function:
    obj <- function(parms, Obs, X)

    {
  par1  <- parms[1]
par2  <- parms[2]
par3  <- parms[3]
par4  <- parms[4]

Est<-rep(0,length(X))
      for(Count in 1:length(X))
{

Est[Count]<- Cds(t,X[1,Count],Kappa_Cox=par1,
Sigma_Cox=par2,Theta_Cox=par3,Lambda0=par4,Freq)
}
     # Calculates RMSE and returns:
      obj <- sqrt((1/length(X))*sum((Est - Obs)^2))
    }

# Minimizing the objectfunction using 'nlminb()'-functionen in R
  opt <- nlminb(start=c(Kappa_Cox,Sigma_Cox,Theta_Cox,Lambda0),   # Initial guess
              objective=obj,     # Object function
              Obs=Obs,           # Extra input to object function
              X=X,             # Extra input to object function
  lower=c(0.0001,0.0001,0.00001,0.00001),
  upper=c(1,1,1,1))

  # Results are in 'opt'. Info in str(opt)
#  Select info:
  parms <- opt$par
  names(parms) <- c("Kappa","Sigma","Theta","Lambda0") # Parameters
```

```
  obj <- opt$objective  # Final value of object function

  estimates<-rep(0,length(X))
  for(Count in 1:length(X))
  {
estimates[Count]<- Cds(t,X[1,Count],Kappa_Cox=parms[1],Sigma_Cox=parms[2],
Theta_Cox=parms[3],Lambda0=parms[4],Freq)
  }
  # Returns in a list:
  return(c(parms=parms, est=estimates, obs=Obs, obj=obj))

}
```

E.2 Swaption Pricing

```
library(deSolve)
library(bvpSolve)

for(i in 1:NoSolutions)
{

Delta1<-complex(real=0,imaginary=i-1) # Delta1

#Complex differential equation dy/dt
Diff <- function(t, b, h, pars)
{
b <- complex(real=b[1],imaginary=b[2]) # Delta2
db <- (Sigma^2*Wt(t)+a)*b-(1/2)*(Delta1*Kt(t)*Sigma+b*Sigma)^2
h <-complex(real=h[1],imaginary=h[2]) # Gamma
dh <- -a * Theta * b
list(c(re=Re(db),im=Im(db),re2=Re(dh),im2=Im(dh)))
}
#Unknown initial condition
init <- c(re=NA,im=NA,re2=NA,im2=NA)

#Known terminal condition
end  <- c(re=0,im=0,re2=0,im2=0)
sol  <- bvptwp(yini = init, yend = end, x = seq(t, Tn, by = 0.5), func = Diff)

delta2_Re[i]<-sol[1,2]
delta2_Im[i]<-sol[1,3]
gamma_Re[i]<-sol[1,4]
gamma_Im[i]<-sol[1,5]
}

#Spline Interpolation of ODE solutions
DeltaReSpline<-splinefun(x=seq(1,NoSolutions,by=1), y = delta2_Re[1:NoSolutions],
method = c("fmm", "periodic", "natural", "monoH.FC"),ties = mean)
```

```
DeltaImSpline<-splinefun(x=seq(1,NoSolutions,by=1), y = delta2_Im[1:NoSolutions],
method = c("fmm", "periodic", "natural", "monoH.FC"),ties = mean)
GammaReSpline<-splinefun(x=seq(1,NoSolutions,by=1), y = gamma_Re[1:NoSolutions],
method = c("fmm", "periodic", "natural", "monoH.FC"),ties = mean)
GammaImSpline<-splinefun(x=seq(1,NoSolutions,by=1), y = gamma_Im[1:NoSolutions],
method = c("fmm", "periodic", "natural", "monoH.FC"),ties = mean)


#Spline Delta2 function
DeltaSpline<-function(u)
{
if(SwaptionType==1) #Payer
{
return(complex(real=DeltaReSpline(u),imaginary=DeltaImSpline(-Alpha)))
}
else if(SwaptionType==0) #Receiver
{
return(complex(real=DeltaReSpline(u),imaginary=DeltaImSpline(Alpha)))
}
}


#Spline Gamma function
GammaSpline<-function(u)
{
if(SwaptionType==1) #Payer
{
return(complex(real=GammaReSpline(u),imaginary=GammaImSpline(-Alpha)))
}
else if(SwaptionType==0) #Receiver
{
return(complex(real=GammaReSpline(u),imaginary=GammaImSpline(Alpha)))
}
}

#Spline CCF
CCFSpline<-function(u)
{
if(SwaptionType==1) #Payer
{
return(exp(GammaSpline(u+1)+complex(real=Alpha,imaginary=u)*y(t)+DeltaSpline(u+1)*r0))
}
else if(SwaptionType==0) #Receiver
{
return(exp(GammaSpline(u+1)+complex(real=-Alpha,imaginary=u)*y(t)+DeltaSpline(u+1)*r0))
}
}

IntegrandSpline<-function(u)
```

```
{
if(SwaptionType==1) #Payer
{
PhiSP<-PVBP(t)*(CCFSpline(u)/(complex(real=Alpha,imaginary=u))^2)
return((exp(-Alpha*K)/pi)*Re(exp(-complex(real=0,imaginary=u)*K)*PhiSP))
}
else if(SwaptionType==0) #Receiver
{

PhiSP<-PVBP(t)*(CCFSpline(u)/(complex(real=-Alpha,imaginary=u))^2)
return((exp(Alpha*K)/pi)*Re(exp(-complex(real=0,imaginary=u)*K)*PhiSP))
}
}

#Swaption Price using Spline CCF
integrate(IntegrandSpline,lower=0,upper=1000, subdivisions=1000)
```

E.3 Monto Carlo Simulation of Interest Rates and Intensities

```
time<-seq(from=t,to=Tn,length=n)

for(j in 1:(j-1))
{
r<-rep(0,n)
swap<-rep(0,n)
lambda<-rep(0,n)
r[1]<-r0
swap[1]<-y(t)
lambda[1]<-Lambda0
norm<-rnorm(n+1,0,1)
for(i in 1:(n-1))

{
lambda[i+1]<-lambda[i]+Kappa_Cox*(Theta_Cox-max(lambda[i],0))*(time[i+1]-time[i])
+Sigma_Cox*sqrt(max(lambda[i],0))*sqrt(time[i+1]-time[i])*norm[i+1]
r[i+1]<-r[i]+(a*Theta-(Sigma^2*Wt(time[i])+a)*max(r[i],0))*(time[i+1]-time[i])
+Sigma*sqrt(max(r[i],0))*sqrt(time[i+1]-time[i])*norm[i+1]
swap[i+1]<-swap[i]+Kt(time[i])*Sigma*sqrt(r[i])*sqrt(time[i+1]-time[i])*norm[i+1]
}
R_T[j]<-r[n-1]
Swap_T[j]<-swap[n-1]
Lambda_T[j]<-lambda[n-1]
}
```

Appendix F

Selected Matlab Code

F.1 Maximum Likelihood Estimation for One-Factor CIR process

```
function ML_CIRparams = CIR_calibration(V_data,dt,params)
% ML_CIRparams = [ alpha theta sigma ]
N= length(V_data);
if nargin <3
x = [ ones(N-1 ,1) V_data (1:N -1)];
ols = (x'*x )^( -1)*(x'* V_data (2:N ));
m=mean ( V_data ); v= var( V_data );
params = [- log( ols (2))/ dt ,m, sqrt (2* ols (2)* v/m)];
end
options = optimset ('MaxFunEvals', 200000 , 'MaxIter', 200000);
ML_CIRparams = fminsearch( @FT_CIR_LL_ExactFull , params , options );
function mll = FT_CIR_LL_ExactFull( params )
alpha = params (1); teta = params (2); sigma = params (3);
c = (2* alpha )/(( sigma ^2)*(1 - exp(- alpha *dt )));
q = ((2* alpha * teta )/( sigma ^2)) -1;
u = c* exp(- alpha *dt )* V_data (1:N -1);
v = c* V_data (2:N);
mll = -(N -1)* log(c)+ sum(u+v- log(v./u)*q /2 -...
log( besseli (q ,2* sqrt (u.*v) ,1)) - abs( real (2* sqrt(u.*v ))));
end
end
```