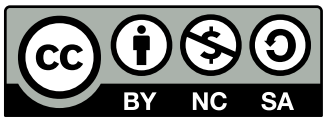
Aalto University
School of Science

Master's Programme in Mathematics and Operations Research

Calibration of Bermudan Swaptions

Ville Närhi

Master's Thesis
2026



Aalto University
School of Science

Author Ville Närhi

Title Calibration of Bermudan Swaptions

Degree programme Mathematics and Operations Research

Major Systems and Operations Research

Supervisor Prof. Ahti Salo

Advisor Dr. Vesa Tähtinen

Date April 24, 2026 **Number of pages** 82+7 **Language** English

Abstract

The accurate valuation of interest rate derivatives is crucial for financial institutions to understand their exposures and assess risk. Bermudan swaptions, derivatives that allow the holder to enter an interest rate swap on a set of predetermined exercise dates, are among the most common complex interest rate derivatives, widely used for hedging and risk management in the derivatives books of institutions.

In this thesis, we develop the necessary framework to price Bermudan swaptions. We begin with the mathematical and financial preliminaries, before discussing volatility and interest rate modeling, calibration methodology, and the pricing process itself. Calibration here refers to finding interest rate model parameters such that model-implied volatilities match observed market volatilities as closely as possible. In this thesis, only the volatility parameter is calibrated, others are left as user inputs.

We then present the numerical setup, covering the target Bermudan swaption, the yield curves used, and the SABR model employed to interpolate market volatilities. We consider two calibration strategies: ATM calibration, which uses instruments struck at current at-the-money rates, and strike-targeted calibration, which uses instruments that matches the strike of the Bermudan at 2.00%. Three interest rate models are examined: the Hull-White one-factor (HW1F) model, the Hull-White two-factor (HW2F) model, and the one-factor Cheyette local volatility model. HW1F serves as the industry-standard benchmark, HW2F extends it with an additional factor to better capture term structure dynamics, and the Cheyette local volatility model further develops this with a local volatility specification that depends on both time and the current level of rates, offering greater flexibility in fitting market volatility quotes.

We find that strike-targeted calibration produces superior pricing results and model-implied volatilities, and is therefore the recommended calibration strategy for Bermudan swaptions. The HW2F model underperforms relative to the other two, and is not recommended. The HW1F and Cheyette local volatility models perform rather similarly, and the choice between them ultimately depends on institutional priorities: HW1F offers simplicity and computational efficiency, while the Cheyette model provides greater flexibility and adaptability across changing rate environments.

Keywords Bermudan swaptions, interest rate derivatives, interest rate models, volatility calibration, Hull-White model, Cheyette model

Aalto-yliopisto
Perustieteiden
korkeakoulu

Tekijä Ville Närhi

Työn nimi Bermudalaisten swaptioiden kalibrointi

Koulutusohjelma Mathematics and Operations Research

Pääaine Systems and Operations Research

Työn valvoja Prof. Ahti Salo

Työn ohjaaja FT Vesa Tähtinen

Päivämäärä 24.4.2026 **Sivumäärä** 82+7 **Kieli** englanti

Tiivistelmä

Korkojohdannaisten tarkka arvostus on rahoitusalan yrityksille keskeistä riskiposition ymmärtämiseksi ja riskienhallinnan kannalta. Bermudalaiset swaptiot ovat johdannaisia, jotka antavat niiden haltijalle oikeuden astua koronvaihtosopimukseen jonain ennalta määriteltynä toteutuspäivänä. Ne ovat yksi yleisimmistä monimutkaisista korkojohdannaisista, ja niitä käytetään laajalti suojaukseen ja riskienhallintaan. Niiden arvostus riippuu korkojen tulevasta kehityksestä ja on aktiivisen tutkimustyön kohde.

Tässä diplomityössä kehitetään tarvittava kehys Bermudalaisten swaptioiden hinnoitteluun. Aluksi käydään matemaattiset sekä rahoitusteoreettiset esitiedot läpi, minkä jälkeen käsitellään volatiliteetti- ja korkomallinnusta, kalibrointimenetelmiä, sekä itse hinnoitteluprosessia. Kalibroinnilla tarkoitetaan korkomallien parametrien määrittämistä siten, että mallin implisiittiset volatiliteetit vastaavat mahdollisimman tarkasti havaittuja markkinavolatiliteetteja. Tässä opinnäytetyössä ainoastaan volatiliteetti kalibroidaan, muut parametrit jätetään käyttäjän määritettäviksi.

Seuraavaksi esitellään työn numeerinen asetelma, joka kattaa hinnoiteltavan Bermudalaisen swaption, korkokäyrät sekä SABR-mallin markkinavolatiliteettien interpolointiin. Tarkasteltavana on kaksi kalibrointistrategiaa: ATM-kalibrointi, joka käyttää nykyisellä tasolla olevia instrumentteja, sekä strike-kohdennettu kalibrointi, mikä käyttää Bermudalaisen swaptionin strike-korkoa vastaavia instrumentteja. Työssä käytetään kolmea korkomallia: Hull-Whiten yhden faktorin mallia (HW1F), Hull-Whiten kahden faktorin mallia (HW2F), sekä yhden faktorin Cheyetten lokaalin volatiliteetin mallia. HW1F toimii alalla vakiintuneena vertailumallina, HW2F laajentaa tätä lisäfaktorilla paremman termirakenteen dynamiikan kuvaamiseksi. Cheyette kehittää tätä edelleen sekä ajasta että korkojen nykyisestä tasosta riippuvalla lokaalilla volatiliteetilla, mikä tarjoaa parempaa joustavuutta markkinavolatiliteettien sovittamisessa.

Strike-kohdennettu kalibrointi tuottaa paremmat hinnoittelutulokset ja mallin implisiittiset volatiliteetit, ja on siten suositeltava kalibrointistrategia Bermudalaisille swaptioille. HW2F-malli suoriutuu heikoimmin malleista, eikä sitä suositella. HW1F ja Cheyette suoriutuvat melko samankaltaisesti, ja valinta niiden välillä riippuu yrityksen prioriteeteista: HW1F on yksinkertainen ja laskennallisesti tehokas, kun taas Cheyette tarjoaa paremman joustavuuden ja sopeutuvuuden muuttuvissa korkoympäristöissä.

Avainsanat Bermudalainen swaptio, korkojohdannaiset, korkomallit, volatiliteetin kalibrointi, Hull-White korkomalli, Cheyette korkomalli

Preface

First of all, I want to extend my gratitude to Professor Ahti Salo for his supervision on this thesis, as well as my advisor Vesa Tähtinen for his expert guidance and support throughout the entire process of conducting this thesis. I also want to thank Jaakko Juntunen for providing the idea for the topic of this thesis, as well as everyone else that has been involved in the process, such as my coworkers at Kuntarahoitus. This thesis has been an extremely rewarding journey for me through which I have learned a lot, and undertaking this journey would not have been possible without the people mentioned above.

It feels weird to be writing this after having studied here for nearly six years, and now this chapter of my life is coming to an end. Countless days, most of them long but some short, spent at the university, both studying and teaching. Thinking about it now, I must say I could hardly be happier with a younger me for choosing this degree. I have made memories and friends that will last me a lifetime, and have learned things I would never have even thought existed before studying, including quite a lot in just this thesis. It has definitely been a long road filled with a lot of good memories (and stress), and while I am happy it is coming to an end, I do also feel a bit sad as my head gets filled with memories from these past years while writing this. I wish to also extend my thanks to every professor and lecturer I have had along the way, whom I have learned from.

Most importantly, I want to thank everyone who has been here along for the ride with me - my friends and family especially. I could not have made it here without the support of everyone, and am truly grateful to everyone who has been here for me along the years. It has been a long road filled with countless exams, assignments and lectures, and while it has been stressful at times, I would not exchange these years for anything. I have also grown a lot as a person and learned a lot about myself throughout these years, and I think that is just as important as the topics I have learned about.

Otaniemi, April 24, 2026

Ville Närhi

Contents

Symbols and abbreviations

Symbols

$A(t)$	Swap annuity
$H_i(T_i)$	Continuation value at exercise date T_i
$F(t;T,S)$	Simply-compounded forward rate
$P(t,T)$	Zero-coupon bond price (discount factor)
$R_{\text{swap}}(t)$	Swap rate
V_{swap}	Value of an interest rate swap
V_{swaption}	Value of a European swaption
V_{Bermudan}	Value of a Bermudan swaption
W_t	Brownian motion
$f(t,T)$	Instantaneous forward rate
r_t	Instantaneous short rate
α	Speed of mean reversion
α_{SABR}	SABR initial volatility
β	SABR elasticity parameter
δ_i	Day-count fraction i
ν	SABR volatility of volatility
σ	Volatility

Abbreviations

ATM	At-the-money
BM	Brownian motion
BSM	Black-Scholes-Merton formula
bps	Basis points
HJM	Heath-Jarrow-Morton framework
HW1F	Hull-White 1-factor model
HW2F	Hull-White 2-factor model
ITM	In-the-money
LIBOR	London Interbank Offered Rate
LSMC	Least-squares Monte Carlo
LV	Local volatility
LVM	Local volatility model
OIS	Overnight indexed swap
OTM	Out-of-the-money
PDE	Partial differential equation
RMSE	Root mean square error
SABR	Stochastic Alpha Beta Rho model
SDE	Stochastic differential equation
SOFR	Secured Overnight Financing Rate
SV	Stochastic volatility

1 Introduction

1.1 Motivation

The 2008 global financial crisis exposed the fragility of the global financial system. What started as a collapse in the United States housing market quickly evolved into a worldwide liquidity crisis, eventually leading to stock market crashes, the downfall of major financial institutions, and notable regulatory reforms. One of the primary reasons behind the financial crisis was the lack of understanding and, consequently, mispricing of complex derivatives such as mortgage-backed securities and collateralized debt obligations. These instruments derive their value from mortgage payments and housing prices and were often built on assumptions of stable market conditions and limited risk. Their valuation models were oversimplified or poorly understood, leading institutions to underestimate their exposure and take excessive positions that they should not have been taking. When housing prices then began to decline and default rates increased, these instruments lost value rapidly, causing widespread margin calls and liquidity shortages. This led to turbulence throughout the global economy, leading to the financial crisis. For a detailed account of the causes of the financial crisis, see Brunnermeier (2009).

In response, a significant regulatory overhaul was implemented. These new regulations included, for example, the Dodd-Frank Act in the United States and EMIR in Europe for stricter oversight of over-the-counter derivatives, the Basel III and IV accords for more emphasis on model transparency, capital adequacy, risk sensitivity, and more. These reforms reshaped the financial world, mandating financial institutions to emphasize and justify their valuation methodologies, pricing models, and accurate risk quantification, including stress testing for extreme conditions such as the financial crisis. As a result, accurate pricing of financial products was no longer just a luxury, but became a regulatory necessity. The importance of these regulations for risk management practices has also been evident in recent years, as systematic vulnerabilities in global markets will always exist. For example, the market shock caused by the COVID-19 pandemic highlighted the need for robust pricing models capable of performing under uncertain and rapidly changing market conditions (Goodell, 2020). The collapse of Silicon Valley Bank in 2023 (Van Vo and Le, 2023), partly caused by insufficient interest rate risk management, served as another reminder of the critical importance of effective risk management and how even small oversight in risk management practices can lead to the downfall of an institution and have systemic consequences. Furthermore, geopolitical tensions around the world add further uncertainty and volatility to global markets every year. These events, among others, emphasize the importance of effective risk management from financial institutions under varying market conditions so that the global economy does not collapse.

With ever-changing market conditions and regulations, it is critical that models not only consider the current market data, but also behave reasonably well under stress so that financial institutions remain within required limits in different risk metrics. This is especially true for interest rate derivatives, the largest class of derivatives,

which represent 79% of the total global over-the-counter derivatives notional value of around $579 trillion (International Swaps and Derivatives Association (ISDA), 2024). As the name suggests, interest rate derivatives derive their value from interest rates such as Euribor, SOFR, and other reference rates used around the world. Instruments such as swaps, caps, floors, and swaptions are among the most common interest rate derivatives, which are essential for hedging interest rate exposure, speculating on future interest rate movements, asset and liability management, and other purposes. Their valuation therefore affects front-office trading decisions, pricing, as well as risk metrics used by the entire organization, and even relatively small inaccuracies can have significant consequences, affecting balance sheets, risk metrics, and overall financial stability of the institution.

European swaptions are interest rate derivatives that grant the holder the right, but no obligation, to enter an interest rate swap - a contract in which one party pays a fixed rate and receives a floating rate - at a given future date. American swaptions, on the other hand, can be exercised whenever during the lifetime of the option, and in between these two are Bermudan swaptions - instruments that allow the holder to enter an interest rate swap on multiple predetermined exercise dates. This feature gives remarkable flexibility compared to European swaptions by allowing the holder to change between fixed and floating interest rates on a known schedule, giving the opportunity to adapt to changing interest rates with relative ease, while also remaining cheaper than American swaptions. However, this flexibility comes at a cost, as the value of a Bermudan swaption is dependent on the evolution of interest rates throughout the lifetime of the instrument. More precisely, their valuation involves an optimal stopping problem, where the holder must decide at each exercise date whether to exercise the option or continue holding it. This makes their valuation considerably more complex than that of European swaptions, as there does not exist a closed-form solution, and therefore Bermudan swaptions must be priced via numerical methods.

In order to price Bermudan swaptions accurately, interest rate models are required to simulate how the short rate evolves over time. These models act as the basis for pricing interest rate derivatives by generating future rate paths and discount factors consistent with observed market data, providing the framework to describe the evolution of the term structure. As such, the choice of interest rate model is crucial for pricing interest rate derivatives, particularly for more complex products like Bermudan swaptions, where the exercise decisions are sensitive to assumptions about the behavior of interest rates over time. In order to match observed market prices as accurately as possible, the parameters of interest rate models are calibrated, i.e., the aim is to find parameter values so that the prices given by the model match market prices as accurately as possible.

For Bermudan swaptions, the calibration instruments employed are the highly liquid European swaptions quoted as implied volatilities. These quotes form the so-called volatility cube, describing the market's assessment of uncertainty (volatility) as a function of tenor, maturity, and strike. Although the cube contains rich information on the distribution of future swap rates, models are often calibrated only to at-the-money (ATM) instruments due to convenience, good market liquidity, and historical practice. However, this simplification can introduce pricing biases for path-dependent

instruments such as Bermudan swaptions, whose early-exercise features heavily depend on future rates as they can be exercised when the underlying swap is far from the ATM level. Bermudans often have long maturities and tenors, sometimes over 20 years - rates can change substantially over such a long time period, and calibrating the model to ATM swaptions may consequently provide inaccurate valuations. Instead, Andersen and Piterbarg (2010b) propose that the strike of the calibration instruments should reflect the strike of the relevant Bermudan itself to localize the calibration to the Bermudan being priced. The European swaptions used for calibration are therefore chosen so that their strike coincides with the fixed rate of the Bermudan. This ensures consistency with the valuation of European swaptions, which is crucial as a Bermudan swaption can be considered as the maximum value swaption of the given set of times.

Andersen and Piterbarg (2010b) also note that the ability to capture the volatility smile or skew (Hull and Basu, 2016) is important for Bermudan swaptions to consider future rates, and hence pricing, accurately. However, many common interest rate models cannot replicate this smile or skew properly and might fail to accurately capture the distribution of future rate paths. In more complicated instruments, such as Bermudan swaptions, this simplification can lead to valuation errors, as the model could underestimate the value of a swap far in-the-money (ITM) or overestimate the value of holding an option that is deep out-of-the-money (OTM). More traditional models, such as the Hull-White 1-Factor model (HW1F) (Hull and White, 1990), offer relative simplicity, computational speed, analytical tractability, and are easy to implement. However, they can only be calibrated to one volatility per expiry, and their deterministic volatility form does not allow smiles and skews to be replicated, which requires more flexible interest rate models with more sophisticated features.

The local volatility version of the Cheyette (or quasi-Gaussian) model (Cheyette, 2001) offers an improved fit to market data and is able to capture the slopes of volatility smiles in addition to volatilities at specific strikes. Naturally, this comes at the cost of some added complexity and computational demands. However, the inability of models such as HW1F to capture volatility smiles is often exchanged for their benefits, especially if the institutions are not trading in non-ATM Bermudans.

This trade-off between model simplicity and pricing accuracy, particularly in capturing the volatility smile for complex derivatives like Bermudan swaptions, is at the heart of modern interest rate modeling. This thesis aims to explore this by investigating how the choice of calibration instruments and interest rate models influence the pricing of Bermudan swaptions. Particularly, we seek to determine whether more trade-specific calibration and more complex models lead to more reliable pricing and justify the trade-offs in simplicity.

1.2 Research Questions

Building upon the identified challenges in accurately pricing non-ATM Bermudan swaptions, this thesis examines how the choice of interest rate model and calibration strategy affect the valuation of Bermudan swaptions and how well each model captures the volatility smile. We consider three interest rate models that are commonly used in the financial sector.

1. Hull-White one-factor model (HW1F) (Hull and White, 1990). A short-rate model that serves as a benchmark due to its highly desirable properties, relative simplicity, and wide acceptance in the industry.

2. Hull-White two-factor model (HW2F) (Hull, 1994). An extension of HW1F with two correlated factors. It is chosen to examine whether increasing model dimensionality (i.e. factor count) improves the fit to the volatility smile and pricing of Bermudan swaptions.

3. Cheyette one-factor local volatility model (Cheyette, 2001). It can be viewed as a Markovian representation of the Heath-Jarrow-Morton (HJM) framework (Heath et al., 1992) that incorporates a local volatility specification and avoids typical problems with HJM models. It is commonly regarded to be well-suited for Bermudan swaptions as it captures both smile effects and state-dependent evolution of volatility, providing an advantage over the Hull-White models, while still remaining relatively simple in contrast to many other models.

With these models, we aim to answer whether more complex models capable of capturing properties of the volatility smile lead to more accurate and consistent pricing for Bermudan swaptions. An additional consideration is whether the number of stochastic factors (HW1F/Cheyette vs. HW2F) impacts Bermudan pricing performance, and whether the additional factor is worth it.

However, the performance of these models strongly depends on how well they are calibrated. Calibration refers to the practice of tuning the parameters of a model so that the model-implied prices or volatilities match the observed market data as closely as possible (Cont and Tankov, 2003). The volatility parameter is often the primary focus of calibration, though additional parameters like the mean-reversion rate and correlation can also be calibrated depending on the model. In this thesis, we only calibrate the volatility due to it being standard practice in the industry, and other parameters are left as user inputs estimated based on the market conditions. The primary research question is the choice of the calibration set, i.e. the set of market instruments to which a model is fitted to align its outputted prices or volatilities with market data. Financial institutions often consider only ATM swaptions in the calibration set, where the parameters are fitted only to the at-the-money European swaption surface, which is the first calibration strategy we consider. The other calibration scheme we consider is strike-targeted calibration, in which the parameters are fitted to European swaptions whose strikes match those of the fixed rate of the Bermudan swaption being valued. This localizes the calibration to the specific Bermudan swaption, theoretically producing the local structure of the volatility smile around the relevant moneyness level and providing a better assessment of exercise opportunities around the relevant rates.

This thesis investigates the impact of model choice and calibration methodology on the pricing of Bermudan swaptions. In particular, the following research questions are addressed:

1. How does the choice between ATM and strike-targeted calibration strategies affect the pricing of Bermudan swaptions, particularly in considering the volatility smile?

2. How does the Cheyette local-volatility model compare with the Gaussian Hull-White frameworks in reproducing the volatility smile and achieving stable Bermudan valuations?

3. Does increasing the number of stochastic factors - from one in HW1F to two in HW2F - improve Bermudan pricing performance?

Each model is calibrated to the same European swaption market dataset with quotes from the same day, ensuring consistency and comparability across the numerical experiments. The calibrated models are then used to value a Bermudan swaption, providing a direct assessment of how calibration strategy and model design affect pricing outcomes as well as the model-implied volatilities. Ultimately, the goal is to assess how different combinations of models and calibration strategies influence the pricing of Bermudan swaptions, and whether one combination is better than another. Although pricing accuracy and volatility modeling are the most important considerations, we also address other aspects such as computational efficiency, simplicity, and overall practicality. These secondary considerations help evaluate whether gains in accurate pricing justify the trade-off in other aspects, which contributes to a more informed understanding of how model and calibration choices influence the valuation and risk management of complex interest rate derivatives.

1.3 Structure

The remainder of this thesis is structured as follows. Chapter 2 establishes the theoretical foundation by introducing the key mathematical concepts underlying interest rate modeling, including stochastic calculus, martingales, and risk-neutral valuation. The section also provides an overview of interest rate derivatives. Starting from the fundamental building blocks of interest rates, it proceeds to discuss interest rate swaps, option pricing, and swaptions. It then reviews the volatility smile and relevant market dynamics before concluding with an introduction to Bermudan swaptions. Chapter 3 presents the methodological framework of the thesis. It begins with a discussion of volatility modeling approaches, including deterministic, local, and stochastic volatility. It then reviews interest rate models, with particular emphasis on the models employed in this study. Finally, it describes the calibration procedure used to fit model parameters to market data and outlines the pricing methodology for Bermudan swaptions.

Chapter 4 describes the implementation and experimental setup used in the study. In particular, it outlines the Bermudan swaption being priced, the data and market instruments used for calibration, and the construction of the calibration sets. The section also discusses the implementation details and parameter specifications of the interest rate models considered. Chapter 5 presents the numerical results, including an assessment of the calibration quality, how the different models and calibration strategies influence the pricing of Bermudan swaptions, and how effectively they replicate the volatility smile. Finally, Chapter 6 summarizes the key findings and their implications, provides a critical evaluation of the results, and suggests potential avenues for further research on the topic.

2 Fundamentals of Financial Mathematics

2.1 Preliminaries

This section introduces the necessary mathematical preliminaries required to understand the stochastic modeling framework used throughout this thesis for the pricing of Bermudan swaptions. We mostly follow the notation used by Øksendal (2003) and Shreve (2004). For a more comprehensive treatment of stochastic calculus and its applications in financial mathematics, we refer to these books, as well as Karatzas and Shreve (1998), Mörters and Peres (2010), and Björk (2009).

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, where Ω is the sample space representing the set of all possible outcomes, $\mathcal{F}$ is a σ-algebra whose elements are events with a corresponding probability, $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$, and $(\mathcal{F}_t)_{t\geq 0}$ is a filtration representing the available information up to time t, defined as an increasing sequence of σ-algebras $\mathcal{F}_s \subseteq \mathcal{F}_t$ when $s \leq t$. The filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions, that is,

- it is right-continuous: $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \geq 0$,
- it is complete: $\mathcal{F}_0$ contains all subsets of $\mathbb{P}$-null sets $\mathcal{N}$ in $\mathcal{F}$, i.e. $\mathcal{N} \subset \mathcal{F}_0$.

Unless otherwise specified, all stochastic processes in this thesis are assumed to be adapted to this filtration, i.e. any stochastic process X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$. Such a filtration satisfying the usual conditions is guaranteed to exist for Brownian motion via the augmented natural filtration, constructed by enlarging the σ-algebras generated by the Brownian motion itself with all $\mathbb{P}$-null sets of $\mathcal{F}$ (Karatzas and Shreve, 1988).

2.1.1 Brownian Motion

Brownian motion is one of the most fundamental stochastic processes in modern probability theory and quantitative finance. The process was first observed by Robert Brown in 1827 in the context of seemingly random movement of pollen grains suspended in liquid (Brown, 1828). The process was later mathematically formulated by Wiener (1923), after whom the process is often (synonymously) called the Wiener process. In financial modeling, it was first used by Bachelier (1900) to model the stochastic evolution of stock prices and became the cornerstone of modern stochastic calculus through the work of Kiyosi Itô.

Brownian motion is central in stochastic calculus and quantitative finance as the stochastic process underlying continuous-time models. It is able to describe random evolution and uncertainty, and as such can inherently capture the unpredictable and ever-fluctuating nature of financial markets. Consequently, it serves as the model of uncertainty in all stochastic differential equations used throughout the thesis.

Definition 2.1 (Brownian motion)**.** A standard Brownian motion is a continuous-time stochastic process with the following properties:

1. $W_0 = 0$ almost surely,

2. For any $0 \leq s < t$, the increments $W_t - W_s$ are normally distributed with mean 0 and variance $t - s$, i.e. $W_t - W_s \sim N(0, t - s)$,

3. The increments $W_t - W_s$ are independent, i.e. for any partition $0 < t_0 < t_1 \cdots < t_n$ the increments $W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent random variables,

4. The sample paths $t \mapsto W_t$ are almost surely (a.s.) continuous.

These properties imply that $\mathbb{E}[W_t] = 0$, $\mathrm{Var}(W_t) = t$, and $\mathbb{E}[W_t W_s] = \min(t, s)$ for $0 \leq s < t$. Despite its path continuity, the trajectories of Brownian motion are almost surely nowhere differentiable. This is a key reason why ordinary calculus fails in stochastic models, as the same operations do not apply for stochastic processes as they do for functions.

In finance, there are often multiple correlated sources of uncertainty, and it is therefore useful to consider multidimensional Brownian motions. Consider a vector of Brownian motions $W_t = (W_t^{(1)}, \ldots, W_t^{(m)}), t \geq 0$ with covariance structure $\mathrm{Cov}(dW_i, dW_j) = \rho_{ij} dt$, where ρ is the correlation matrix of W_t. The process $W_t = (W_t^{(1)}, \ldots, W_t^{(m)})$ is then a vector of correlated Brownian motions with correlation matrix ρ.

2.1.2 Conditional Expectation and Martingales

Conditional expectation is a key concept in pricing theory, as it formalizes how information is incorporated over time and provides the foundation for martingales. The filtration $\mathcal{F}_t$ contains the information up to the current time t, and the conditional expectation provides the best estimate of the future value of the process based on $\mathcal{F}_t$.

Definition 2.2 (Conditional expectation)**.** Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ be a random variable. The conditional expectation $\mathbb{E}[X \mid \mathcal{G}]$ is any random variable that satisfies the following:

- $\mathbb{E}[X \mid \mathcal{G}]$ is integrable
- $\mathbb{E}[X \mid \mathcal{G}]$ is $\mathcal{G}$-measurable
- $\int_G \mathbb{E}[X \mid \mathcal{G}] d\mathbb{P} = \int_G X d\mathbb{P}$ for all $G \in \mathcal{G}$.

Below we list some properties that conditional expectations satisfy (almost surely).

- Positivity: If $X \geq 0$, then $\mathbb{E}[X \mid \mathcal{G}] \geq 0$
- Linearity: $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[\mathbb{X} \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$
- Unbiased: $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$
- Tower property: If $\mathcal{H} \subseteq \mathcal{G}$, then $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}]$
- Known can be taken out: If Y is $\mathcal{G}$-measurable, then $\mathbb{E}[XY \mid \mathcal{G}] = Y\mathbb{E}[X \mid \mathcal{G}]$

- Jensen's inequality: If f is a convex function such that $f(X)$ is integrable, then $\mathbb{E}[f(X) \mid \mathcal{G}] \geq f(\mathbb{E}[\mathbb{X} \mid \mathcal{G}])$.

For more properties and proofs, and further details on conditional expectations in general, see Kallenberg (1997).

Conditional expectations in financial mathematics provide the framework for modeling the evolution of information. If $(\mathcal{F}_t)_{t\geq 0}$ represents the information available to traders at time t, the prices of assets can be formulated as conditional expectations given $\mathcal{F}_t$. This is essential when formulating the risk-neutral valuation framework and when considering martingales, for which the definition is given below.

Definition 2.3 (Martingale)**.** A stochastic process $(M_t)_{t\geq 0}$ is a martingale if it satisfies the following properties:

- M_t is adapted to the filtration $\mathcal{F}_t$, i.e. M_t is $\mathcal{F}_t$-measurable for all $t \geq 0$,
- $\mathbb{E}[|M_t|] < \infty$ for all t, i.e. M_t is integrable,
- For all $0 \leq s \leq t, \mathbb{E}[M_t \mid \mathcal{F}_s] = M_s$.

The third property, called the martingale property, is the core of the concept. The property states that, given the current information $\mathcal{F}_s$at time s, the best prediction for a future value of the process at time t is its current value. A common interpretation is that a martingale is a model of a fair game where the expected future winnings are equal to the current amount, regardless of the history of the process. M_t is called a submartingale if $\mathbb{E}[M_t \mid \mathcal{F}_s] \geq M_s$ and a supermartingale if it satisfies $\mathbb{E}[M_t \mid \mathcal{F}_s] \leq M_s$. Well-known examples of martingales are Brownian motion and Itô integrals mentioned in Section 2.1.3. Proofs for these can be found in Shreve (2004).

In financial mathematics, it is often of interest to characterize random times at which some event occurs or a decision is made, such as exercising an option. Such problems can be formulated through the concept of stopping times.

Definition 2.4 (Stopping time)**.** A stopping time is a random variable $\tau : \Omega \to [0, \infty]$ such that $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. That is, for every $t \geq 0$, the event $\{\tau \leq t\}$ is $\mathcal{F}_t$-measurable.

In practice, a stopping time represents a point in time t when a decision to stop the process is made based on the information available at that time. This provides a natural bridge to the optimal stopping problem, in which one seeks a stopping time that maximizes the expected payoff.

Definition 2.5 (Optimal stopping)**.** Let X_t be an adapted process representing a payoff observed through time. The optimal stopping problem consists of finding a stopping time τ^* that maximizes the expected gain

$$V_t^T = \sup_{\tau} \mathbb{E}[X_\tau \mid \mathcal{F}_t],$$

where the supremum is taken over all stopping times τ satisfying $t \leq \tau \leq T$.

The optimal stopping problem is closely tied to the valuation of many financial instruments, such as Bermudan swaptions, where the decision to exercise is made on a finite set of discrete dates, and the goal is to find the optimal exercise strategy to maximize the value of the option.

2.1.3 Itô Calculus

Because the paths of Brownian motion are almost surely nowhere differentiable, classical calculus cannot be applied. To analyze and model systems driven by Brownian dynamics, we instead rely on Itô calculus, which, as a framework, extends traditional calculus to continuous-time stochastic processes by defining stochastic integrals and introducing a compatible differential rule. The key mathematical concept that makes this extension possible is quadratic variation.

Definition 2.6 (Quadratic variation)**.** Let X_t be a continuous stochastic process. The quadratic variation of X_t over the interval $[0, t]$ is

$$\langle X, X\rangle_t = \lim_{||\Pi||\to 0} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2 \quad \text{in probability,}$$

where $\Pi = \{t_0, t_1, \ldots, t_n\}$ is a partition of $[0, T]$ and $0 < t_0 < t_1 < \cdots < t_n = T$.

More generally, the covariation of two different continuous processes X and Y is

$$\langle X, Y\rangle_t = \lim_{||\Pi||\to 0} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}),$$

For Brownian motion W_t, one can show that its quadratic variation is

$$\langle W, W\rangle_t = \lim_{n\to\infty} \sum_{i=1}^{n} (W_{t_i} - W_{t_{i-1}})^2 = t \quad \text{almost surely.}$$

Analogously, for correlated Brownian motions $W_t^{(1)}, W_t^{(2)}$ with instantaneous correlation ρ, we have

$$\langle W^1, W^2\rangle_t = \rho t,$$

which we refer to as $(dW_t)^2 = dt$.Since the quadratic variation of Brownian motion satisfies $\langle W, W\rangle_t = t$, the Itô multiplication rule follows:

$$(dW_t)^2 = dt.$$

Consequently, the following differential rules apply

- $dW_t dW_t = dt$
- $dW_t dt = 0$
- $dt dt = 0$.

As the sample paths of Brownian motion have unbounded variation (Mörters and Peres, 2010), classical Riemann-Stieltjes integration with respect to W_t is undefined. However, its finite quadratic variation allows one to define an integral, called the Itô integral, with a stochastic process as an integrand with respect to W_t. We provide a concise overview of how the Itô integral is constructed. For a complete and rigorous treatment, see Baldi (2017) or Øksendal (2003).

Let $M^2([0,T])$ denote the space of all real-valued progressively measurable processes H such that $\mathbb{E}\left[\int_0^T |H_t|^2 dt\right] < \infty$. Here, H is said to be progressively measurable with respect to $(\mathcal{F}_t)_{t\geq 0}$ if the mapping $H : [0,t]\times\Omega \to \mathbb{R}$ is $\text{Borel}([0,t])\otimes \mathcal{F}_t$-measurable for every t. The construction begins by defining the stochastic integral for elementary processes. For a partition $0 = t_0 < t_1 < \cdots < t_n = T$, an elementary process is of the form

$$H_t = \sum_{i=0}^{n-1} H_i \mathbf{1}_{[t_i, t_{i+1}[}(t),$$

where each H_i is $\mathcal{F}_{t_i}$-measurable and square-integrable. For such a process, the stochastic integral of H with respect to W over $[0,T]$ is defined as

$$\int_0^T H_t dW_t := \sum_{i=0}^{n-1} H_i (W_{t_{i+1}} - W_{t_i}).$$

A key property of this discrete integral follows from the independence of Brownian increments. Because the increments $W_{t_{i+1}} - W_{t_i}$ are independent of $\mathcal{F}_{t_i}$, the expectation of the cross-terms in the squared integral vanishes. This yields the following result, known as the Itô isometry.

Theorem 2.1 (Itô isometry)**.** *Let $H \in M^2([0,T])$ be an elementary process. Then,*

$$\mathbb{E}\left[\left(\int_0^T H_t dW_t\right)^2\right] = \mathbb{E}\left[\int_0^T H_t^2 dt\right].$$

Proof. See Lemma 7.1 in Baldi (2017). □

This isometry establishes that the discrete integral is an isometric mapping from the elementary processes of $M^2([0,T])$ to L^2. Since elementary processes are dense in $M^2([0,T])$ (see Lemma 7.2 in Baldi (2017)), the isometry extends to the whole space $M^2([0,T])$, defining the stochastic integral for every $H \in M^2([0,T])$.

Definition 2.7 (Itô integral)**.** For any progressively measurable and square-integrable stochastic process $H \in M^2([0,T])$, the Itô integral is defined as

$$I_t = \int_0^t H_s dW_s = \lim_{n\to\infty} \int_0^t H_s^{(n)} dW_s \quad \text{in } L^2, \quad 0 \leq t \leq T,$$

where $H^{(n)}$ is a sequence of elementary processes such that

$$\mathbb{E}\left[\int_0^t \left(H_s - H_s^{(n)}\right)^2 ds\right] \to 0.$$

By continuity of the norm in L^2, the Itô isometry established for simple processes extends to all general integrands in $M^2([0,T])$, as demonstrated in Theorem 7.1 of Baldi (2017). Furthermore, the Itô integral exhibits several desirable properties, such as the following (Shreve, 2004):

- The paths of I_t are continuous.
- For each t, I_t is $\mathcal{F}_t$-measurable.
- The quadratic variation of the integral is $\langle I, I\rangle_t = \int_0^t H_s^2 ds$.
- I_t is a martingale.

However, perhaps the most notable property is the Itô isometry, which provides a direct expression for computing the variance of Itô integrals.

With the Itô integral defined, we now define the most central class of processes in financial mathematics.

Definition 2.8 (Itô process)**.** A process $X = (X_t)_{t\geq 0}$ is an Itô process if it can be written as

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s,$$

where $\mu, \sigma \in L^2$ are adapted stochastic processes known as the drift and diffusion coefficients, where the drift μ represents the deterministic evolution of the process and the diffusion σ represents the magnitude of randomness driven by the Brownian motion W_t. As a shorthand notation in differential form, we can write

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

Virtually all stochastic processes, except those with jumps, are Itô processes. This naturally extends to finance, meaning that nearly every continuous-time model employed is an Itô process. For example, the Black-Scholes-Merton model (see Section 2.2.3) describes a stock price as a geometric Brownian motion, while interest rate models describe the stochastic evolution of rates using processes of this form.

The key result for manipulating such processes is Itô's lemma, the stochastic counterpart of the chain rule. Being the core formula of stochastic calculus, Itô's lemma is used to find the differential of a time-dependent function of a stochastic process in a relatively simple form.

Theorem 2.2 (Itô's lemma)**.** *For an Itô process $dX_t = \mu_t dt + \sigma_t dW_t$ and any twice differentiable scalar function $Y_t = f(t, X(t))$ of two real variables t and x, Y_t is itself an Itô process with stochastic differential*

$$dY_t = df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \frac{\partial f}{\partial x}(t, X_t)dX_t + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}(t, X_t)dt,$$

or equivalently

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}\right)dt + \sigma_t \frac{\partial f}{\partial x} dW_t. \tag{1}$$

The additional term involving $\frac{1}{2}\sigma_t^2 \partial_{xx} f$ arises from the non-zero quadratic variation of Brownian motion and is what distinguishes stochastic calculus from its deterministic counterpart. Below we provide a sketch of proof to demonstrate what is meant here.

Sketch of proof. The lemma can be derived heuristically by forming the second-order Taylor series expansion of the function around (t, X_t). Given an Itô process X_t that satisfies $dX_t = \mu_t dt + \sigma_t dW_t$, where $f(t, X_t)$ is a twice-differentiable scalar function, we note that its expansion in a Taylor series is

$$\frac{\Delta f(t)}{dt} dt = f(t + dt, x) - f(t, x) = \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \cdots$$

$$\frac{\Delta f(x)}{dx} dx = f(t, x + dx) - f(t, x) = \frac{\partial f}{\partial x} dx + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dx)^2 + \cdots .$$

Using the definition of the partial derivative,

$$f_y = \lim_{dy \to 0} \frac{\Delta f(y)}{dy}$$

we get

$$df = f_t dt + f_x dx = \lim_{dx \to 0, dt \to 0} \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{1}{2} \left(\frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial^2 f}{\partial x^2} (dx)^2 + \cdots \right).$$

We then substitute $x = X_t$. The differential of x then becomes $dx = dX_t = \mu_t dt + \sigma_t dW_t$, which yields

$$df = \lim_{dW_t \to 0, dt \to 0} \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} (\mu_t dt + \sigma_t dW_t) +$$
$$\frac{1}{2} \left[\frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial^2 f}{\partial x^2} (\mu_t^2 (dt)^2 + 2\mu_t \sigma_t dt dW_t + \sigma_t^2 (dW_t)^2 \right] + \cdots .$$

In the limit $dt \to 0$ the terms $(dt)^2$ and $dt dW_t$ tend to zero faster than dt. We can then use $dt dt = 0, dW_t dt = 0, dW_t dW_t = dt$ to get

$$df = \lim_{dt \to 0} \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dW_t$$

as desired. For a fully rigorous derivation, see Øksendal (2003) pp.46-48 or Shreve (2004) pp.138-141. □

We now consider the multi-dimensional case where we have a vector of Itô processes, i.e. an n-dimensional Itô process $\mathbf{X}_t = (X_t^{(1)}, \ldots, X_t^{(n)})$. We then have, in matrix notation,

$$d\mathbf{X}_t = \boldsymbol{\mu} dt + \boldsymbol{\sigma} d\mathbf{W}_t,$$

where

$$X_t = \begin{pmatrix} X_t^{(1)} \\ \vdots \\ X_t^{(n)} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nm} \end{pmatrix}, \quad dW_t = \begin{pmatrix} dW_t^{(1)} \\ \vdots \\ dW_t^{(m)} \end{pmatrix}.$$

Each component $X_t^{(i)}$ then has a stochastic differential of the form

$$dX_t^{(i)} = \mu_t^{(i)} dt + \sum_{j=1}^{m} \sigma_t^{(ij)} dW_t^{(j)},$$

where $W_t^{(1)}, \ldots, W_t^{(m)}$ are Brownian motions with correlation matrix ρ. The proof for the following theorem is similar to the one-dimensional version and is therefore omitted.

Theorem 2.3 (Multidimensional Itô's lemma)**.** *Let $d\boldsymbol{X}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\sigma}_t d\boldsymbol{W}_t$ be an n-dimensional Itô process as above. Let $f : [0, \infty) \times R^n \to R$ be a function in C^2. The process $Y_t = f(t, X_t)$ is then again an Itô process, whose $k-$th component is given by the differential*

$$dY_t^{(k)} = \frac{\partial f^{(k)}}{\partial t}(t, X_t)dt + \sum_{i=1}^{n} \frac{\partial f^{(k)}}{\partial x_i} dX_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f^{(k)}}{\partial x_i \partial x_j} dX_i dX_j,$$

where the differential products satisfy the Itô multiplication rules

$$\begin{cases} dW_i dW_j = \rho_{ij} dt, \\ dW_i dt = 0, \quad i = 1, \ldots, n, \\ (dt)^2 = 0. \end{cases}$$

The Itô integral, its properties, and Itô's lemma form the core foundation of stochastic differential calculus, the analytical framework for continuous-time finance. Itô processes arise as solutions of differential equations driven by Brownian motions called stochastic differential equations (SDEs), where the presence of Brownian motion (or a stochastic process in general) introduces an element of randomness into the dynamics. Consequently, SDEs themselves are stochastic processes as they have a stochastic process driving their evolution and are used to model the evolution of a stochastic process over time. A typical SDE is of the form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t,$$

where μ is the drift and σ is the diffusion coefficient. In integral form, the previous equation can be expressed as

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s,$$

where the first integral is a Riemann integral and the second is an Itô integral. An SDE therefore specifies that infinitesimal changes in X_t consist of a deterministic drift part and a stochastic diffusion part.

A particularly important SDE in finance is the geometric Brownian motion, given by

$$dX_t = \mu X_t dt + \sigma X_t dW_t,$$

where μ and σ are constants called drift and volatility, respectively. The drift parameter dictates whether the process tends to go up or down in the long run, whereas the volatility parameter represents the instantaneous standard deviation of returns of the process. Geometric Brownian motion is the equation for the dynamics of the price of a stock in the Black-Scholes-Merton model for option pricing (discussed in Section 2.2.3), due to which here we denote $X_t = S_t$. Applying Itô's lemma to $\ln S_t$, we can obtain for this the analytic solution

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right).$$

SDEs are the mathematical backbone of the models used in this thesis, allowing processes such as interest rates to be represented as SDEs driven by Brownian motions. Itô calculus provides the tools to manipulate SDEs, derive functions of them, and compute conditional expectations in pricing, providing a consistent framework for expressing and analysing stochastic interest-rate dynamics.

2.1.4 Risk-Neutral Pricing Framework

Risk-neutral pricing is a key concept in modern finance, providing the theoretical foundation for valuing assets and derivatives in arbitrage-free markets. Imagine a hypothetical world where all investors are assumed to be risk-neutral, caring only about expected payoffs and not about risk. In such a world, the fair price of a gamble is simply its expected value. For example, consider a coin toss where the player wins 100 if the coin lands on heads and 0 if tails. The expected payoff of this coin toss is 50 and in a risk-neutral world, everybody would play this because 50 is the fair price. This, of course, does not apply in the real world, as investors are generally considered risk averse and would need a positive expected net payoff to accept risk, which market prices reflect through risk premia. A risk-neutral player would play the game at the fair price of 50 for an expected net profit of 0, whereas a risk-averse player would require a positive expected net payoff to play the game. Although the assumption of universal risk neutrality in markets is unrealistic, it provides a powerful mathematical tool. Under a risk-neutral measure $\mathbb{Q}$, arbitrage-free prices can be expressed as expected discounted payoffs rather than requiring us to model risk aversion and market risk premia by explicitly incorporating them into the risk measure. The transformation from the probability measure $\mathbb{P}$ to $\mathbb{Q}$ thus allows prices to be represented as conditional expectations under which all discounted asset prices are martingales.

Definition 2.9 (Risk-neutral probability measure)**.** A risk-neutral probability measure (or equivalent martingale measure) $\mathbb{Q} \sim \mathbb{P}$ has the property that every traded asset, when discounted by the risk-free rate, with the price process S_t is a martingale under $\mathbb{Q}$. Here, the equivalence of measures ($\mathbb{Q} \sim \mathbb{P}$) means that they share the same null sets, i.e., for every $A \in \mathcal{F}$, $\mathbb{P}(A) = 0$ if and only if $\mathbb{Q}(A) = 0$.

This principle is formalized by the fundamental theorems of asset pricing (Delbaen and Schachermayer, 1994), considered one of the most important results in mathematical finance. However, before stating the theorems, we formalize key concepts, such as the market and what constitutes a valid trading strategy.

Definition 2.10 (Financial market)**.** A financial market is an $\mathcal{F}_t^{(m)}$-adapted $(n+1)$-dimensional Itô process $S_t = (S_t^{(0)}, S_t^{(1)}, \ldots, S_t^{(n)})$, $0 \le t \le T$. The risk-free asset, or money-market account, satisfies

$$dS_t^{(0)} = r_t S_t^{(0)} dt, \quad S_0^{(0)} = 1,$$

and the risky assets numbered $1, \ldots, n$ follow Itô processes of form

$$dS_t^{(i)} = \mu_t^{(i)} dt + \sum_{j=1}^{m} \sigma_t^{(ij)} dW_t^{(j)} = \mu_t^{(i)} dt + \sigma_t^{(i)} dW_t, \quad S_0^{(i)} = s_i,$$

where σ_i is the row number i of the $n \times m$ matrix $[\sigma_t^{(ij)}]$, $1 \le i \le n$. The assets $S_t^{(i)}$, $i = 1, \ldots, n$, are called risky assets because their dynamics contain a diffusion term $\sigma_t^{(i)}$.

Definition 2.11 (Portfolio and Value Process)**.** A portfolio (or trading strategy) in this market over the trading interval $[0, T]$ is defined as an $(n+1)$-dimensional, progressively measurable stochastic process

$$\phi_t = (\phi_t^{(0)}, \phi_t^{(1)}, \ldots, \phi_t^{(n)}),$$

where the components $\phi_t^{(i)}$ represent the number of units held in the i-th asset at time t. Its corresponding value process V_t^{ϕ} is defined as

$$V_t^{\phi} = \phi_t \cdot S_t = \sum_{i=0}^{n} \phi_t^{(i)} S_t^{(i)}, \quad t \in [0, T],$$

where $\cdot$ denotes the inner product in $\mathbb{R}^{n+1}$. The initial value of the portfolio V_0^{ϕ} represents the initial investment.

Definition 2.12 (Self-financing portfolio)**.** The portfolio ϕ_t is called self-financing if its value process satisfies

$$dV_t^{\phi} = \phi_t \cdot dS_t = \sum_{i=0}^{n} \phi_t^{(i)} dS_t^{(i)}, \qquad 0 \le t \le T.$$

That is, a portfolio is self-financing if its value changes only through profits and losses on the held assets without additional capital deposits or withdrawals.

Definition 2.13 (Admissible portfolio)**.** A self-financing portfolio ϕ_t is called admissible if its value process is almost surely lower bounded from below, i.e. there exists a constant $C \ge 0$ such that

$$V_t^{\phi} \ge -C \quad \text{for all } t \in [0, T] \quad \text{almost surely.}$$

This lower bound rules out strategies with unbounded losses.

Definition 2.14 (Arbitrage opportunity)**.** An admissible portfolio ϕ_t over $[0, T]$ is called an arbitrage opportunity if its value process satisfies

$$V_0^\phi = 0, \quad V_T^\phi \geq 0 \quad \text{a.s.}, \quad \text{and} \quad \mathbb{P}[V_T^\phi > 0] > 0.$$

In other words, an arbitrage opportunity is a portfolio that requires zero initial capital, carries absolutely no risk of a terminal loss, yet offers a strictly positive probability of generating a profit without the possibility of loss.

Definition 2.15 (Arbitrage-free market)**.** A market model is arbitrage-free if there exists no admissible portfolio on $[0, T]$ that is an arbitrage opportunity.

Definition 2.16 (Contingent Claim)**.** A (European) contingent claim with maturity T, or $T-$claim, is a (lower) bounded $\mathcal{F}_T$-measurable random variable X. A contingent claim X with maturity T is called attainable if there exists an admissible portfolio ϕ such that $V_T^\phi = X$. Such a strategy ϕ is said to replicate, or hedge, X.

Definition 2.17 (Complete market)**.** A market model is complete if every bounded $T-$claim with payoff X is attainable.

With these definitions in hand, we now state the fundamental theorems of asset pricing which connect the absence of arbitrage and market completeness to the existence and uniqueness of risk-neutral measures.

Theorem 2.4 (**First Fundamental Theorem of Asset Pricing**)**.** *A market model is arbitrage-free if and only if there exists a risk-neutral probability measure* $\mathbb{Q}$*.*

Proof. See Theorem 5.4.7 in Shreve (2004), p. 231. □

Theorem 2.5 (**Second Fundamental Theorem of Asset Pricing**)**.** *Assuming a market model that is arbitrage-free, the market is complete if and only if the risk-neutral probability measure* $\mathbb{Q}$ *is unique.*

Proof. See Theorem 5.4.9 in Shreve (2004), pp. 232-234. □

The Fundamental Theorems provide a link between economic consistency and probabilistic representation: the first theorem establishes the equivalence between the absence of arbitrage and the existence of an equivalent martingale measure, while the second ensures that completeness implies its uniqueness. Consequently, every arbitrage-free price process can be represented as a martingale under an appropriate measure, providing the theoretical basis for the risk-neutral valuation framework used throughout this thesis.

Building on the above, the pricing of a derivative can be reduced to a conditional expectation of its discounted payoff under $\mathbb{Q}$. The price of a derivative V_t at time $t \leq T$ can therefore be expressed as

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s ds \right) V_T \middle| \mathcal{F}_t \right], \tag{2}$$

where r_s is the risk-free interest rate and V_T is the derivative payoff at maturity T.

This formula (2) is written under the assumption that we use the money-market account, a deposit account that grows at the risk-free rate, as the unit of account for pricing. However, this choice is not unique, and other assets can be used in its place. Any such asset that is a strictly positive tradable asset used as a reference unit is called a numéraire. Just as we use euros to value goods and services or meters to measure distances, a numéraire is used to value other financial assets. If N_t is a numéraire, the relative price of another asset S_t can be expressed as the ratio $\frac{S_t}{N_t}$. We note a few commonly used numéraires or their respective measures below.

Definition 2.18 (Money-market account (Bank account))**.** The numéraire used by the standard risk-neutral measure $\mathbb{Q}$. Under this measure, discounted asset prices are martingales. Let B_t denote the value of a money-market account at time $t \geq 0$, where we assume that the evolution of the account is given by

$$dB_t = r_t B_t dt, \quad B_0 = 1,$$

where $r_t \geq 0$ is a function of time. The money-market account is then

$$B_t = \exp\left(\int_0^t r_s ds\right).$$

This definition implies that a single unit of currency deposited at time 0 accumulates to $B_t = \exp\left(\int_0^t r_s ds\right)$ at time t, where r_t is the instantaneous rate at which the bank account grows. This quantity is consequently known as the short rate, representing the theoretical interest rate at which money can be borrowed over an infinitesimally short period of time.

Definition 2.19 (T-forward measure)**.** Introduced by Jamshidian (1989), the T-forward measure uses a zero-coupon bond $P(t,T)$ (see Section 2.2.1) whose maturity T coincides with that of the derivative to price as the numéraire. This leads to the T-forward measure, denoted $\mathbb{Q}^T$, under which any asset price divided by $P(t,T)$ is a martingale. This is especially helpful for some assets, as any simply-compounded forward rate through a time interval ending in T is a martingale under this measure (see Proposition 2.5.1 Brigo and Mercurio (2001)).

Definition 2.20 (Swap measure)**.** The swap annuity is defined as the present value of the fixed leg of a swap (see Section 2.2.2). Choosing the annuity as the numéraire defines the associated swap measure, denoted $\mathbb{Q}^A$, under which the forward swap rate is a martingale. This is particularly useful for pricing European swaptions, as it simplifies the valuation by expressing the price of the complex derivative relative to the annuity numéraire.

We can move from one numéraire to another via the change of numéraire technique. The idea is that if N_t is a strictly positive traded asset, then there exists an equivalent probability measure $\mathbb{Q}^N$ under which every price expressed in terms of N_t is a martingale. That is, we move from one pricing measure to another by choosing a

different numéraire, and the ratio of the prices of two assets when discounted by a numéraire is a martingale. More precisely, the Radon-Nikodym derivative between $\mathbb{Q}^N$ and the risk-neutral measure $\mathbb{Q}$ is

$$\left.\frac{d\mathbb{Q}^N}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = \frac{N_t/B_t}{N_0/B_0}.$$

Under $\mathbb{Q}^N$, the arbitrage-free price of a payoff H_T can be written as

$$V_t = N_t \mathbb{E}^{\mathbb{Q}^N}\left[\frac{H_T}{N_T}\middle|\mathcal{F}_t\right].$$

This generalizes the risk-neutral valuation formula, which is recovered when $N_t = B_t$.

When asset price dynamics are specified under one measure and we wish to understand their form under another, Girsanov's theorem is required. This theorem is a fundamental result in stochastic calculus as it provides a mathematical framework for transitioning between equivalent probability measures. Recall that under $\mathbb{P}$, asset prices incorporate a drift reflecting real-world expected returns and risk premia, whereas under $\mathbb{Q}$, discounted asset prices are martingales. While the existence of $\mathbb{Q}$ is guaranteed by the First Fundamental Theorem of Asset Pricing, Girsanov's theorem tells us precisely how the two measures are related. It shows that shifting between equivalent probability measures adjusts the drift in the underlying Brownian motion while leaving the volatility structure unchanged. This makes it possible to quantify the drift adjustment required to move from $\mathbb{P}$ to $\mathbb{Q}$, forming the foundation of risk-neutral pricing frameworks such as the Black-Scholes formula (see Section 2.2.3).

Theorem 2.6 (Girsanov's Theorem)**.** *Let $(W_t)_{t\geq 0}$ be a Brownian motion under $\mathbb{P}$, and let $(\theta_t)_{t\geq 0}$ be an adapted process satisfying Novikov's condition*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \theta_s^2 ds\right)\right] < \infty.$$

Then the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^T \theta_s dW_s - \frac{1}{2}\int_0^T \theta_s^2 ds\right)$$

defines an equivalent probability measure $\mathbb{Q} \sim \mathbb{P}$. Under the new measure $\mathbb{Q}$, the process

$$W_t^{\mathbb{Q}} = W_t + \int_0^t \theta_s ds$$

is a $\mathbb{Q}$-Brownian motion.

Proof. See Theorem 5.2.3 in Shreve (2004), pp.212-214. □

Remark 1. Novikov's theorem is crucial so that the exponential process in the Radon-Nikodym derivative is well-defined. This ensures that the probability mass under the new probability measure $\mathbb{Q}$ remains 1, validating the measure change so that the new measure is a probability measure (as $\mathbb{E}(Z(T)) = 1$, where $Z(T) = \exp\left(-\int_0^T \theta_s dW_s - \frac{1}{2}\int_0^T \theta_s^2 ds\right)$.

In essence, Girsanov's theorem formalizes how a transition between equivalent probability measures is captured by a specific adjustment to the drift without changing the volatility structure. In practice, this means that when moving from $\mathbb{P}$ to $\mathbb{Q}$, the risk premium is removed from the asset price dynamics. By applying the correct drift adjustment, Girsanov's theorem permits the transition between the risk-neutral, forward, and swap measures, ensuring that prices expressed in terms of the chosen numéraire are martingales under the corresponding measure.

2.2 Interest Rate Derivatives

Interest rate derivatives are financial contracts whose value is derived from the level of interest rates, which are rates outside of the control of the parties of the contract, commonly reference rates such as Euribor (Euro Interbank Offered Rate) and SOFR (Secured Overnight Financing Rate). Before SOFR, LIBOR (London Interbank Offered Rate) was a primary benchmark, along with Euribor, for short-term interest rates but was terminated in 2023 after a decision to discontinue in 2021 due to concerns of manipulation and its accuracy (Klingler and Syrstad, 2021). Interest rate derivatives form the foundation of the fixed-income market. They are among the most actively traded financial instruments, accounting for 79% of the total global notional value of OTC derivatives (International Swaps and Derivatives Association (ISDA), 2024). They allow institutions to hedge interest rate exposures, manage the duration of portfolios, structure complex funding or investment strategies, among many other purposes. Their valuation is highly connected to the entire term structure of interest rates, which describes the relationship between interest rates and the time to maturity.

This section establishes the framework for modeling and valuing interest rate derivatives. The most fundamental concepts of the term structure, such as market conventions, zero-coupon bonds, spot rates, and forward rates, are first introduced. These describe how interest rates are represented and evolve over time, forming the foundation for the valuation of interest rate swaps, the most fundamental linear derivatives from which more complex instruments are constructed. Afterwards, the framework for option pricing is introduced and applied to swaptions, which grant the right (but no obligation) to enter a swap at a predetermined rate.

Market dynamics and the volatility smile are then discussed, linking theory to observed market data and highlighting the limitations of classical models that assume a constant volatility across strikes. Finally, Bermudan swaptions are introduced alongside their key features and differences from European swaptions.

2.2.1 Fundamental Interest Rate Concepts

Market conventions dictate how interest rates and cash flows are defined in practice, providing consistency in the valuation, curve construction, and comparison of instruments across markets and currencies.

The day-count convention specifies how the time elapsed between two dates, t and T, is measured and expressed as a fraction of a year. The three most common

day-count conventions are as follows. For a more comprehensive discussion on day-count conventions, see Miron and Swannell (1991).

- Actual/365. Each year consists of 365 days, so that for two dates $D_1 = (d_1, m_1, y_1)$ and $D_2 = (d_2, m_2, y_2)$ the year fraction is $\frac{D_2 - D_1}{365}$, where $D_2 - D_1$ denotes the actual number of days between them. Standard in SONIA and sterling markets.
- Actual/360. Each year consists of 360 days, giving a year fraction of $\frac{D_2 - D_1}{360}$. Typical for Euribor-based instruments.
- 30/360. Each month consists of 30 days, resulting in years consisting of 360 days. The year fraction between D_1 and D_2 is then given by

$$\frac{\max(30 - d_1, 0) + \min(d_2, 30) + 360(y_2 - y_1) + 30(m_2 - m_1 - 1)}{360}.$$

  This convention is typical for corporate and government bonds.

In general, the accrual factor is therefore computed as

$$\delta(D_1, D_2) = \frac{\text{days between } D_1 \text{ and } D_2}{\text{basis days per year}}.$$

Payment frequencies are typically annual, semi-annual, or quarterly, with dates adjusted for the weekends and holidays according to business-day conventions such as modified following (MF) or preceding. Floating-rate payments reference benchmark interest rate indices, such as Euribor, each of which defines its own day-count basis and fixing calendar.

The valuation of interest rate derivatives is largely determined by the term structure of interest rates, which is the relationship between interest rates or bond yields and different terms or maturities. Typically represented by a yield curve, the term structure is formalized using zero-coupon bond prices, or discount factors. These factors quantify the present value of one unit of currency payable at a given maturity.

Definition 2.21 (Zero-Coupon Bond)**.** A zero-coupon bond is a financial contract that pays a single unit of currency, called the face value, at maturity T, with no intermediate cash flows. The price of the bond at time $t \leq T$ is denoted by $P(t, T)$, representing the market value of receiving 1 unit of currency at time T. We assume that $P(T, T) = 1$ for all T to avoid arbitrage. Under $\mathbb{Q}$, its price is given by the risk-neutral valuation formula

$$P(t, T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(- \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right],$$

where r_t is the short-rate process. The set $\{P(t, T)\}_{T \geq t}$ forms the discount-factor curve, where $P(t, T)$ generally decreases with T due to the time value of money.

Remark 2 (Zero-Coupon Bond vs. Discount Factor). The terms discount factor and zero-coupon bond are often used interchangeably, yet rigorously they represent slightly

different concepts. As stated in Brigo and Mercurio (2001), the discount factor $D(t,T)$ denotes the equivalent amount of currency at time t corresponding to one unit of currency to be received at time T, expressing time value in a purely numerical sense, while zero-coupon bond prices $P(t,T)$ represent the market value of a contract delivering one unit at time T. If interest rates are deterministic, $D(t,T) = P(t,T)$ for all $t \leq T$. However, when rates are stochastic, $D(t,T)$ becomes a random quantity at t that depends on how rates evolve in the future between t and T, whereas $P(t,T)$ is its deterministic expectation at time t $P(t,T) = \mathbb{E}^{\mathbb{Q}}[D(t,T) \mid \mathcal{F}_t]$. As such, the discount factor can be viewed as the instantaneous random equivalent of one unit of future currency, while the zero-coupon bond price is the market value of that future payoff at time t. We use the notation $P(t,T)$ for both concepts in this thesis, as the distinction is negligible here.

Zero-coupon bond prices are the fundamental quantities in interest rate theory, as all interest rates can be expressed using them. However, zero-coupon bonds are theoretical constructs that cannot be directly observed in the market, and so interest rates themselves are quoted instead. To translate these quoted rates into zero-coupon bond prices, a compounding convention must be specified.The three most common compounding types are continuous, simple, and k-times-per-year, each of which can be defined directly in terms of zero-coupon bond prices.

- A continuously-compounded spot rate $R(t,T)$ assumes that interest accrues continuously over time, and is the standard convention utilized in continuous-time finance. It is defined through the relation

$$P(t,T) = \exp(-R(t,T)(T-t)) \iff R(t,T) = -\frac{\ln P(t,T)}{T-t}.$$

- A simply-compounded spot rate assumes that interest accrues proportionally to the time period without compounding on itself. It satisfies

$$P(t,T) = \frac{1}{1 + L(t,T)(T-t)} \iff L(t,T) = \frac{1 - P(t,T)}{(T-t)P(t,T)}.$$

- k-times compounded spot rate assumes that interest is compounded periodically k times per year, e.g. $k = 2$ for semiannual compounding. It is defined by

$$P(t,T) = \frac{1}{\left(1 + \frac{Y^k(t,T)}{k}\right)^{k(T-t)}} \iff Y^k(t,T) = k\left(P(t,T)^{-\frac{1}{k(T-t)}} - 1\right)$$

A key property of $k-$times-per-year compounding is that it converges to continuous compounding as the frequency k approaches infinity. That is,

$$\lim_{k \to +\infty} \frac{k}{[P(t,T)]^{1/(k(T-t))}} - k = -\frac{\ln(P(t,T))}{T-t} = R(t,T).$$

As a result, for each fixed Y,

$$\lim_{k\to+\infty}\left(1+\frac{Y}{k}\right)^{k(T-t)} = e^{Y(T-t)}.$$

In infinitesimal time intervals, the previous definitions of spot rates are equivalent. That is, for each t, the short rate is

$$r_t = \lim_{T\to t^+} R(t,T) = \lim_{T\to t^+} L(t,T) = \lim_{T\to t^+} Y^k(t,T) \quad \text{for each } k.$$

A natural continuation from spot rates is the concept of forward rates. They represent interest rates that are fixed today for a future investment period and are determined directly by the current term structure. In other words, the forward rate is the future yield on a bond, calculated from the yield curve, connecting the short rate to the term structure. The simply-compounded forward rate for the period $[T,S]$ at time t is

$$F(t;T,S) = \frac{1}{S-T}\left(\frac{P(t,T)}{P(t,S)} - 1\right), \tag{3}$$

which can equivalently be written as

$$1 + (S-T)F(t;T,S) = \frac{P(t,T)}{P(t,S)}.$$

With continuous compounding, the corresponding forward rate for the same period is

$$f(t;T,S) = \frac{1}{S-T}\ln\frac{P(t,T)}{P(t,S)}.$$

As the maturity S approaches the expiry T, both the simply-compounded and continuously-compounded forward rates converge to the instantaneous forward rate, obtained by taking the limit

$$\lim_{S\to T+} F(t;T,S) = -\lim_{S\to T+}\frac{1}{S-T}\left(\frac{P(t,T)}{P(t,S)} - 1\right) = -\frac{1}{P(t,T)}\frac{\partial P(t,T)}{\partial T} = -\frac{\partial \ln P(t,T)}{\partial T}.$$

Definition 2.22 (Instantaneous forward interest rate)**.** For maturity T, the instantaneous forward rate at time t is given by

$$f(t,T) = \lim_{S\to T^+} F(t;T,S) = -\frac{\partial \ln P(t,T)}{\partial T}.$$

The function $T \mapsto f(t,T)$ is called the forward curve at time t. The zero-coupon bond prices can be expressed in terms of the instantaneous forward rates as

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right).$$

By letting the maturity T approach the current time t, we obtain the instantaneous short rate.

Definition 2.23 (Instantaneous short rate)**.** The instantaneous short rate at time t is given by

$$r_t = f(t,t) = \lim_{T\to t^+} R(t,T).$$

2.2.2 Interest Rate Swaps

Interest rate swaps (IRSs) are contracts agreed between two parties to exchange periodic interest rate payments over a specified period based on a notional principal amount. The most common type is a fixed-for-floating swap in which one party pays a fixed interest rate on the notional principal and receives a floating interest rate based on a reference rate such as Euribor or SOFR. Swaps are widely used for hedging interest rate risk at a low cost, transforming fixed-rate liabilities into floating-rate ones or vice versa, among other purposes. Swap rate quotes are available up to 60 years, though maturities from 1 to 30 years are often considered standard.

A swap is called a payer swap if the holder of the contract pays the fixed leg and receives the floating leg, and it is specified by the notional amount N, the fixed rate K, a set of future dates $T_0 < T_1 \cdots < T_n$ with $T_i - T_{i-1} = \delta$, and the discount factor $P(t, T_i)$ to a payment date T_i. The dates, and hence the day-count fractions δ, being equidistant is not necessary and is assumed mostly for convenience here. Cash flows occur on coupon dates $T_1, \ldots, T_n$, and the value of an interest rate swap at time t is determined by considering the valuation of the floating and fixed legs separately. At each T_i, the holder of the contract pays a fixed amount δNK, and receives a floating amount $\delta NF(T_{i-1}, T_i)$, where $F(T_{i-1}, T_i) = F(T_{i-1}; T_{i-1}, T_i)$ denotes the simply-compounded forward rate for the period $[T_{i-1}, T_i]$ at time T_{i-1}. As a result, the net cash flow at T_i is

$$\delta N(F(T_{i-1}, T_i) - K). \tag{4}$$

Using the definition of forward rates (3), the value of this cash flow at $t \leq T_0$ can be computed as

$$N(P(t, T_{i-1}) - P(t, T_i) - K\delta P(t, T_i)),$$

and therefore the total value of the swap $V_{\text{swap}}(t)$ at time $t \leq T_0$ is

$$V_{\text{swap}}(t) = N\left(P(t, T_0) - P(t, T_n) - K\delta \sum_{i=1}^{n} P(t, T_i)\right). \tag{5}$$

A receiver swap is exactly the same, except the cash flows are switched so that the holder now pays the floating leg and receives the fixed leg. The value of a receiver swap is therefore identical to that of a payer swap, only that we change the sign of the cash flows at times $T_1, \ldots, T_n$, and hence the value at $t \leq T_0$ is

$$V_{\text{receiver}}(t) = -V_{\text{swap}}(t).$$

The forward swap rate (or the par swap rate) $R_{\text{swap}}(t)$ at time $t \leq T_0$ is the fixed rate that sets the fixed leg equal to the floating leg and the net present value of the swap equal to zero $V_{\text{swap}}(t) = V_{\text{receiver}}(t) = 0$, making the contract fair for both parties. The swap rate is defined as

$$R_{\text{swap}}(t) = \frac{P(t, T_0) - P(t, T_n)}{A(t)}, \tag{6}$$

where the denominator $A(t) = \delta \sum_{i=1}^{n} P(t, T_i)$ is called the swap annuity, which represents the value of one unit of fixed-leg payments. Using the annuity and swap rate, the value of a swap can be expressed equivalently as

$$V_{\text{swap}}(t) = NA(t)(R_{\text{swap}}(t) - K).$$

Remark 3. The above formulas follow the single-curve framework, where a single curve is used for both discounting and projection. In practice, a dual-curve OIS-LIBOR framework is used post-financial crisis, where discounting is performed using an OIS curve while the projected forward rates are derived from a separate projection curve that mirrors the reference rate in the floating leg. Consequently, the valuation formulas look slightly different in the dual-curve framework than they do in the single-curve framework. However, the literature universally presents all formulas in the single-curve framework and since the dual-curve adjustments do not materially affect the theoretical results derived, all formulas in this thesis are stated in the single-curve framework. Nonetheless, the dual-curve framework is employed in the numerical parts of this thesis, as described in Section 4.2.

2.2.3 Fundamentals of Option Valuation

Options are a fundamental class of financial derivatives whose distinctive feature is their asymmetric payoff. While swaps or forwards involve an obligation to exchange cash flows making their payoff structure linear, an option grants its holder the right - but no obligation - to engage in a transaction at a predetermined price. This asymmetric payoff structure introduces convexity and makes option valuation inherently nonlinear, relying on expectations under a risk-neutral measure instead of straightforward deterministic discounting. As the holder possesses a right to buy without an obligation, the holder must pay a premium to the seller to compensate for the flexibility. The option premium is the price paid at initiation of the contract to obtain the potentially infinite upside of the position while limiting downside risk. As such, the premium represents the market's assessment of the value of optionality, primarily driven by the volatility of the underlying as well as the time to maturity.

An option is specified by its underlying asset S_t, the strike price K, and the maturity T. At maturity, the holder decides whether exercising the option is advantageous or not, depending on the payoff the holder were to receive at the time. This payoff depends on whether the option is a call or a put option - a call option gives the holder the right to buy the underlying asset at the strike price K, whereas a put option grants the right to sell it. Formally, the payoffs are defined as follows.

Definition 2.24 (Call and put payoffs)**.** For maturity T and strike K, the payoffs of a European call and put are

$$\text{Call: } (S_t - K)^+ = \max(S_t - K, 0), \quad \text{Put:} (K - S_t)^+ = \max(K - S_t, 0).$$

As can be seen, a long position in a call benefits from a rise in the underlying, while a long position in a put benefits from a fall in the underlying. Conversely, a short position

represents the obligation to deliver the corresponding payoff to the holder - being short on an option means acquiring the premium from the buyer, and is beneficial if markets remain calm and the option expires out of the money.

At time $t \leq t_0$, a call option with strike price K is said to be at the money (ATM), in the money (ITM), out of the money (OTM) if

$$K = S_t, \quad K < S_t, \quad K > S_t,$$

respectively.

Remark 4 (Option styles). The above holds for European options where the holder may only exercise the option at maturity. However, there are many other types of options as well, such as American options, which may be exercised at any point during the lifetime of the option, as well as Bermudan and Canary options, which can be exercised on a predetermined set of dates. Bermudan options can be exercised more frequently than Canary options. For more information on different option styles and their payoffs and valuations, see Haug (2007) or Hull and Basu (2016).

Options are difficult to value, as the value of the contract depends on more than simply the value of the underlying asset. In practice, the value of an option is usually decomposed into two parts. The intrinsic value represents the immediate benefit of exercising the option and is defined as the positive difference between the current underlying price and the strike price, while the time value captures the additional value of maintaining the right to exercise later, reflecting uncertainty about future market conditions. The time value primarily depends on volatility, the time to maturity, and the prevailing interest rates. Higher volatility increases the probability and magnitude of favorable outcomes, increasing the expected payoff. Interest rates, in turn, affect the time value through both discounting and their influence on the expected forward value of the underlying. Indeed, an increase in interest rates tends to raise the time value of options whose payoff benefits from higher underlying prices (calls) and reduce it for those that benefit from lower prices (puts).

The put-call parity defines a model-independent relationship between the price of European calls and puts on the same underlying with identical strike prices and maturities. Namely, it dictates that a portfolio of a long call option and a short put option is equivalent to a single forward at this strike price and expiry. The parity simply dictates that if the price at maturity is above the strike, the call will be exercised, while the put will be exercised if the price is below the strike at maturity. The put-call parity is formulated as follows.

Proposition 2.1. *For any European call C_t and put P_t with identical strike price K and maturity T,*

$$C_t - P_t = S_t - KP(t,T),$$

where $P(t,T)$ is the risk-free discount factor (i.e. $KP(t,T)$ represents the present value of the strike price at the current risk-free rate). The risk-free rate is often assumed to be constant, in which case $P(t,T) = e^{-r(T-t)}$.

Proof. A portfolio long one call option and short one put option yields a terminal payoff $S_t - K$. If $S_t > K$, the call is exercised and the payoff is $S_t - K$. If $S_t < K$, the put is exercised against the seller, in which case the payoff still equals $S_t - K$. In both cases the payoff is therefore $S_t - K$, and the situation is identical to a forward contract delivering one unit of the underlying for K at T. By no-arbitrage, the value of the portfolio must thus be equal to that of the equivalent forward, $S_t - KP(t,T)$. □

The put-call parity establishes a crucial no-arbitrage constraint between call and put prices. Any deviation would then allow arbitrage opportunities through synthetic replication, where one option could be reconstructed using the other together with positions in the underlying and a bond. The parity also ensures that the implied volatilities of European calls and puts with identical strike and maturity are equal when priced with any arbitrage-free model.

One of the most important parts of modern option theory is the Black-Scholes-Merton (BSM) model (Black and Scholes, 1973; Merton, 1973), which provides a closed-form solution for the price of European options. The model makes the following assumptions:

- The risk-free rate r is constant and known.
- The underlying asset price follows a geometric Brownian motion with constant drift μ and volatility σ,

$$dS_t = S_t \mu dt + \sigma S_t dW_t.$$

- The market is frictionless, i.e. transactions do not incur any fees or costs.
- There are no arbitrage opportunities, i.e. there is no way to make a riskless profit in excess of the risk-free rate.
- The stock does not pay dividends (an extension can be added to the model to accommodate this)
- The underlying is perfectly divisible and short selling is permitted.

Theorem 2.7 (Black-Scholes-Merton formula)**.** *Under the assumptions, the value of a European call option C_t at time t is*

$$C(S_t, t) = N(d_+)S_t - N(d\_)Ke^{-r(T-t)},$$

where

$$d_\pm = \frac{\ln(S_t/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

and $N(\cdot)$ denotes the standard normal cumulative distribution function. Via the put-call parity, the price of a corresponding put option is $P_t = Ke^{-r(T-t)}N(-d\_) - N(-d_+)S_t$.

Proof. See Björk (2009), pp.104-105. □

Although the BSM model assumes constant volatility, markets typically quote options in terms of their implied volatility - the volatility that, when inserted into the formula, reproduces the observed market price. As all other inputs (underlying price S_t, strike K, maturity T and interest rate r) are directly observable, implied volatility provides a convenient summary of how the market prices uncertainty.

The existence and uniqueness of this quantity follow from a standard monotonicity argument. Since vega

$$\frac{\partial C_t}{\partial \sigma} = S_t\sqrt{T-t}n(d_+) > 0 \quad \text{for all } \ t < T$$

where $n(\cdot)$ denotes the standard normal probability density function, the Black-Scholes-Merton price is continuous and strictly increasing in σ, mapping onto the full no-arbitrage interval as σ ranges over $(0, \infty)$. Consequently, for any market price $C^M(S_t, K, T)$ satisfying

$$\max\left(0, S_t - Ke^{-r(T-t)}\right) \leq C^M(S_t, K, T) \leq S_t,$$

there exists a unique implied volatility σ_{imp} such that

$$C_{\text{BSM}}(S_t, K, T-t, r, \sigma_{\text{imp}}(K, T)) = C^M(S_t, K, T).$$

Repeating this inversion across all traded strikes and maturities defines the implied volatility surface

$$\sigma_{\text{imp}} : (K, T) \mapsto \sigma_{\text{imp}}(K, T),$$

providing a one-to-one correspondence between observed prices and their implied volatilities within the no-arbitrage region. If the BSM model held exactly in the markets, this surface would be flat across strikes and maturities since the assumptions of constant volatility and log-normal returns leave no room for strike or maturity dependence. However, in practice, asset return distributions tend to be skewed and exhibit fat tails, and the surface is neither flat nor stable over time. These deviations are known as volatility smiles (symmetric curvature in implied vol across strikes) and skews (asymmetric tilt), whose implications for derivative pricing are examined in Section 2.2.5.

2.2.4 European Swaptions

Building on the two previous sections, we introduce the derivatives called swaptions (or swap options), which are options on an interest rate swap. More precisely, a European swaption with strike rate K grants the holder the right, but no obligation, to enter a swap with fixed rate K at a predetermined date in the future (the swaption maturity T_0) under specified terms. A payer swaption gives the right to enter a payer swap at the swaption maturity to pay fixed and receive floating, while a receiver swaption gives the right to enter a receiver swap to receive fixed and pay floating. The swaption maturity is commonly determined to coincide with the first reset date of the

underlying swap. The length of the underlying swap $T_n - T_0$ is called the tenor of the swaption. Swaps are denoted in the form maturity × tenor, i.e. $T_0 \times (T_n - T_0)$. For instance, a $5Y \times 10Y$ swaption means that, in 5 years at maturity T_0, the holder has the right (without obligation) to enter a 10-year swap.

Assuming that the underlying swap starts on the maturity date T_0 of the option, from (5) we get the payoff for a payer swaption with strike K at maturity T_0 as

$$V_{\text{swaption}}(T_0) = V_{\text{swap}}(T_0)^+ = N\left(\sum_{i=1}^{n} P(T_0, T_i)\delta(F(T_0; T_{i-1}, T_i) - K)\right)^+. \tag{7}$$

Note that, as the summation is inside the positive part operator instead of outside, the payoff cannot be decomposed into more elementary products. As a result, the stochastic dependence between different forward rates must be taken into account in the valuation.

At time $t \leq T_0$, the payer swaption with strike rate K is said to be at the money (ATM), in the money (ITM), out of the money (OTM) if

$$K = R_{\text{swap}}(t), \quad K < R_{\text{swap}}(t), \quad K > R_{\text{swap}}(t),$$

respectively. Here $R_{\text{swap}}(t)$ is the same as in (6). For a receiver swaption, we switch the inequality signs for the above to hold equivalently.

Under the risk-neutral valuation framework, using the change of numéraire method, we can switch to the annuity measure (or swap measure) $\mathbb{Q}^A$. Under this, the value of a European payer swaption at time $t \leq T_0$ can be expressed as

$$V_{\text{swaption}}(t) = NA(t)\mathbb{E}_t^A(R_{\text{swap}}(T_0) - K)^+,$$

where the forward swap rate is now a martingale under the swap measure.

From the above formula, we can see that a payer European swaption is a call option, and a receiver swaption is a put option, on the forward swap rate with the strike at the fixed rate of the swap K. Similarly to options, swaptions can be priced using a vanilla model similar to the Black-Scholes-Merton model called the Black model (Black, 1976).

Definition 2.25 (Black's formula for European swaptions)**.** Assuming a constant risk-free rate r, the price of a payer swaption V_{swaption} at time $t \leq T_0$ is

$$V_{\text{swaption}}(t) = NA(t)(R_{\text{swap}}(t)N(d_+) - KN(d_-)),$$

where

$$d_\pm = \frac{\ln\left(\frac{R_{\text{swap}}(t)}{K}\right) \pm \frac{1}{2}\sigma^2(T_0 - t)}{\sigma\sqrt{T_0 - t}}.$$

Here, the constant σ is known as Black's swaption volatility.

As is the case with options, it is market practice to quote European swaption prices in terms of implied Black volatilities, i.e. the volatilities that recover the market price when using the Black formula. The swaptions market is highly liquid with a great variety of option maturities and swap underlyings traded actively, where the quotes are organised by expiry and tenor. Market quotes on swaptions in terms of implied volatilities provide the most accessible information on the volatility structure of interest rates in the form of volatility surfaces, which consider maturity, tenor, and volatility as dimensions. When multiple volatility surfaces are considered at the same time for different strikes, the object is called a volatility cube. These implied volatilities provide the primary calibration targets for the interest rate models discussed later in Section 3.2.

2.2.5 The Volatility Surface and Smile

Market participants tend to quote options and swaptions in terms of implied volatilities rather than prices because they provide a convenient measure of option values across different expiries and tenors. Formally, for an observed market price V_{M} and a model price $V_{\text{model}}(\sigma)$, the implied volatility σ_{imp} is the unique solution to

$$V_{\text{model}}(\sigma_{\text{imp}}) = V_{\text{M}}.$$

Under the assumptions of constant volatility and log-normally distributed rates in the Black framework, the implied volatilities should, in principle, be independent of strike and maturity. If this were accurate, all swaptions with the same expiry would imply identical volatilities, and thus the corresponding implied volatility surface would be flat. However, empirical market data exhibits a systematic dependence on both strike and maturity, giving rise to the volatility smiles or skews (Gatheral, 2011; Brigo and Mercurio, 2001). These phenomena show that the implied volatility is higher for OTM or ITM swaptions than it is for ATM. This deviation from theory indicates that the lognormal assumption does not capture the true distribution of rate movements, which typically display heavier tails and asymmetry. As such, the market uses Black's formula as a standardized measure to quote prices as volatilities, rather than as a description of rate behavior.

Market quotes for implied volatilities are arranged over a grid of tenors and maturities, producing what is called the volatility surface. To include strike dependence, surfaces can be extended to what is known as the volatility cube - multiple surfaces together with different strikes in one object. Naturally, the volatility cube contains richer information, as it includes another dimension by including strikes. Figure 1 shows the implied volatility surface of the Euribor 6-month ATM volatilities on June 30th, 2025. Here, we can see that the underlying rate indeed has a volatility skew, since the volatilities at shorter tenors and maturities are considerably higher than those at longer maturities and tenors.

So far, we have only considered the log-normally distributed Black implied volatilities, as they are more present in the literature and historically dominant. However, after the financial crisis of 2008, negative rates are not considered an impossibility and thus worthy of consideration. For this, normal (or Bachelier)

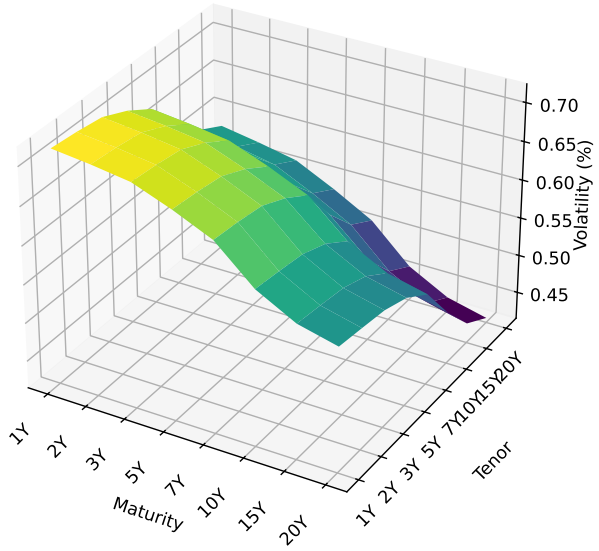


Figure 1: The surface of Euribor 6-Month ATM volatilities with tenors and maturities from 1Y to 20Y as of 30.6.2025.

volatilities are used, originating from Bachelier (1900). In the Bachelier model, the underlying is assumed to follow an arithmetic Brownian motion, allowing negative rate movements. That is, the Bachelier model assumes that the T-forward price of an asset S_t at time t follows an arithmetic Brownian motion with volatility σ_{Normal}

$$dS_t = \sigma_{\text{Normal}} dW_t.$$

The model is particularly useful when interest rates are low or negative as it provides stable price sensitivity when volatilities are near zero. As such, normal volatilities are often considered standard in some markets.

In practice, market data providers usually provide both measures and allow for direct conversion between them. Most theoretical derivations and interest rate models in the literature are expressed under the Black framework, and consequently the theoretical discussion in this thesis is done using the Black framework for consistency with the established literature. However, European swaptions are quoted in terms of normal volatilities and are often considered more realistic in EUR markets, and as such the numerical results in Chapter 5 are reported using normal volatilities. For more on this and the Bachelier model in general, see Choi et al. (2022).

The observed dependence of implied volatility on strike and maturity reflects the collective view of uncertainty by the market. The overall level of the volatility surface reflects the expected volatility, while its slope and curvature reveal how the market prices the likelihood of extreme rate scenarios. The implied volatility surface therefore

provides a concise, market-based representation of the distribution of future rates. The empirical volatility smile or skew motivates the later discussion in Section 3.1 on how volatility is modeled and represented in pricing frameworks.

2.2.6 Bermudan Swaptions

A Bermudan swaption is an option that grants its holder the right, but not the obligation, to enter into a fixed-floating swap on any of its predetermined set of call dates before final maturity. It can therefore be viewed as an intermediate case between a European swaption and an American swaption. Bermudan swaptions are commonly used as hedges for callable coupon bonds and for hedging interest rates in general, due to their flexible exercise schedules.

Formally, the holder of a Bermudan swaption can exercise the option on any of the call dates T_i that are typically a subset of the tenor structure of the swap $\{T_i\}_{i=0}^{n}$. Typically, the exercise structure is restricted to the subset $\{T_i\}_{i=s}^{n-1}$, where $s \geq 1$ and the time up to T_s is known as the no-call period where the swaption can not yet be exercised. For convenience, in this thesis we assume $s = 1$ and the set of exercise dates is therefore $\{T_i\}_{i=1}^{n-1}$ and with a no-call period from T_0 to T_1. When exercised at T_i, the holder enters into an interest rate swap to pay the fixed rate K and receive the floating rate from T_i to T_n. Once the option is exercised at some T_i no subsequent swaps can be entered, i.e. there are many exercise dates but only one option. If a Bermudan swaption only has one exercise date, it is clearly a European swaption. Similarly, each exercise opportunity of a Bermudan at times T_i can be considered a separate European swaption and thus the Bermudan can be considered to consist of European swaptions, from which we exercise the one that yields the highest payoff. However, the valuation process of Bermudan swaptions is not quite as simple as that of European swaptions, as we need to also account for the continuation value. The continuation value $H_i(T_i)$ is defined as the time-t value of the remaining option if it is not exercised, that is, the value of the Bermudan when only considering exercise dates after T_i. If the option has not been exercised by the last exercise date T_{n-1}, the option is worthless and the last continuation value $H_{n-1}(T_{n-1}) = 0$. As such, at the last exercise date T_{n-1} the value of a Bermudan swaption is

$$V_{\text{Bermudan}}(T_{n-1}) = \max(V_{\text{swap}}(T_{n-1}), 0),$$

whereas the value at some exercise date T_i is the maximum of the underlying swap value and the continuation value,

$$V_{\text{Bermudan}}(T_i) = \max(V_{\text{swap}}(T_i), H_i(T_i)).$$

Based on the above, one might think that Bermudan swaptions are simple generalizations of European swaptions and easy to price, yet they are considerably more complex and pricing them requires further knowledge.

3 Pricing of Interest Rate Derivatives

With Bermudan swaptions and the underlying framework introduced, we move on to discuss the methodology used to actually price them. We first pick up where Section 2.2.5 left off, and discuss volatility modeling (deterministic, local, and stochastic volatility models). After this, we move on to interest rate models, where we cover short-rate models as well as models in the Heath-Jarrow-Morton framework, providing the dynamics of each along with examples. The emphasis is placed on the Cheyette, Hull-White 1-Factor, and Hull-White 2-Factor models, as they are the models used in the numerical section to price Bermudan swaptions. We then discuss the calibration procedure for how interest rate models are calibrated, and conclude the section with a thorough overview of the pricing process of Bermudan swaptions.

3.1 Volatility Modeling

So far, volatility has been present only as a single constant parameter. When this is the case, the implied volatility surface is flat, which is not representative of real market data. This motivates a more accurate representation of volatility as a function that may depend on time, the underlying, and potentially additional sources of randomness. Through these more advanced volatility models, we may fit a larger set of prices with greater accuracy so that the instrument prices match market data.

A natural first evolution of the constant volatility in the Black framework is deterministic time-dependent volatility. Consider a simple Itô process

$$dX_t = \mu_t dt + \sigma(t) dW_t,$$

where $\sigma(t)$ is a deterministic function fixed at time zero. This differs from the constant-volatility case only by allowing time dependence, yet the implications are significant. The entire future level of instantaneous volatility is known at time zero, and all uncertainty in future prices then arises solely from the Brownian motion part. Deterministic volatility models are analytically convenient, often allowing closed-form or semi-closed-form pricing formulae, and can be calibrated to the term structure of option prices by choosing $\sigma(t)$ appropriately. However, because the volatility is not random or dependent on the level of the underlying, such models cannot replicate volatility smiles or skews. Any curvature present in the implied volatility surface is due to nonlinearities in pricing formulae rather than a feature of the model and is typically insufficient to reproduce the shapes observed in markets.

To capture the implied-volatility surface across both strikes and maturities better, one can let volatility depend deterministically on time and, additionally, the current level of the underlying. Introduced by Dupire (1994) and Derman and Kani (1994), these so-called local volatility models (LVM) have the form

$$dX_t = \mu_t dt + \sigma_{\text{loc}}(t, X_t) dW_t,$$

where the local volatility function $\sigma_{\text{loc}}(t, X_t)$ is chosen so that the model reproduces observed market prices exactly. When the volatility is formulated this way, all

randomness is embedded in the price path of the asset, and the instantaneous volatility of the underlying asset can therefore be deduced from the observed prices for all strikes and maturities. That is, local volatility models consider that volatility can change as the price of the underlying asset changes. As such, LVMs can capture smile and skew effects, providing improved flexibility compared to the constant volatility approach. Common examples of local volatility models include the Dupire model (Dupire, 1994), as well as the constant elasticity of variance (CEV) model (Cox, 1996).

While local volatility models are popular because they are relatively simple, easily implementable, and highly tractable, they do have their limitations. As the volatility is not random beyond its dependence on the current price of the underlying, future smile dynamics are determined by the current state, and therefore these models are unable to capture the inherent randomness of volatility itself. Local volatility models hence fail to exhibit realistic volatility smile dynamics. See, for example, Hagan et al. (2002) for further discourse on this and the shortcomings of LVMs in general.

This limitation introduces the approach of stochastic volatility models, which treat the volatility itself as a stochastic process and can consequently capture the evolution of the smile. The Heston model (Heston, 1993) is one of the most popular and earliest examples of a stochastic volatility model, in which the variance follows a mean-reverting square-root process. The SABR (Stochastic Alpha, Beta, Rho) model developed by (Hagan et al., 2002) is another popular stochastic volatility model, described by two equations for the forward rate F_t and its volatility α_{SABR}

$$\begin{aligned} dF_t &= \alpha_{\text{SABR}} F_t^{\beta} dW_t^{(1)}, \\ d\alpha_{\text{SABR}} &= \nu \alpha_{\text{SABR}} dW_t^{(2)}, \end{aligned}$$

where $dW_t^{(1)} dW_t^{(2)} = \rho dt$, $0 \leq \beta \leq 1$, and $\nu \geq 0$. Here, α_{SABR} is the initial volatility level, ν is the volatility of the volatility that indicates how pronounced a volatility smile or skew is, ρ represents the correlation between the forward rate and its volatility, and β is an elasticity parameter determining the stochastic model used to model the underlying process. When $\beta = 1$, the model resembles Black-Scholes and the forward rate follows a lognormal distribution. When $\beta = 0$, the model reduces to a stochastic normal model with normally distributed forward rates, and when $0 < \beta < 1$, the distribution of the forward rate interpolates between the two cases. SABR provides a way to capture smiles and skews through intuitive parameters that can be flexibly calibrated. An interesting extension of SABR is the free boundary specification developed by Antonov et al. (2015), which allows the SABR model to be generalized to negative rates. This extension models the forward rate as

$$dF_t = |F_t|^{\beta} \alpha_{\text{SABR}} dW_t^{(1)},$$

where $0 \leq \beta \leq \frac{1}{2}$, with a free boundary condition at $F = 0$ permitting negative rates.

In practice, the model is most commonly used through asymptotic closed-form approximations for implied volatilities, originally derived by Hagan et al. (2002). These analytical formulas allow the SABR parameters to be mapped directly to market volatility surfaces, making calibration extremely quick and convenient.

Approximations for both log-normal (Black) and normal (Bachelier) implied volatilities were formulated by Hagan et al. (2002). In interest rate markets, normal volatilities are more relevant due to the possibility of negative rates. For a European call option with a given forward rate F, strike K and time to maturity T, Hagan et al. (2002) derived that the SABR asymptotic normal implied volatilities $\sigma_N(K, F)$ can be computed as

$$\sigma_N(K,F) = \nu \frac{F-K}{x(\zeta)} \left\{ 1 + \left[\frac{\beta(\beta-2)\alpha_{\text{SABR}}^2}{24 F_{\text{mid}}^{2-2\beta}} + \frac{\rho\beta\nu\alpha_{\text{SABR}}}{4F_{\text{mid}}^{1-\beta}} + \frac{2-3\rho^2}{24}\nu^2 \right] T \right\},$$

where

$$F_{\text{mid}} = \frac{F+K}{2}, \quad \zeta = \frac{\nu}{\alpha_{\text{SABR}}} \left(\frac{F^{1-\beta} - K^{1-\beta}}{1-\beta} \right), \quad x(\zeta) = \ln\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2} + \zeta - \rho}{1-\rho} \right).$$

After the first asymptotic formulas developed by Hagan et al. (2002), many others have been proposed. Despite their great efficiency and accuracy near ATM, all of the asymptotic formulas have weaknesses. For example, in the original formula proposed by Hagan et al. (2002), the approximation accuracy deteriorates for long maturities and for strikes far from the forward level (i.e., in the wings of the volatility smile). For a more detailed treatment of these issues, see Section 5.8 in Crispoldi et al. (2016).

With the exception of the SABR model, used for volatility interpolation as described in Chapter 4, stochastic volatility models are left out of the numerical part of this thesis for two key reasons, even though theoretically they constitute an improvement in pricing and smile replication. First, they are often considerably more computationally demanding and difficult to calibrate compared to local and deterministic volatility models. Additionally, stochastic volatility alternatives for the interest rate models used were not available in the software used to obtain the results, nor were they easily accessible.

3.2 Interest Rate Modeling

Interest rate models are mathematical frameworks, typically SDEs, that determine how the stochastic evolution of interest rates is modeled. As the evolution of interest rates is complex, the models must be flexible and able to capture various statistical features of interest rates. The choice of interest rate model for a financial instrument depends on many criteria and is an actively debated topic in quantitative finance as, ideally, the chosen interest rate model should be able to represent the empirical volatilities of the term structure of rates, fit the market prices of vanilla derivatives such as swaptions, and be relatively easy to implement and apply. Interest rate models can broadly be divided into three classes: short-rate models, which describe the evolution of the instantaneous short rate; Heath-Jarrow-Morton (HJM) framework models, which model the entire forward rate curve; and market models, which directly model rates observable in the market. Our focus is on short-rate and HJM framework models, since market models are not considered in this thesis because they are considerably

more complex to calibrate and implement without likely providing a meaningful improvement for Bermudan swaptions. For a more thorough overview of interest rate models, see Brigo and Mercurio (2001) or Filipovic (2009).

3.2.1 Short-Rate Models

Short-rate models describe the evolution of the instantaneous short rate r_t, which is the fundamental driver of the term structure. As established in Section 2.2.1, under the risk-neutral measure $\mathbb{Q}$, the price of zero-coupon bonds can be expressed as

$$P(t,T) = \mathbb{E}^{Q}\left[\exp\left(-\int_t^T r_s ds\right)\middle|\mathcal{F}_t\right],$$

assuming the money-market account is used as numéraire. That is, the value of a zero-coupon bond maturing at time T can be determined by integrating the short rate from time t to time T. This relation highlights why modeling the short rate well is crucial; the entire term structure of bond prices, and therefore the prices of swaps, swaptions, and other more complex interest rate derivatives, can be derived once the short rate is modeled. The accuracy with which the short rate is modeled therefore also determines how accurately these interest rate derivatives are priced. The simplest short-rate models are single-factor models where the entire term structure is driven by a single state variable, implying perfectly correlated instantaneous movements across all maturities.

One of the first examples of short-rate models is the Vasicek model (Vasicek, 1977). In the Vasicek model, the instantaneous short rate is modeled by the following SDE:

$$dr_t = \alpha(b - r_t)dt + \sigma dW_t,$$

where $\alpha > 0$ is the mean reversion parameter representing the speed of mean reversion towards the long-term mean of the rate b, and σ is the instantaneous volatility. The idea of the model is that rates can fluctuate randomly but tend to return to a stable long-term average. The main advantage of the Vasicek model is that it allows for closed-form solutions for zero-coupon bond prices. When first introduced, its main disadvantage was that the state variable r_t is normally distributed, allowing the possibility of negative rates. This was an undesirable feature before the financial crisis, when rates went negative for the first time. With the possibility of negative rates nowadays, this is actually an advantage rather than a drawback, but the Vasicek model is not as common due to it being too simple for current rate environments.

The Cox-Ingersoll-Ross (CIR) model (Cox et al., 1985) was introduced as an extension of the Vasicek model to address the possibility of negative rates in the Vasicek model. Until negative rates were observed, this gave CIR a significant advantage over the Vasicek model, but this stance has been reconsidered in recent years as negative rates have become plausible. The CIR model is given by the following SDE:

$$dr_t = \alpha(b - r_t)dt + \sigma\sqrt{r_t}dW_t,$$

where the parameters are the same as in the Vasicek model. The CIR model differs from the Vasicek model by adding the square root term r_t, which ensures that the

rates given by CIR cannot be negative, as the distribution of rates is a non-central chi-squared distribution rather than a normal distribution.

The Hull-White 1-Factor (HW1F) model (Hull and White, 1990) is another extension of the Vasicek model that offers a key advantage over the CIR and Vasicek models: it is a no-arbitrage model, which means it can be calibrated exactly to the current term structure of interest rates, while the distribution of rates remains normal. The SDE of the model is

$$dr_t = (\theta(t) - \alpha r_t)dt + \sigma dW_t,$$

where $\theta(t)$ is a time-dependent drift term that is calibrated to ensure that the model matches the current term structure of rates, α is the speed of mean reversion commonly left as a user input, and σ is the volatility determined by calibration. The drift adjustment $\theta(t)$ is set by imposing the requirement that the model zero-coupon bond prices match the market curve for all maturities uniquely. In the model, it is given by

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + \alpha f(0,t) + \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}),$$

as shown in Hull and White (1990).

Under the rate dynamics of the Hull-White model, the price at time t of a zero-coupon bond with maturity T is given by

$$P(t,T) = A(t,T)e^{-r_t B(t,T)}, \quad 0 \le t \le T,$$

where

$$A(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left\{ - B(t,T)\frac{\partial \ln P(0,t)}{\partial t} - \frac{1}{4\alpha^3}\sigma^2(e^{-\alpha T} - e^{-\alpha t})^2(e^{2\alpha t} - 1)\right\},$$

and

$$B(t,T) = \frac{1}{\alpha}\left(1 - e^{-\alpha(T-t)}\right).$$

The proof and derivation can be found in Hull and White (1990) or Brigo and Mercurio (2001). The model supports the pricing of zero-coupon bonds and can thus be used to derive swap rates and other building blocks required to price interest rate derivatives.

European swaptions can also be priced analytically in the HW1F model using Jamshidian's decomposition (Jamshidian, 1989). Because a swap can be viewed as a coupon-bearing bond, the swaption payoff can be decomposed into options on zero-coupon bonds, which have closed-form prices in the Hull–White one-factor model as noted above. Consider a European swaption with strike K, maturity T_0, and nominal N with cash flows occurring at T_i for $i = 1, \ldots, n$. To apply Jamshidian's decomposition efficiently, it is convenient to express the underlying fixed leg of the swap, including the return of the notional principal, as a single portfolio of zero-coupon bonds. Assuming a unit notional ($N = 1$), the generalized fixed-leg cash flows c_i paid at each swap payment date T_i are defined as

$$c_i = \begin{cases} K\delta_i, & i = 1, \ldots, n-1, \\ 1 + K\delta_n, & i = n. \end{cases}$$

With this, the value of the fixed leg can now conveniently be written as

$$\sum_{i=1}^{n} c_i P(t, T_i).$$

Jamshidian's decomposition aims to find a critical short rate, r^*, at maturity T_0. This rate is the value of the spot rate at time T_0 for which the price of the coupon-bearing bond equals the strike, i.e., the underlying swap has a present value of zero at T_0. Therefore, r^* is the solution to

$$\sum_{i=1}^{n} c_i P(T_0, T_i; r^*) = 1.$$

In other words, the cash flows of the swap are discounted, added together, and set equal to the notional amount to find r^*, where the option will be exercised when $r_t < r^*$. Using the formula for zero-coupon bond prices admitted by the HW1F model gives

$$\sum_{i=1}^{n} c_i A(T_0, T_i) \exp(-B(T_0, T_i) r^*) = 1.$$

The value of r^* can then be found using a numerical root-finding algorithm, such as the Newton-Raphson algorithm. Let K_i be the time-T_0 value of a pure discount bond maturing at T_i when the spot rate is r^*. Once r^* is found, the value of K_i is given by

$$K_i = P(T_0, T_i; r^*) = A(T_0, T_i) \exp(-B(T_0, T_i) r^*).$$

By Jamshidian's decomposition, the swaption payoff can now be calculated as the sum of n options on discounted bonds with the exercise price of the i-th option equal to K_i. The payoff of the swaption at time $t < T_0$ is therefore

$$V_{\text{swaption}}(t) = \sum_{i=1}^{n} c_i \text{ZBP}(t, T_0, T_i, K_i),$$

where $\text{ZBP}(t, T_0, T_i, K_i)$ is the value of a European put option on a zero-coupon bond calculated as

$$\text{ZBP}(t, T_0, T_i, K_i) = K_i P(t, T_0) N(-d_2) - P(t, T_i) N(-d_1),$$

where $N(\cdot)$ is the cumulative standard normal distribution function and

$$d_1 = \frac{\ln\left(\frac{P(t,T_i)}{P(t,T_0)K_i}\right) + \frac{1}{2}\sigma_p^2}{\sigma_p}$$

$$d_2 = d_1 - \sigma_p, \tag{8}$$

where σ_p represents the volatility term specific to the HW1F model,

$$\sigma_p = \sigma \sqrt{\frac{1 - e^{-2\alpha(T_0 - t)}}{2\alpha}} B(T_0, T_i).$$

A more thorough derivation of the results, including an explanation of Jamshidian's decomposition, can be found in Brigo and Mercurio (2001) or Andersen and Piterbarg (2010a), as well as the original decomposition theorem by Jamshidian (1989).

The Hull-White 1-factor model combines many attractive features of short-term interest rate models and is therefore commonly seen as an appealing option. The inclusion of the $\theta(t)$-term allows the model to fit the current term structure exactly, and the inclusion of mean reversion ensures stability and prevents unrealistic behavior of rates. The rates given by the model are distributed normally, which is almost a necessity in current rate environments, and the model provides closed-form solutions to simplify and speed up pricing computations.

Numerous other one-factor short-rate models have also been proposed and played important roles in the development of interest rate modeling. Among these are the Black-Karasinski model (Black et al., 1990), the Dothan model (Dothan, 1978), the Black-Derman-Toy model (Black et al., 1990), the Ho-Lee model (Ho and Lee, 1986), and more. All of these models addressed specific modeling needs in their era and thus differ by properties such as distribution of rates, calibration flexibility, parameters, and so forth. Some of these models, such as the Ho-Lee or Black-Derman-Toy models, are still in use, albeit mostly for benchmarking purposes. In practice, models such as Hull-White are often preferred because of greater flexibility with regard to calibration or tractability and a better ability to capture various features of markets.

Although one-factor models greatly vary among themselves, they all share the limitation that the entire term structure is driven by a single state variable. Consequently, all yields are perfectly correlated which restricts the ability of the model to reproduce observed behavior in the market, particularly when it comes to the shape of swaption volatility surfaces. To capture the complex dynamics of the term structure more accurately, multi-factor short-rate models were introduced. These models, as the name suggests, use multiple stochastic factors to drive the evolution of the short rate, allowing them to model not just parallel shifts, but also changes in slope and curvature. Some prominent examples are, for example, the Longstaff-Schwartz model (Longstaff and Schwartz, 1992) and the Chen model (Chen, 1996). We focus on an extension of HW1F, the Hull-White two-factor model (HW2F) (Hull, 1994), which, compared to the one-factor version, contains an additional stochastic term whose mean reverts to zero. The model dynamics are described by

$$r_t = x_t + y_t + \theta(t), \quad r(0) = r_0,$$

where the two stochastic factors evolve according to Ornstein-Uhlenbeck processes,

$$\begin{aligned} dx_t &= -\alpha x_t dt + \sigma_1 dW_t^{(1)}, \\ dy_t &= -b y_t dt + \sigma_2 dW_t^{(2)}, \\ x_0 &= y_0 = 0 \end{aligned}$$

with instantaneous correlation ρ such that

$$dW_t^{(1)} dW_t^{(2)} = \rho dt.$$

Here, $\alpha, b > 0$ are mean-reversion parameters, σ_1, σ_2 volatilities of the respective factors, and $\theta(t)$ is a deterministic function, calibrated so that the model reproduces the initial discount curve exactly. Typically, the model is configured so that one factor models short-term changes while the other captures long-term movements of the yield curve. In order for the model to perfectly fit the market term structure of the discount factors, we must have (Brigo and Mercurio, 2001)

$$\theta(T) = f^M(0,T) + \frac{1}{2}\sum_{i,j=1}^{2} \frac{\sigma_i \sigma_j}{\alpha b}(1 - e^{-\alpha T})(1 - e^{-bT}),$$

where $f^M(0,T)$ is the market instantaneous forward rate at time 0 for a maturity T implied by the market term structure $T \mapsto P^M(0,T)$, i.e., $f^M(0,T) = -\frac{\partial \ln P^M(0,T)}{\partial T}$.

Similarly to the one-factor Hull-White model, the two-factor model admits a closed-form solution for zero-coupon bond prices. As shown and proven in detail in Brigo and Mercurio (2001), the zero-coupon bond price at time t for maturity T is given by

$$P(t,T) = A(t,T)\exp\big(-B_1(t,T)x(t) - B_2(t,T)y(t)\big).$$

where

$$A(t,T) = \frac{P^M(0,T)}{P^M(0,t)}\exp\Big(\tfrac{1}{2}\big(V(t,T) - V(0,T) + V(0,t)\big)\Big),$$

$$B_1(t,T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha}, \qquad B_2(t,T) = \frac{1 - e^{-b(T-t)}}{b},$$

$$\begin{aligned} V(t,T) = {} & \frac{\sigma_1^2}{\alpha^2}\left[T - t - \frac{2}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) + \frac{1}{2\alpha}\left(1 - e^{-2\alpha(T-t)}\right)\right] \\ & + \frac{\sigma_2^2}{b^2}\left[T - t - \frac{2}{b}\left(1 - e^{-b(T-t)}\right) + \frac{1}{2b}\left(1 - e^{-2b(T-t)}\right)\right] \\ & + \frac{2\rho\sigma_1\sigma_2}{\alpha b}\left[T - t - \frac{1 - e^{-\alpha(T-t)}}{\alpha} - \frac{1 - e^{-b(T-t)}}{b} + \frac{1 - e^{-(\alpha+b)(T-t)}}{\alpha + b}\right]. \end{aligned}$$

By substituting these back into the affine-form solution, the complete closed-form formula for the zero-coupon bond price at time t is given by

$$P(t,T) = \frac{P^M(0,T)}{P^M(0,t)}\exp\Big(-B_1(t,T)x(t) - B_2(t,T)y(t) + \tfrac{1}{2}\big(V(t,T) - V(0,T) + V(0,t)\big)\Big).$$

For the complete derivation and further details, see Section 4.2.2 in Brigo and Mercurio (2001).

This formulation of the Hull-White 2-factor model, also known as the G2++ model, shares the normally distributed rates of HW1F which makes it attractive in modern times due to the possibility of negative rates. Recall that, in the one-factor model, every zero-coupon bond price at a fixed time t can be written as a function of

a single random variable, namely the short rate r_t. This allows the swaption payoff to be reduced to a portfolio of options via Jamshidian's decomposition, leading to the closed-form pricing formula. As the two-factor model depends on two stochastic factors x_t and y_t rather than one, each bond price is a function of (x_t, y_t). Consequently, the swaption payoff depends on two random variables, as the swap value is a linear combination of bond prices. Because the zero-coupon bond prices are now driven by two distinct stochastic processes, their joint distribution cannot be collapsed into a single univariate Gaussian variable without violating the correlation structure across different cash flow maturities, and there is no longer a critical state similar to r^* where the underlying swaption is valued at par. Therefore, Jamshidian's decomposition cannot be applied, and the HW2F model does not have an analytical formula for pricing European swaptions. However, it still yields a semi-analytical formula that one can use rather than having to exclusively resort to numerical pricing methods.

Because the option cannot be decomposed into individual bond options like in HW1F, its expected payoff must be evaluated using a double Gaussian integral over the state variables x_t and y_t. To avoid the computational burden of two-dimensional numeric integration, the problem can be reduced to a one-dimensional integral by freezing the variable x_t in the integrand. The complete derivation is rather complex and extensive, and is therefore omitted from this thesis. A detailed treatment can be found in Appendix D of Section 4.2 in Brigo and Mercurio (2001). Following this derivation, the arbitrage-free price of a European swaption at time $t = 0$ is given by the following one-dimensional integral

$$V_{\text{swaption}}(0) =$$

$$N\omega P(0,T_0)\int_{-\infty}^{+\infty}\frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x\sqrt{2\pi}}\left[N(-\omega h_1(x)) - \sum_{i=1}^{n}\lambda_i(x)e^{\kappa_i(x)}N(-\omega h_2(x))\right]dx,$$

where $\omega = 1$ for a payer swaption and $\omega = -1$ for a receiver swaption, $N(\cdot)$ is the cumulative standard normal distribution function,

$$h_1(x) = \frac{\bar{y}(x)-\mu_y}{\sigma_y\sqrt{1-\rho_{xy}^2}} - \frac{\rho_{xy}(x-\mu_x)}{\sigma_x\sqrt{1-\rho_{xy}^2}}$$

$$h_2(x) = h_1(x) + B_2(T_0,t_i)\sigma_y\sqrt{1-\rho_{xy}^2},$$

$$\lambda_i(x) = c_i A(T_0,t_i)e^{-B_1(T_0,t_i)x}$$

$$\kappa_i(x) = -B_2(T_0,t_i)\left[\mu_y - \frac{1}{2}(1-\rho_{xy}^2)\sigma_y^2 B_2(T_0,t_i) + \rho_{xy}\sigma_y\frac{x-\mu_x}{\sigma_x}\right],$$

and

$\bar{y} = \bar{y}(x)$ is the unique solution of the following equation

$$\sum_{i=1}^{n} c_i A(T,t_i)\, e^{-B_1(T,t_i)x - B_2(T,t_i)\bar{y}} = 1.$$

The distributional parameters $\mu_x, \mu_y, \sigma_x, \sigma_y$, and the correlation ρ_{xy} are the means, standard deviations, and correlation of the two factors, evaluated under the T_0-forward measure. In practice, this one-dimensional integral is highly tractable and is typically solved efficiently using Gauss-Hermite quadrature (Ye, 2023).

3.2.2 Heath-Jarrow-Morton Framework Models

The Heath-Jarrow-Morton (HJM) framework (Heath et al., 1992) is a framework to model the evolution of interest rates by modeling the instantaneous forward rate curves directly. While short-rate models only capture the dynamics of a point in the curve, i.e. the short rate, HJM models are capable of capturing the full dynamics of the entire forward rate curve. In the HJM framework, under the risk-neutral measure $\mathbb{Q}$, the dynamics of the instantaneous forward rate $f(t,T)$ are given by

$$\begin{aligned} df(t,T) &= \alpha(t,T)dt + \sigma(t,T)dW_t, \quad 0 \le t \le T, \\ f(0,T) &= f^M(0,T), \end{aligned}$$

where $T \mapsto f^M(0,T)$ is the market instantaneous forward rate curve at time $t = 0$, $W = (W_1, \ldots, W_n)$ is an n-dimensional Brownian motion, $\sigma(t,T) = (\sigma_1(t,T), \ldots, \sigma_n(t,T))$ is a vector of adapted processes and $\alpha(t,T)$ is itself an adapted process. A model can be classified as an HJM model if the model follows the stochastic differential equation above.

The key behind the HJM framework is the HJM drift condition. Heath et al. (1992) showed that in order for a unique risk-neutral measure to exist, the function α cannot be chosen arbitrarily, and must equal a quantity depending on σ and on the drift rates in the dynamics of N selected zero-coupon bond prices. In other words, the absence of arbitrage uniquely determines the drift in terms of the volatility, and no drift estimation is needed. In particular, the drift condition states we must have

$$\alpha(t,T) = \sigma(t,T) \int_t^T \sigma(t,s)ds. \tag{9}$$

Consequently, the dynamics of $f(t,T)$ and therefore the whole model are fully determined once the volatility structure $\sigma(t,T)$ is specified. As such, many short-rate models can also be shown to be a part of the HJM framework if the volatility structure is of a very specific form. For example, Brigo and Mercurio (2001) show that Vasicek and HW1F are HJM models if the volatility is of form $\sigma(t,T) = \sigma e^{-a(T-t)}$, and Filipovic (2009) shows that the Ho-Lee model is HJM with constant volatility $\sigma(t,T) = \sigma > 0$. In contrast to short-rate models, the HJM framework ensures consistency of interest-rate movements across all maturities rather than the yield curve being implied by only the short-rate dynamics.

However, since HJM models indeed capture the full dynamics of the entire forward rate curve $f(t,T)$, they are generally non-Markovian and infinite-dimensional. Unlike short-rate models such as HW1F, where the entire yield curve is determined at each point in time by a single state variable, HJM models evolve the forward curve across

all maturities simultaneously. Specifying the curve exactly therefore requires knowing its value at every maturity T, making the state space infinite-dimensional in general.

The non-Markovian nature of HJM models is most clearly seen from the drift condition (9), where t appears as both a limit of integration and inside the integrand. The drift of $f(t,T)$ at any point T thus depends on an integral of the volatilities across all maturities between t and T, meaning the evolution of the curve at time t depends not just on its current value but on the entire history of the forward curve. This path-dependence is what makes the general HJM framework non-Markovian, and is what poses a computational challenge, since standard numerical pricing methods are either inapplicable or extremely demanding.

For a practical implementation, a finite-dimensional Markovian representation is then required. This can be done by finding a suitable specification of σ so that the HJM model becomes Markovian in a finite-dimensional state space, which means that there exists a process $X_t \in \mathbb{R}^n$ such that all bond prices $P(t,T)$ can be written as functions of X_t rather than the entire forward curve. The one-factor Cheyette local volatility model (Cheyette, 2001) is an especially interesting example of such a model, intended to overcome these very limitations of the HJM framework. The dynamics of the one-factor Cheyette local volatility model are

$$dx_t = [y_t - \alpha x_t]dt + \sigma(t, x_t, y_t)dW_t, \tag{10}$$

$$dy_t = [\sigma^2(t, x_t, y_t) - 2\alpha y_t]dt, \tag{11}$$

$$r_t = \theta(t) + x_t, \tag{12}$$

$$x_0 = y_0 = 0, \tag{13}$$

where r_t is the short rate, x_t is the main driver state that governs the level of interest rates, y_t is the convexity state that controls the local variance of rate changes, α is the mean reversion speed, $\sigma(t,x,y)$ is the local volatility function, and $\theta(t)$ is the deterministic shift to fit the current term structure. In practice, it is usually assumed that σ does not depend on y. Because this thesis focuses exclusively on the one-factor local volatility formulation of the Cheyette framework, all subsequent references to the Cheyette model in this thesis explicitly denote this specification.

As Andersen and Piterbarg (2010a) describe, the assumption that r_t is Markovian requires imposing a separability condition on the deterministic volatility structure of instantaneous forward rates, i.e., $\sigma(t,T) = g(t)h(T)$, where g and h are deterministic functions. In addition to making the process Markovian, the separability condition forces the short-rate model to be driven by the two factors x_t and y_t as in 10–13. A one-factor Cheyette local volatility model is then obtained by requiring g to be a deterministic time-dependent function of the state variables, $g(t) = g(t, x_t, y_t)$. Many different local volatility structures exist for the Cheyette model, but we choose one where a smart initial guess (SIG) of a Hull-White 1F model is used that is calibrated to the data and used as SIG for Cheyette, as this is the version we use later on in the numerical section. This is explained in further detail in Section 4.3, after the calibration methodology and Bermudan pricing have been discussed.

Following Andersen and Piterbarg (2010a), we define the auxiliary variable $B(t,T) = \frac{1-e^{-\alpha(T-t)}}{\alpha}$, and note that zero-coupon bond prices can be expressed as

deterministic functions of x_t and y_t,

$$P(t,T) = P(t,T,x_t,y_t) = \frac{P(0,T)}{P(0,t)} \exp\left(- B(t,T)x_t - \frac{1}{2}B(t,T)^2 y_t \right).$$

The Cheyette model does not admit a closed-form solution for European swaption prices, as the payoff depends jointly on both state variables through bond prices and the exercise boundary in the (x_t, y_t) plane is non-linear. Pricing is therefore performed numerically via numerical pricing methods, such as the finite-difference method or Monte Carlo method discussed in Section 3.4. In the context of the one-factor local volatility Cheyette model specifically, a detailed treatment of the finite-difference method used for pricing derivatives can be found in Andersen and Piterbarg (2010a). It is also worth noting that approximate semi-analytic formulas for pricing European swaptions exist for the one-factor local volatility Cheyette model (Andersen and Piterbarg, 2010a). However, they are rarely used in practice due to potential pricing inaccuracies and because the numerical methods are efficient enough to be considered practical.

The one-factor local volatility Cheyette model offers an attractive combination of several properties. Similarly to the HW models, it fits the initial yield curve exactly and maintains enough flexibility to capture complex interest rate structures while remaining Markovian and finite-dimensional. This allows Bermudan swaptions and other derivatives to be priced efficiently, while the local volatility specification enables the model to reproduce volatility smiles and skews more accurately than standard short-rate models. Additionally, it is simple enough for efficient numerical pricing methods such as finite difference methods or Monte Carlo, and can fit a large number of option quotes if needed.

Despite these advantages, the model still has some drawbacks. Although relatively simplified compared to the general HJM framework, it is still more complex than the Vasicek, CIR, and HW models. These classical models benefit from a longer history of industry use, wider implementation for easier access, and deeper familiarity in general. Furthermore, compared to short-rate models, the Cheyette parameters can depend on both time and maturity, which requires more extensive data and calibration effort, making it more computationally demanding. As a result, practitioners often prefer short-rate models due to their intuitiveness, speed, and historical precedence. Nevertheless, the Cheyette model is widely regarded as a powerful and flexible alternative, as its pricing capabilities heavily outweigh these computational drawbacks.

3.3 Calibration Methodology

In order for interest rate models to produce accurate prices for derivatives, their parameters must first be determined from observed market data. This process is referred to as calibration, where the objective is to find a set of model parameters so that the model reproduces market prices as closely as possible. Interest rate models typically contain parameters governing both the level dynamics of interest rates and the volatility structure, typically through the mean-reversion rate and the volatility.

Although the mean-reversion rate can in principle also be calibrated, it is a common practice to leave it as a user input estimated from the current market conditions or matched to counterparty quotes. This is because the mean reversion calibration procedure can often cause instability in the pricing process (Andersen and Piterbarg, 2010a). As such, in this thesis we treat the mean-reversion parameters as user inputs, and the calibration primarily focuses on determining the volatility parameters of the models. The volatility is calibrated product-specific to match relevant vanilla options, namely European swaptions for Bermudan swaptions.

The calibration problem can be formulated as a non-linear least-squares optimization problem, where the objective is to find the vector of volatility step parameters $\Sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ that minimizes the discrepancies between model-implied and market-observed swaption volatility quotes. In this thesis, the discrepancy is measured using relative normal volatility errors, i.e. the relative difference between the model-implied volatilities and the target normal volatilities observed in the market. The choice of normal volatility is motivated by the presence of negative interest rates in modern market environments, while the use of relative errors prevents instruments with large absolute volatilities from dominating the objective function.

Let $S(\Sigma)$ denote the sum of the squared relative deviations. The calibration objective function is formulated

$$S(\Sigma) = \sum_{i=1}^{N} \left(\frac{\sigma_{\text{model}}(i, \Sigma) - \sigma_{\text{market}}(i)}{\sigma_{\text{market}}(i)} \right)^2,$$

where $\sigma_{\text{model}}(i, \Sigma)$ denotes the model-implied volatility for instrument i under the parameter set Σ, and $\sigma_{\text{market}}(i)$ denote the corresponding market volatility. The goal of the calibration procedure itself is to find the parameters $\hat{\Sigma}$ that minimize this sum so that

$$\hat{\Sigma} \in \arg\min_{\Sigma} S(\Sigma),$$

which is assumed to be a non-empty set.

Although the objective function is expressed in terms of implied normal volatility, the calibration is performed indirectly via prices. For a given set of parameters Σ, we first compute the model prices of the calibration instruments, which we then convert to implied normal volatilities. The relative volatility residuals are then evaluated from these implied volatilities. This approach ensures internal consistency between the pricing framework and volatility representation across all models.

The optimization problem is solved using Jacobian-based non-linear least-squares methods. In this thesis, the calibration is performed using a Levenberg-Marquardt algorithm (Marquardt, 1963), which is widely used to solve non-linear least-squares problems. The algorithm initializes by taking a starting guess x and systematically making steps of Δx along all parameter axes in sequence. This allows the solver to obtain an initial finite-difference approximation to the Jacobian matrix, defined as

$$J(\Sigma) = \frac{\partial V}{\partial \Sigma},$$

where V denotes the vector of model outputs, in this case swaption volatilities. The Jacobian is evaluated at each iteration, and the solver therefore iteratively updates both the parameter solution and the Jacobian matrix. For further technical details regarding the algorithm, see Marquardt (1963), Press (2007), or Nocedal (2006).

In the documentation of the software used in the numerical parts, the specific algorithm used is stated to be a non-diagonal extension of the Levenberg-Marquardt method where the Jacobian matrix is obtained from finite differences and updated by Broyden corrections. The algorithm converges rapidly, but can also often exit too early on highly nonlinear equations. Nevertheless, the software documentation states it to often be the best method for calibration over other common numerical methods, such as quasi-Newton class algorithms, Gauss Newton Subspace Solver, and Simplex.

3.4 Bermudan Swaption Pricing

Having specified the modeling framework and calibration methodology, we build on Section 2.2.6 and describe how a Bermudan swaption is priced. Recall that the key feature of Bermudan swaptions is multiple exercise dates $\{T_i\}_{i=1}^{n-1}$ on which the holder has the right to enter an underlying swap at any of these dates. If exercised at T_i, the holder receives the value of the swap $V_{\text{swap}}(T_i)$ by entering it, and the option then ceases to exist. If not exercised on some T_i, the exercise decision is postponed to the next exercise date until the option expires on T_{n-1}. Therefore, when pricing Bermudan swaptions, we need to account for the continuation value $H_i(t)$ which is the time-t value of the remaining option if the option is not exercised. The option becomes worthless if it is not exercised by the last exercise date T_{n-1}, and consequently $H_{n-1}(T_{n-1}) = 0$. Therefore,

$$V_{\text{Bermudan}}(T_{n-1}) = \max(V_{\text{swap}}(T_{n-1}), 0).$$

Generally, the value of a Bermudan swaption is then determined by evaluating at each exercise date whether to exercise it or whether to postpone the decision to a later exercise date. This can be formulated as

$$V_{\text{Bermudan}}(T_i) = \max(V_{\text{swap}}(T_i), H(T_i)).$$

Here, $V_{\text{Bermudan}}(T_i)$ represents the Bermudan option value at exercise date T_i. As such, we must also have for the continuation value that

$$H_{i-1}(T_i) = V_{\text{Bermudan}}(T_i),$$

i.e., the continuation value at time T_{i-1} is equal to the value of the Bermudan swaption at the next exercise date T_i. Using the general derivative pricing formula under the risk-neutral measure (2), we get

$$\begin{aligned} H_{i-1}(T_{i-1}) &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_{T_{i-1}}^{T_i} r_u du \right) V_{\text{Bermudan}}(T_i) \middle| \mathcal{F}_{T_{i-1}} \right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_{T_{i-1}}^{T_i} r_u du \right) \max(V_{\text{swap}}(T_i), H_i(T_i)) \middle| \mathcal{F}_{T_{i-1}} \right]. \end{aligned}$$

Alternatively, we can use the bank account numéraire $B(t) = \exp\left(\int_0^t r_u du\right)$ to represent the discount factor, and the continuation value can then be written as

$$\begin{aligned} H_{i-1}(T_{i-1}) &= B(T_{i-1})\mathbb{E}^{\mathbb{Q}}\left[\frac{H_{i-1}(T_i)}{B(T_i)}\right] \\ &= B(T_{i-1})\mathbb{E}^{\mathbb{Q}}\left[\frac{\max(V_{\text{swap}}(T_i), H_i(T_i))}{B(T_i)}\middle|\mathcal{F}_{T_{i-1}}\right]. \end{aligned}$$

As can be seen, the continuation values can be derived recursively using backward induction. This is because the optimal stopping decision at any given exercise date inherently depends on the unknown expected value of maintaining the option into the future, i.e., the continuation value, and therefore the pricing process cannot be simulated forward. The valuation is therefore performed via backward induction by anchoring to the final exercise date, where the continuation value is zero and the payoff is known, and recursively going backward in time through the exercise schedule to compute the continuation value and determine the optimal exercise decision at each previous exercise date. It is also worth noting here that the valuation of a Bermudan swaption is sometimes described in a more intuitive way such that

$$V_{\text{Bermudan}}(t) = V_{\text{MaxEuropean}}(t) + V_{\text{SwitchOption}}(t),$$

where MaxEuropean is the European swaption with the highest value out of the set of exercise dates, i.e.

$$V_{\text{MaxEuropean}}(t) = \max_k \left\{ V^{(i)}_{\text{swaption}}(T_i) \middle| i = 1, \ldots, n-1 \right\},$$

and SwitchOption is a theoretical construct that measures the value of the flexibility to postpone the exercise decision. That is, it can be interpreted as the value of the additional optionality in a Bermudan swaption relative to a European swaption.

Formally, the recursion algorithm for pricing Bermudan swaptions is derived via optimal stopping theorem and Hamilton-Jacobi-Bellman equation, see Section 18.2.2 in Andersen and Piterbarg (2010b) for further details. Although the recursion algorithm is relatively straightforward and holds for any payoff $V_{\text{swap}}(T_i)$, the conditional expectations can prove a computational challenge. This is because calculating the conditional expectation depends on the entire distribution of future paths of rates which must be integrated over, and the swaption payoffs $V_{\text{swaption}}(T_i)$ are non-linear due to the presence of the maximum operator (and the ability to choose whether to exercise or not). Consequently, these continuation values do not admit closed-form solutions even for more tractable models such as HW1F, and we must therefore resort to backward numerical pricing methods to compute the continuation values and the prices of the Bermudan swaptions. These numerical pricing methods are discussed more in Section 3.4.1.

The valuation process of a Bermudan swaption is not a simple isolated calculation, but rather the culmination of the entire framework discussed so far in this thesis. We

now describe the full pricing process from start to finish, assuming that there is an existing Bermudan swaption contract that has been specified already with a set of exercise dates, maturity of the underlying swap, fixed rate, notional, floating index, and market conventions.

1. Yield curve construction: Discounting and projection curves are bootstrapped from market instruments to obtain the discount factors $P(0,T)$ and forward rates used to accurately price the cash flows of the underlying swap.

2. Volatility cube construction via SABR interpolation: Because market swaption quotes are not available for all maturities and strikes required for calibration, the volatility surface is interpolated using a SABR model (as detailed in Section 3.1) for the strike dimension, while bilinear interpolation is used across the expiry and tenor dimensions, to generate a fully realized, consistent volatility cube. This is described in further detail in Section 4.1.

3. Choice of interest rate model and specification: We choose and specify an interest rate model to use for the pricing, and adjust it so that the model reproduces the initial term structure. The choice of interest rate model defines the evolution of the yield curve, the dynamics of zero-coupon bond prices, and the distribution of future rates.

4. Model calibration: The model parameters are calibrated to European swaption prices or volatilities derived from the volatility cube as per the calibration strategy employed. The calibration methodology in this thesis is the one described in Section 3.3, although there are many other methods.

5. Bermudan pricing: Using the calibrated model, the continuation values are calculated, and the recursive valuation

$$V_{\text{Bermudan}}(T_i) = \max(V_{\text{swap}}(T_i), H_i(T_i))$$

   is applied backward through the exercise schedule to compute the value. Here, a numerical pricing method, such as the finite-difference PDE method, is applied to compute the continuation values (see Section 3.4.1).

3.4.1 Numerical Pricing Methods

A commonly used framework for Bermudan pricing is the backward finite-difference method. In this approach, the finite-difference solution of the backward pricing partial differential equation associated with the interest rate model is solved to obtain the value of the Bermudan swaption. The Bermudan value is represented as a function of time and the model state variable(s), and the partial differential equation is solved numerically on a discretized grid backward through time. The main steps of this method are as follows.

1. The state space, such as the short-rate in a one-factor model, is discretized on a grid $[x_0, \ldots, x_n]$ with upper and lower bounds, using time step size Δt and $\theta \in [0, 1]$ which is a weighting parameter that defines the finite-difference scheme used. For one-factor models such as HW1F and Cheyette this is a one-dimensional grid, while two-factor models like HW2F require a two-dimensional grid which increases computational complexity.

2. The terminal condition for the current valuation step is specified. At the final exercise date T_{n-1}, the option value equals the exercise payoff

$$V_{\text{Bermudan}}(T_{n-1}) = \max(V_{\text{swap}}(T_{n-1}), 0).$$

3. The continuous pricing PDE associated with the interest rate model is discretized into a system of linear equations to step backward from a future time $t + \Delta t$ to the current time t. This results in a linear system that relates the unknown continuation values at the current time step to the known option values from the next time step in the backward sweep, effectively providing a numerical approximation to the conditional expectations in the recursion algorithm.

4. Appropriate product-specific boundary conditions are imposed at the edges of the state grid to ensure numerical stability and to prevent the computations from exploding to infinity at extreme interest rate levels.

5. The PDE is solved backward in time using a finite-difference time-stepping scheme, such as the Crank-Nicolson method (Crank and Nicolson, 1947). The solution gives the precise continuation value at each point on the grid just before each exercise date.

6. At each exercise date T_i, the Bermudan early-exercise condition is enforced and the intrinsic value of the swap is computed for each node on the grid. The option value is then updated according to

$$V_{\text{Bermudan}}(T_i) = \max(V_{\text{swap}}(T_i), H_i(T_i)).$$

   If the current time step does not correspond to an exercise date, the option value is simply the continuation value.

7. The backward stepping, i.e. steps 3 through 6, are repeated until the valuation reaches the initial time (the valuation date) $t = 0$, yielding the Bermudan price.

A detailed treatment of finite-difference methods for derivative pricing can be found in Duffy (2006) and Andersen and Piterbarg (2010a).

While the backward finite-difference method is the primary engine employed for the pricing in this thesis, a common alternative approach to computing continuation values is Monte Carlo simulation. Here, the interest rate model is simulated forward in time along N independent paths, and the continuation values are estimated from the simulated data rather than solved directly via a PDE. Regular Monte Carlo methods

cannot directly handle the inherent optimal stopping problem in Bermudan swaptions, as they simulate forward and the valuation instead requires backward induction. Therefore least-squares Monte Carlo (LSMC) is used instead Longstaff and Schwartz (2001).

The key idea of LSMC is to approximate the continuation value at each exercise date T_i by regressing the discounted future payoffs against a set of basis functions of the state variables. Formally, at each exercise date T_i, the continuation value is approximated as

$$\hat{H}_i(T_i) \approx \sum_{k=1}^{K} \beta_k \psi_k(X_{T_i}),$$

where ψ_k are basis functions, X_{T_i} denotes the model state variables at time T_i, and β_k are coefficients estimated by ordinary least squares from the simulated paths. Common choices for basis functions are Laguerre polynomials as in Longstaff and Schwartz (2001), as well as Hermite or Chebyshev polynomials. To ensure numerical stability and avoid multicollinearity in the regression, the basis functions are chosen to be orthogonal polynomials, typically of low degree like two or three (Glasserman, 2004). The regression coefficients can be efficiently estimated via standard matrix computations. Following Glasserman (2004), we define a $K \times K$ matrix B_ψ representing the sample cross-products of the basis functions, and a $K \times 1$ vector $B_{\psi V}$ representing the sample cross-covariances between the basis functions and the discounted future payoffs. We then obtain the ordinary least squares solution as

$$\hat{\beta} = B_\psi^{-1} B_{\psi V},$$

which only involves matrix multiplication and inversion, remaining highly efficient even with a large number of paths.

An important consideration is the bias introduced by using the same paths for both estimating the regression coefficients and computing the continuation values. If the same paths are reused, the estimated continuation values are positively biased, as the regression model is evaluated on the same data it was fitted to, leading to an overestimate of the option value. To mitigate this, we first use one set of paths to estimate the regression coefficients $\hat{\beta}$ via backward induction through the exercise schedule. Once the coefficients have been estimated, a fresh set of paths is simulated independently, and the already estimated regression model is applied to these new paths to compute the continuation values and determine the optimal exercise decision. The Bermudan price is then obtained by averaging the discounted payoff across the simulated paths

$$V_{\text{Bermudan}}(0) \approx \frac{1}{N} \sum_{n=1}^{N} \exp\left(- \int_0^{\tau^{(n)}} r_s ds \right) V_{\text{swap}}(T_{\tau^{(n)}}),$$

where $\tau^{(n)}$ denotes the optimal exercise time on path n.

4 Experimental Setup

This chapter presents the setup for the numerical analysis conducted in this thesis. Having established the theoretical framework in the previous sections, we now transition to the practical implementation and calibration of the chosen interest rate models: the Hull-White One-Factor, Hull-White Two-Factor, and the Cheyette models. The primary objective is to evaluate how these models can be calibrated to market data and how well they replicate the market volatility smile, as well as how well they perform in pricing complex derivatives. The interest rate models are calibrated to market implied volatilities of European swaptions, and the calibrated models are used to price a Bermudan swaption, whose contract characteristics and structure are summarized in Table 1.

Contract Parameter	**Value**
Position	Short
Notional	€5,000,000
Swap Start Date	27.10.2022
Swap End Date	27.10.2042
Strike Rate	2.00%
Floating Leg	Euribor 6-month
Fixed Leg Day Count Convention	30/360
Fixed Leg Payment Frequency	Annual
Exercise Frequency	Annually on October 6th (2025–2037)
Payment Date	27 October (year of exercise)

Table 1: Specification of the Bermudan swaption used in the numerical experiments.

The remainder of this chapter details the practical framework required to price this instrument. First, we outline the market data utilized, including the yield curves, and how the SABR model is used to create a fuller volatility cube that provides refined market data for the calibration and pricing process. Next, we define the calibration instruments and their conventions, and discuss the two calibration strategies we use and how they differ, especially concerning Bermudan swaptions. Finally, we consider the specific implementations and specifications of the interest rate models employed.

4.1 Market Data

The valuation date for all market data utilized in this thesis is 30 June 2025. Due to licensing agreements and contractual obligations with the market data provider, the raw market quotes used in this thesis cannot be explicitly disclosed. Instead, we describe the structure and characteristics of the data and describe the requirements to replicate the numerical results, given the necessary market data is available.

The available dataset consists of European swaption normal volatility surfaces. The market quotes span a comprehensive grid of option expiries ranging from one week to 50 years, and swap tenors ranging from one year to 50 years. Together, the

volatility surfaces form a swaption volatility cube, defined across swaption expiries, underlying swap tenors, and strike levels. For each available expiry-tenor pair, normal volatilities are quoted across a relative strike grid ranging from ATM − 400bps to ATM + 400bps in increments of 50bps. Standard market conventions for European swaptions apply throughout the dataset.

In addition to volatility quotes, the pricing framework requires yield curves for discounting and forward rate computations. A dual-curve framework is employed, where discounting is performed using an OIS discount curve and projected forward rates for the floating leg are generated from an Euribor 6-month projection curve as per the floating leg of the underlying swap in the Bermudan. The specifications required to reproduce the yield curves used in this thesis are provided in Appendix A, which details the relevant day-count conventions, interpolation methods, and discount factors. Given these specifications and access to comparable market instruments, the curves can be reconstructed by applying standard bootstrapping methods in conjunction with the specified interpolation and day-count conventions.

To obtain a complete volatility cube across the strike dimension, the available market quotes are interpolated using a SABR model to obtain more extensive market data. This is essential because the calibration procedure considers strikes and expiry-tenor combinations based on the structure of the Bermudan swaption, which may not coincide with market quotes due to the market data not being liquid enough to have available quotes for the strikes and dates we need to calibrate the models to. Without the interpolation, we would have to resort to the closest available quotes in terms of strike or expiry/tenor, which would result in inaccurate pricing of the Bermudan swaption. By calibrating the parameters of the SABR model to the available market quotes, the model can generate SABR volatilities for any moneyness, strike, and tenor, eventually producing a smooth and arbitrage-consistent representation of volatility surfaces. Repeating this process for every surface available gives the full volatility cube that we can use to price European swaptions outside of the market quotes. For further discussion on the SABR model and its calibration, see Hagan et al. (2002), Hagan et al. (2015), and Antonov et al. (2013).

Appendix B details the SABR model specifications used in this thesis, while Appendix C provides the calibrated parameters α, ν, and ρ, with β fixed at 0.5. An illustrative example of the resulting volatility surface interpolation is in Figure 2, which displays the SABR-implied volatility surface for the 10-year expiry. Based on the interpolated volatility cube, a subset of European swaptions mirroring the exercise schedule of the Bermudan swaption is selected as calibration instruments.

4.2 Calibration Instruments and Strategies

To reliably price the Bermudan swaption specified in Table 1, the interest rate models must be calibrated to a set of plain-vanilla market instruments that represent the characteristics of the target derivative. For Bermudan swaptions, the standard calibration instruments are European swaptions on the same underlying index. In this thesis, Euribor 6-month European swaptions are used, reflecting the fact that the floating leg of the underlying Bermudan swaption is indexed to the Euribor 6-month

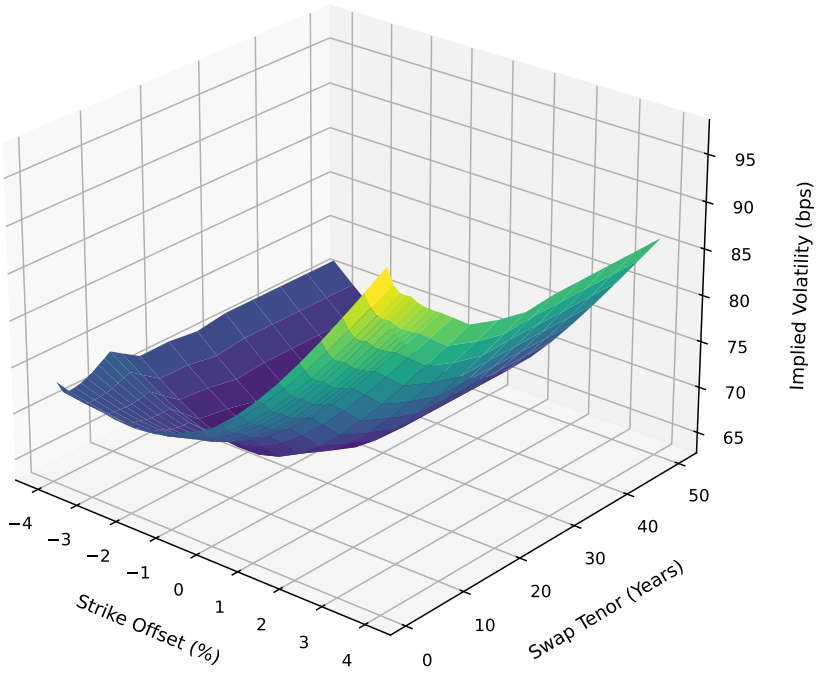


Figure 2: Volatility surface interpolated with SABR, 10-year expiry

reference rate. The swaption volatility quotes are obtained from the interpolated SABR volatility cube described in Section 4.1. European swaptions have high liquidity and are directly connected to the underlying swap in a Bermudan swaption, and are therefore a logical choice. However, selecting the appropriate subset of European swaptions and determining which strike level to use are critical choices that directly impact the accuracy of the early-exercise valuation. A common industry practice is to calibrate models to diagonal swaptions, meaning swaptions whose maturity and the underlying swap tenor are equal, such as a $5Y \times 5Y$ swaption, usually with a 10-year or 30-year swaption strip depending on the contract. The primary issue is that diagonal swaptions often do not correspond to the maturities that are actually relevant for the Bermudan swaption priced, and the model is forced to fit instruments that the Bermudan swaption does not actually consider in its lifetime. For example, pricing a 20-year Bermudan using a 30-year diagonal swaption strip introduces unnecessary noise that extends significantly beyond the horizon of the lifetime of the Bermudan affecting the model parameters. While this does not necessarily prevent accurate valuation, selecting calibration instruments that are more closely aligned with the exercise dates of the Bermudan can improve consistency and ensure that the model dynamics are most relevant for the underlying swap that decides when the Bermudan is exercised.

To align the model dynamics with the structure of the derivative, this thesis utilizes a co-terminal calibration strategy. In this approach, each swaption used for the

calibration has an expiry and tenor that coincides with a Bermudan exercise date (in our case, 27.10.2042), and therefore the calibration set consists of the exact instruments that the Bermudan swaption considers as potential exercise opportunities. For example, if the Bermudan swaption has an option in 5 years to enter a swap for 10 years, the corresponding European swaption would be a 5Y $\times$ 10Y swaption. This strategy thus ensures that the calibration instruments directly reflect the relevant market dynamics for the Bermudan exercise decisions. Co-terminal calibration therefore provides a more intuitive and targeted basis for fitting the models, improving estimation of continuation values, more accurate exercise boundary placement, and overall more stable and accurate pricing. The primary drawback of this approach is that co-terminal expiry-tenor pairs might often have lower liquidity than diagonal options, but we resolve this by using the SABR model to interpolate between the quotes. The co-terminal calibration instruments used in this thesis correspond to the expiry-tenor grid shown in Table 2, where each instrument represents a European swaption aligned with one of the Bermudan exercise dates, totaling 13 instruments for 13 call dates.

Expiry $\times$ Tenor	**Option Call Date**	**Underlying Swap Period**
3M $\times$ 17Y	06.10.2025	27.10.2025 – 27.10.2042
1Y $\times$ 16Y	06.10.2026	27.10.2026 – 27.10.2042
2Y $\times$ 15Y	06.10.2027	27.10.2027 – 27.10.2042
3Y $\times$ 14Y	06.10.2028	27.10.2028 – 27.10.2042
4Y $\times$ 13Y	06.10.2029	27.10.2029 – 27.10.2042
5Y $\times$ 12Y	06.10.2030	27.10.2030 – 27.10.2042
6Y $\times$ 11Y	06.10.2031	27.10.2031 – 27.10.2042
7Y $\times$ 10Y	06.10.2032	27.10.2032 – 27.10.2042
8Y $\times$ 9Y	06.10.2033	27.10.2033 – 27.10.2042
9Y $\times$ 8Y	06.10.2034	27.10.2034 – 27.10.2042
10Y $\times$ 7Y	06.10.2035	27.10.2035 – 27.10.2042
11Y $\times$ 6Y	06.10.2036	27.10.2036 – 27.10.2042
12Y $\times$ 5Y	06.10.2037	27.10.2037 – 27.10.2042

Table 2: Co-terminal European swaptions used for interest rate model calibration, with call dates and underlying swap periods.

Using this co-terminal set of calibration instruments, we consider two different calibration strategies based on the strike level of the instrument:

1. ATM Calibration: The model is fitted to the at-the-money volatilities of the co-terminal swaptions at each expiry. ATM swaptions are highly liquid with tight bid-ask spreads, making them attractive due to their practicality and stability. This is often considered standard practice because it was in use historically, when Bermudans traded before pronounced market smiles and non-ATM points of the volatility cube became liquid enough to track and therefore has had time to be established and accepted (Andersen and Piterbarg, 2010a). However, its main theoretical drawback is the inconsistency between the Bermudan swaption and

its underlying core European swaptions. If the exercise region of the Bermudan swaption is off the ATM level, calibrating to ATM instruments anchors the volatility of the model to a region of the rate distribution the Bermudan rarely occupies. As such, the model may not fully capture the volatility dynamics in the relevant regions and may cause pricing inaccuracies. For example, a Bermudan with a non-ATM strike and a single exercise date calibrated to ATM swaptions yields a significantly different price than the corresponding European swaption, which is clearly not an attractive feature.

2. To resolve the main drawback of the ATM approach, Andersen and Piterbarg (2010a) suggest using a strike-targeted calibration strategy. In this approach, the model is calibrated to the co-terminal European swaptions at the Bermudan strike K (2.00% in this thesis). This localizes the calibration to the strike region most relevant to the Bermudan contract and accurately considers the rate movements at the moneyness where the exercise decisions are made based on the strike. With this strategy, a European swaption is modeled exactly as a single-exercise Bermudan and has the same value in the market as it does in the model.

These two calibration strategies allow us to assess the effect of the calibration instruments and the strike level on Bermudan pricing. ATM calibration provides a simple and stable benchmark that is consistent for multiple trades at once, but may be inaccurate for contracts that are rarely around the ATM level. Strike-targeted calibration is more localized, focusing on the strike region most relevant for the exercise decisions of the Bermudan swaption.

4.3 Model Implementations and Considerations

With the calibration strategies and market data presented above, we next detail the specific implementations and configurations of the three interest rate models evaluated in this thesis: the Hull-White one-factor model, the Hull-White two-factor model, and the one-factor Cheyette local volatility model. The three models are chosen based on their distinct theoretical properties and varying levels of complexity, allowing us to evaluate the trade-offs between flexibility, analytical tractability, and accuracy.

- HW1F: The simplest model. Features a single factor driving the short rate, and allows for closed-form zero-coupon bond pricing as well as analytical European swaption pricing. This model serves as the benchmark due to its efficiency, stability, and good theoretical features. Its primary limitation is its inability to replicate market volatility smiles and skews due to its deterministic volatility, and it can only be calibrated to a single volatility per expiry.

- HW2F: Andersen and Piterbarg (2010b) note that a common argument against one-factor models is that they underprice Bermudan swaptions even though they are widely used. The rationale is that higher de-correlation in rates has a positive impact on Bermudan swaption prices, and can intrinsically de-correlate

rates more than a single-factor model could, where the instantaneous correlation is always one between forward rate movements. However, they continue to state that the argument is flawed in that the relevant correlations for Bermudan swaptions are inter-temporal and can be easily manipulated through the choice of mean reversion. The HW2F model is included here to empirically test these claims and to observe whether the additional factor constitutes an improvement or not.

- One-factor local volatility Cheyette: This model has been viewed as a direct theoretical improvement over the HW1F framework by Andersen and Piterbarg (2010b) and we evaluate it as such. By incorporating a local volatility function rather than a deterministic one, the model should be able to capture volatility skews better without being too computationally intense or mathematically complex, theoretically managing to combine the best of both worlds.

With these three different interest rate models employed, we aim to evaluate whether the claims concerning them hold true, and whether one model is superior to another for Bermudan pricing and replicating market volatilities. Although the models vary significantly in their mathematical formulation, the calibration process for each involves calibrating only the volatility parameters, while the mean reversion and other parameters are left as user inputs. In this thesis, the mean reversion is fixed at 0.03 in all models. As noted by Andersen and Piterbarg (2010b), the mean reversion is often set to match the market prices of Bermudan swaptions in case they are observable - this was not the case in this thesis, and 0.03 was chosen as it is standard in current market conditions. Furthermore, with this value the models matched counterparty quotes from other observable interest rate derivatives, such as step-coupon callable bonds.

All interest rate models used in this thesis also share a unified calibration framework to ensure consistency across the numerical experiments. All three models are calibrated to the EUR market using mid-market quotes (the average of the bid and ask quotes) to eliminate bid-ask spread bias. All models employ the same numerical nonlinear least-squares solver, a Levenberg-Marquardt type, with step size 0.1 to fit the model parameters to the market quotes. The calibration minimizes the relative volatility, meaning the objective function is to optimize the relative difference between the model-implied and market normal volatilities. Normal volatility is specifically chosen because it allows the models to handle negative rates. Finally, all models utilize a step volatility structure combined with forward step interpolation, modeling the volatility parameters as piecewise constant functions of time to accurately capture the dynamic term structure of volatility. With forward step interpolation, the volatility is kept constant from one time node until the next, excluding the next time node. All simulations and pricing routines also adopt the same day-count conventions and dual-curve framework.

Apart from the unified framework, each model has its own features that deviate from those of the other models. These model-specific configurations used in the numerical software are summarized in Tables 3, 4, and 5.

Parameter	Value	Description
Pricing Method	Adaptive	Optimal numerical method between analytical formulas and Backward PDE automatically selected. Typically defaults to efficient analytic formulas when fitting European swaptions.

Table 3: HW1F model-specific configuration.

Parameter	Value	Description
Model Flavor	AG	Uses the Average Gaussian (AG) formulation, also known as G2++, as given in Section 3.2.
Pricing Method	Adaptive	Optimal numerical method between analytical formulas and BackwardPDE automatically selected. Typically defaults to efficient analytic formulas when fitting European swaptions.
Mean Reversion 1	0.03	The speed of mean reversion for the first factor (the "level" factor). A low value creates high persistence, modeling the long-term trend of the yield curve.
Mean Reversion 2	0.5	The speed of mean reversion for the second factor (the "slope" factor). A higher value allows this factor to capture short-term curve fluctuations and shocks.
Correlation	-0.9	The instantaneous correlation between the two factors. A strong negative correlation is essential to generate the "humped" volatility term structure observed in data and to decorrelate the short and long ends of the curve (Andersen and Piterbarg, 2010a).

Table 4: HW2F model-specific configuration.

Parameter	Value	Description
Pricing Method	Backward Finite Difference	Unlike the closed-form HW models, Cheyette requires a PDE grid solver to calculate option prices during the calibration process.
Tension (λ)	0.001	A regularization parameter to control the stiffness of the volatility surface. The calibrator includes the additional penalty $\lambda^2 \sum \frac{(\sigma_{j+1}-\sigma_j)^2}{(x_{j+1}-x_j)^2}$ in the overall calibration effect. A low value allows the model to fit quotes tightly with minimal smoothing.
Time Steps	1000	Defines the granularity of the time grid used in the finite difference solver. High granularity is required to capture the dynamics of the local volatility function.
Local Volatility Flavor	HW1F SIG	Initializes the local volatility calibration using a smart initial guess (SIG) derived from a standard HW1F model. Requires mean reversion as an input for both the underlying HW1F model and the Cheyette model, value 0.03 used for both.

Table 5: Cheyette model-specific configuration.

As Table 5 states, the Cheyette model uses the HW1F smart initial guess in its local volatility calibration. Solving for the full local volatility surface from an arbitrary starting point using a backward finite-difference grid can be computationally intensive, which is why the software utilizes the HW1F model as a smart initial guess for the local volatility surface calibration.

This approach takes advantage of the mathematical relationship between the Cheyette model and the standard HW1F model. If we consider the special case where the Cheyette local volatility is purely deterministic, the state variables of the model are given by

$$
\begin{aligned}
dx_t &= [y_t - \alpha x_t]dt + \sigma(t)dW_t \\
dy_t &= [\sigma^2(t) - 2\alpha y_t]dt \\
r_t &= \theta(t) + x_t \\
x_0 &= y_0 = 0.
\end{aligned}
$$

In this special case, the variable y_t becomes deterministic and satisfies the ODE $y' = \sigma^2 - 2\alpha y$. By looking for a function $\phi(t)$ such that the dynamics match a standard

short-rate evolution

$$d(x_t + \phi(t)) = -\alpha[x_t + \phi(t)]dt + \sigma(t)dW_t,$$

which gives

$$\phi' + \alpha\phi + y = 0,$$

which is an ODE that could be easily solved as usual. However, the solution is irrelevant as $\phi(t)$ is simply absorbed into $\theta(t)$ without affecting the model dynamics. This shows that x_t follows a standard HW1F model with mean reversion α and volatility σ. Consequently, the Cheyette local volatility model can be viewed as an extension of the HW1F model with a convexity correction. Using this approach, the calibration then proceeds as follows:

1. The HW1F model volatility is calibrated to the same European swaption data that is given to the Cheyette model.

2. The calibrated HW1F volatility is used to compute the initial nodes of the local volatility grid.

3. Going forward through the option maturities, the full local volatility surface is calibrated to the options data by minimizing the difference between Cheyette model-implied and market volatilities.

5 Numerical Results

In this chapter, we assess the results for the Bermudan swaption priced according to the setup described in Chapter 4. We first assess how well the calibration of the models succeeds, how efficient they are, how they model the volatilities, and what the call probabilities look like. We then compare the pricing of the Bermudan swaption by each model and calibration strategy, examining how they differ. We conclude the chapter by analysing how well the model-implied volatilities match market quotes outside of the calibration set, as the calibrated volatilities are naturally extremely close to the quotes. We evaluate each combination of model and calibration strategy and aim to determine which combination of model and strategy performs best.

5.1 Calibration Results

Table 6 shows how long each model takes to price the Bermudan swaption on average out of 20 runs. The results show that HW1F and HW2F have an essentially negligible difference, while Cheyette takes notably longer. This could matter in a scenario where there are many contracts to price, and is thus worth noting. However, pricing is often done in parallel and financial institutions have highly efficient computers for this purpose, and this issue is therefore unlikely to be a significant concern in practice.

Model	Average Pricing Time (s)
HW1F	3.10
HW2F	3.88
Cheyette	19.35

Table 6: Average Bermudan swaption pricing time for each interest rate model.

Table 7 displays each model and calibration set, and their respective Root Mean Square Error (RMSE) for the calibration procedure, both for price and volatility. All combinations of model and calibration strategy produce very small RMSEs around the same magnitude, and therefore every model appears to perform the calibration procedure well. This is expected, given that all models and calibration strategies are well-established in the literature and widely used in the industry.

Model and Calibration Set	RMSE (Market Price)	RMSE (Implied Volatility)
HW1F ATM	2.824×10^{-9}	1.808×10^{-11}
HW1F Strike-targeted	4.519×10^{-9}	3.013×10^{-11}
HW2F ATM	1.113×10^{-9}	1.555×10^{-11}
HW2F Strike-targeted	1.001×10^{-9}	1.314×10^{-11}
Cheyette ATM	4.842×10^{-9}	3.269×10^{-11}
Cheyette Strike-targeted	6.065×10^{-9}	3.998×10^{-11}

Table 7: Calibration errors for all models and calibration strategies.

Tables 8, 9, 10, and 11 present the calibrated volatilities of each model. The strike-targeted calibration generally produces a wider range of volatilities compared to ATM calibration. For example, the HW1F strike-targeted volatilities range from 0.6434% to 0.9631%, while the ATM-calibrated volatilities range from 0.6819% to 0.8421%. It is also useful to note the volatility structure of each model. As a single-factor model, HW1F contains one state variable and deterministic volatility, resulting in one calibrated volatility per expiry. In contrast, the HW2F model has two state variables that independently capture short-term and long-term rate movements, necessitating two volatility parameters per expiry and providing additional flexibility in fitting the term structure. The Cheyette local volatility model differs from both short-rate models through its local volatility formulation, which produces a volatility curve with three values for each expiry, providing the most complex volatility structure and the greatest flexibility with it. As the calibration is only performed on one instrument at expiry, the Cheyette local volatility function is only weakly constrained in the rate dimension and remains relatively flat across rate levels, behaving similarly to the HW models. Calibrating to multiple strikes would impose stricter shape constraints on $\sigma(t, x_t)$ and better utilise the flexibility of local volatility in the model. However, this introduces additional calibration complexity and risks overfitting or instability.

The calibrated volatility levels of each model are also of interest. The volatilities of the HW1F and Cheyette models are relatively similar in magnitude, while the HW2F model produces somewhat different volatility levels. This is consistent with the theoretical expectations, as the two state variables in HW2F fundamentally alter its calibration dynamics, distinguishing its parameter outputs from the single-factor models. It is also worth noting the behavior of the second volatility component in HW2F. The second factor primarily captures the slope and curvature of the term structure, and therefore its volatility remains relatively stable at shorter horizons and increases notably at longer expiries, where the single-factor dynamics are insufficient to capture the full shape of the term structure.

Date	Strike-Targeted Volatility	ATM Volatility
06.10.2025	0.9031%	0.8549%
06.10.2026	0.9427%	0.8748%
06.10.2027	0.9366%	0.8709%
06.10.2028	0.8929%	0.8451%
08.10.2029	0.8573%	0.8263%
07.10.2030	0.8475%	0.8261%
06.10.2031	0.8113%	0.8023%
06.10.2032	0.7714%	0.7784%
06.10.2033	0.7451%	0.7612%
06.10.2034	0.7250%	0.7545%
08.10.2035	0.6800%	0.7211%
06.10.2036	0.6434%	0.6819%

Table 8: HW1F calibrated volatilities.

	HW2F Strike-Targeted		HW2F ATM	
Date	**Vol 1**	**Vol 2**	**Vol 1**	**Vol 2**
06.10.2025	1.1052%	1.0299%	0.9467%	1.0287%
06.10.2026	1.0120%	1.0040%	0.9272%	1.0025%
06.10.2027	1.0223%	1.0017%	0.9192%	1.0021%
06.10.2028	0.9965%	1.0014%	0.9006%	1.0021%
08.10.2029	0.9403%	1.0037%	0.8682%	1.0030%
07.10.2030	0.8979%	1.0071%	0.8468%	1.0044%
06.10.2031	0.8902%	1.0076%	0.8491%	1.0039%
06.10.2032	0.8535%	1.0122%	0.8253%	1.0055%
06.10.2033	0.8147%	1.0209%	0.8029%	1.0090%
06.10.2034	0.7987%	1.0316%	0.7919%	1.0142%
08.10.2035	0.7934%	1.0476%	0.7953%	1.0223%
06.10.2036	0.7580%	1.0829%	0.7683%	1.0463%
06.10.2037	0.7427%	1.1359%	0.7434%	1.0951%

Table 9: HW2F calibrated volatilities.

Vol	Crv 1	Crv 2	Crv 3	Crv 4	Crv 5	Crv 6	Crv 7	Crv 8	Crv 9	Crv 10	Crv 11	Crv 12	Crv 13
1	0.9552%	0.9080%	0.9375%	0.9744%	0.8785%	0.8158%	0.8520%	0.8633%	0.7359%	0.6955%	0.7584%	0.6524%	0.6657%
2	0.9476%	0.9099%	0.9370%	0.9749%	0.8778%	0.8167%	0.8519%	0.8644%	0.7364%	0.6967%	0.7592%	0.6536%	0.6644%
3	0.9292%	0.9206%	0.9311%	0.9766%	0.8767%	0.8158%	0.8530%	0.8656%	0.7354%	0.6963%	0.7604%	0.6546%	0.6650%

Table 10: Cheyette local volatility curves, strike-targeted.

Vol	Crv 1	Crv 2	Crv 3	Crv 4	Crv 5	Crv 6	Crv 7	Crv 8	Crv 9	Crv 10	Crv 11	Crv 12	Crv 13
1	0.8413%	0.8689%	0.8795%	0.8712%	0.8402%	0.8234%	0.8373%	0.7995%	0.7374%	0.7502%	0.7385%	0.7136%	0.7328%
2	0.8421%	0.8685%	0.8789%	0.8709%	0.8407%	0.8244%	0.8368%	0.8003%	0.7391%	0.7519%	0.7395%	0.7148%	0.7318%
3	0.8412%	0.8691%	0.8793%	0.8712%	0.8404%	0.8238%	0.8373%	0.7995%	0.7388%	0.7522%	0.7395%	0.7143%	0.7322%

Table 11: Cheyette local volatility curves, ATM.

The call probabilities given by each model and calibration strategy are presented in Tables 12, 13, and 14. The total probability of exercise is similar across both calibration strategies within each model. In particular, the strike-targeted and ATM calibrations produce comparable aggregate call probabilities, although the distribution of those probabilities differs considerably.

When comparing the models, the probability of exercising differs considerably, indicating that the chosen interest rate model has a significant impact on optimal exercise behavior. The Cheyette model exhibits the most conservative early-exercise profile with a non-call probability exceeding 90% in both strategies, while HW1F provides call probabilities of around 89% and HW2F is notably lower at around 85%. This variation highlights that the underlying rate dynamics of the models substantially reshape the exercise boundary. In particular, the higher exercise probabilities of HW2F indicate a more aggressive early-exercise behavior driven by its additional factor, which helps to explain the divergence in its pricing profile discussed in Section 5.2.

These observations are consistent with the theoretical differences between the models considered. Since the Bermudan swaption should only be exercised when doing so is economically favorable, the optimal exercise decision depends strongly on the expected evolution of future rates.

Date	Strike-Targeted Call Probability	ATM Call Probability
06.10.2025	0.0000%	0.0000%
06.10.2026	1.2723%	0.6154%
06.10.2027	0.0694%	0.0591%
06.10.2028	0.7883%	1.2532%
06.10.2029	0.2642%	0.3008%
06.10.2030	0.6487%	0.7841%
06.10.2031	0.7919%	0.6819%
06.10.2032	0.5627%	0.6656%
06.10.2033	0.9172%	0.4465%
06.10.2034	1.3766%	1.5182%
08.10.2035	1.0472%	0.8336%
06.10.2036	2.4601%	2.2785%
06.10.2037	0.6225%	0.8523%
Non Call	**89.1814%**	**89.7123%**

Table 12: HW1F call probabilities.

Date	Strike-Targeted Call Probability	ATM Call Probability
06.10.2025	0.0000%	0.0000%
06.10.2026	0.6069%	1.8235%
06.10.2027	0.0804%	0.0000%
06.10.2028	1.4111%	0.5630%
06.10.2029	0.0170%	0.2001%
06.10.2030	0.7700%	1.1146%
06.10.2031	0.7832%	0.8083%
06.10.2032	1.1537%	1.0866%
06.10.2033	0.6428%	0.7086%
06.10.2034	1.7972%	1.1859%
08.10.2035	1.4719%	1.6159%
06.10.2036	4.1053%	4.1862%
06.10.2037	1.5042%	1.5259%
Non Call	**85.6565%**	**85.2020%**

Table 13: HW2F call probabilities.

Date	Strike-Targeted Call Probability	ATM Call Probability
06.10.2025	0.0000%	0.0000%
06.10.2026	1.0398%	0.4045%
06.10.2027	0.1568%	0.1671%
06.10.2028	0.6201%	1.0411%
06.10.2029	0.5026%	0.5029%
06.10.2030	0.3378%	0.9458%
06.10.2031	0.6320%	0.5184%
06.10.2032	1.0160%	0.5452%
06.10.2033	0.5696%	0.7454%
06.10.2034	0.8189%	0.5509%
08.10.2035	0.9007%	0.4752%
06.10.2036	1.6371%	2.5463%
06.10.2037	0.7384%	0.9637%
Non Call	**91.0302%**	**90.5938%**

Table 14: Cheyette call probabilities.

5.2 Bermudan Pricing Results

Before presenting the numerical pricing results, a brief note on sign convention is required. In the tables and figures throughout this section, the Bermudan swaption valuations are reported as negative figures, reflecting that we hold a short position in the Bermudan described in Table 1. However, for clarity and convenience during the comparative analysis, discussions regarding these valuations will focus on their absolute magnitude. Therefore, any reference to a price being ‘higher‘, ‘larger‘, or ‘increasing‘ indicates a more expensive option premium (a larger absolute liability), rather than a numerical value closer to zero.

Figure 3 presents the Bermudan swaption pricing results detailed in Table 1. As can be observed, the HW1F and Cheyette models produce comparable valuations under both calibration strategies, whereas the HW2F model noticeably diverges. This alignment is mathematically consistent; because HW1F and Cheyette are both driven by a single state variable, their underlying pricing mechanics remain similar. In contrast, the two state variables in the HW2F framework introduce additional dynamics in the evolution of interest rates that inherently alter the valuation.

Turning to the calibration strategies, Figure 4 details the pricing comparisons between strike-targeted and ATM calibrations for each model. Across the board, strike-targeted calibration consistently yields higher prices than ATM calibration, with the magnitude of this difference remaining relatively uniform across the models. This difference makes sense from a theoretical standpoint: because the strike of the Bermudan studied is currently relatively far from the ATM level, it sits on the wings of the volatility smile, where implied volatilities are higher. In turn, ATM volatilities typically represent the lowest point on this curve. Therefore, calibrating

to the specific strike captures this higher volatility premium and results in a more realistic valuation. Relying on ATM calibration for an instrument positioned far from the forward rate incorrectly anchors the volatility, ultimately misplacing the optimal exercise boundaries of the Bermudan swaption.

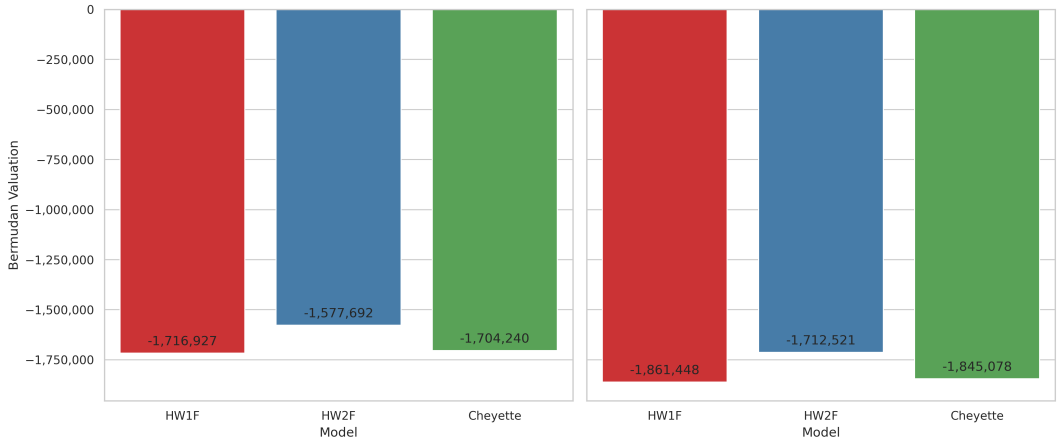


Figure 3: Bermudan swaption prices with different models and calibration strategies (ATM left, strike-targeted right).

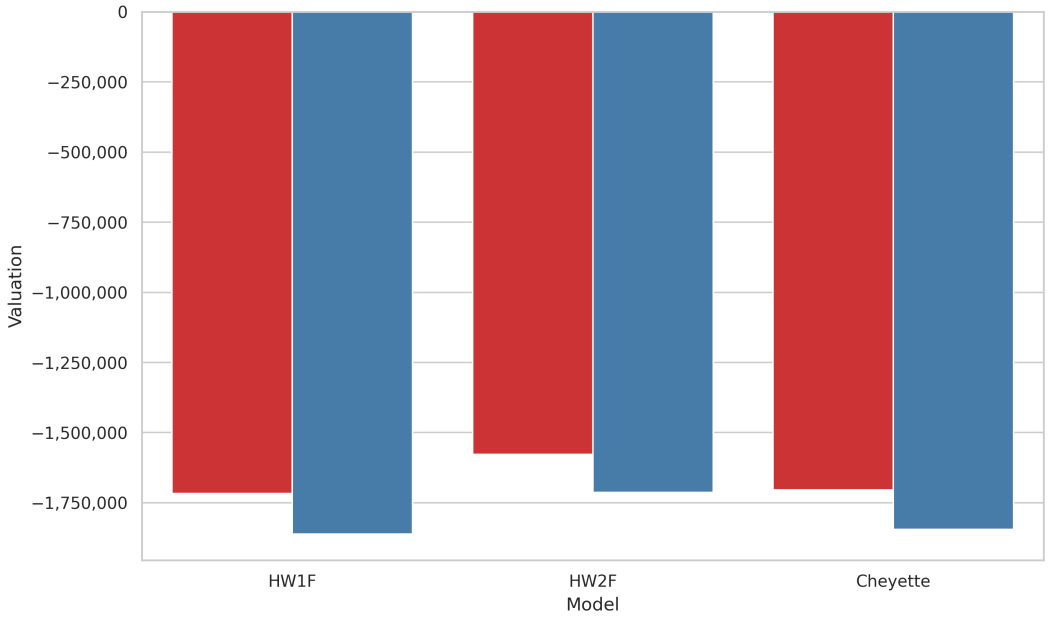


Figure 4: ATM (red) vs strike-targeted (blue) pricing comparison per model.

Figure 5 displays the prices in red for the annual Bermudan swaption specified in Table 1, and the prices in blue for a biennial version of the same swaption with call dates every other year. In theory, the Bermudan with more call dates should always be more expensive as it introduces additional optionality and more opportunities for

gains. The biennial version has all the same call dates as the annual version but fewer exercise dates, and should therefore be cheaper due to less optionality. The results are consistent with this, as for every combination of model and strategy, apart from the strike-targeted strategy for HW2F, the annual version has a higher price. While HW1F and Cheyette indeed follow this and produce relatively similar prices that are notably lower for the biennial swaption, the HW2F model performs poorly by comparison. In ATM calibration it produces only a minimal difference, and under strike-targeted calibration the pricing actually violates the theory. The most likely explanation for this is the parameterization of the model itself, as with different mean reversion or correlation values, more consistent results would presumably follow. Additionally, the model might simply require a richer calibration with more instruments and strike information to properly accommodate the model's additional factor, though this comes at the cost of added complexity. Figure 6 displays these differences between the annual and biennial versions more explicitly.

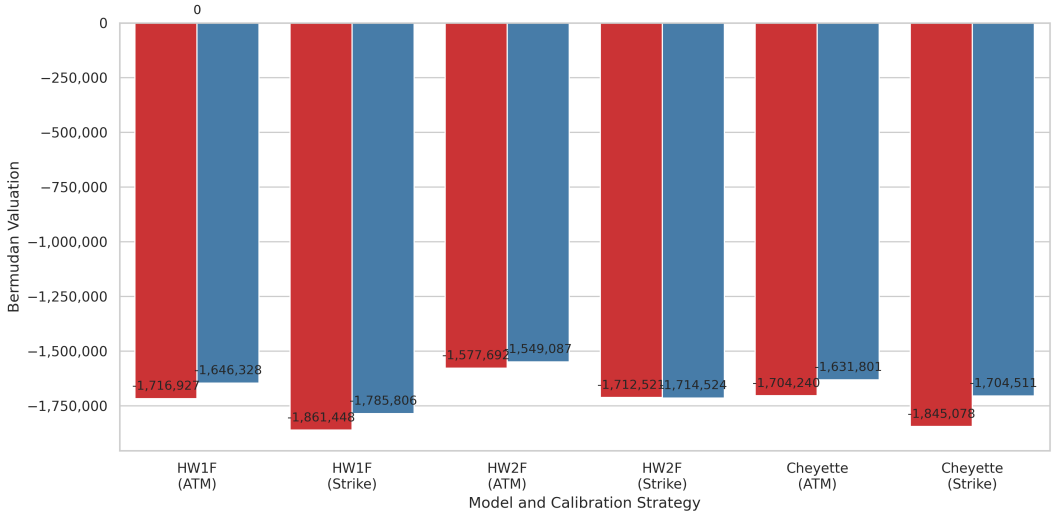


Figure 5: Effect of exercise date frequency on price, annual calls (red) vs biennial calls (blue).

Table 15 details the price of the Bermudan swaption with the Cheyette model and both calibration strategies using both the finite difference method as well as the Monte Carlo method (both described in Section 3.4.1) for comparative reasons. As can be observed, the differences between the prices given by the two methods are essentially negligible, though Andersen and Piterbarg (2010a) do explain that the finite-difference method is often better for the Cheyette model.

Calibration strategy	Monte Carlo	Finite difference method	Difference
ATM calibration	-1 703 969	-1 704 511	542
Strike-targeted calibration	-1 845 921	-1 844 234	-1 687

Table 15: Bermudan swaption prices for different pricing methods with Cheyette.

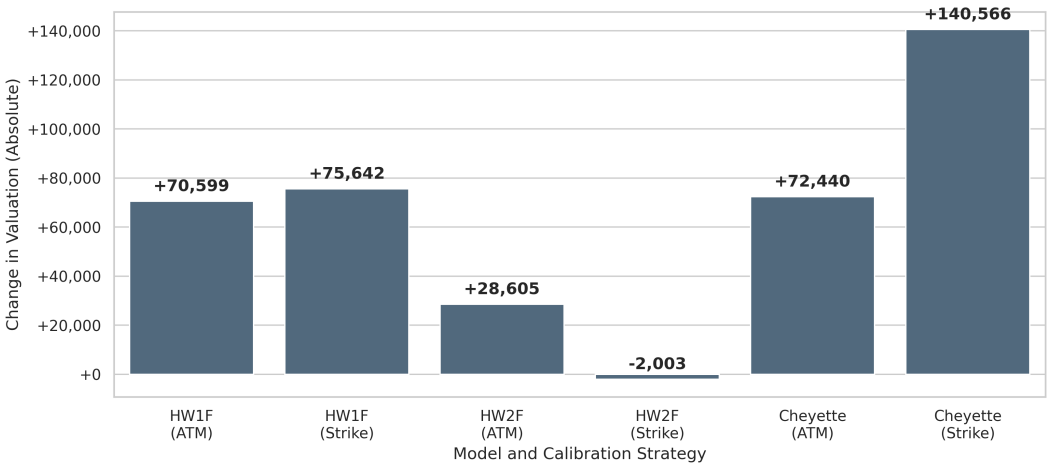
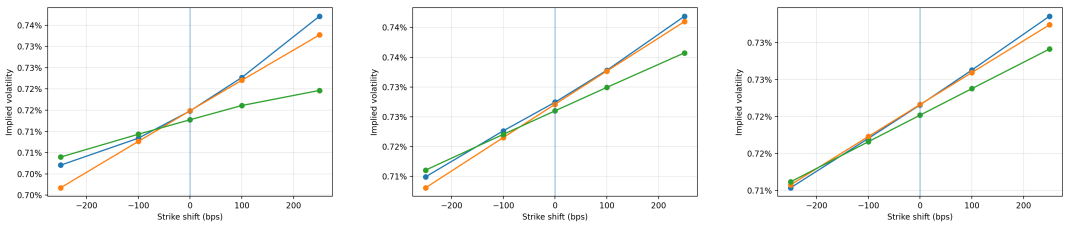


Figure 6: Annual vs. biennial call date schedule difference for models and strategies, derived from the prices in Figure 5.

5.3 Model-Implied Volatilities

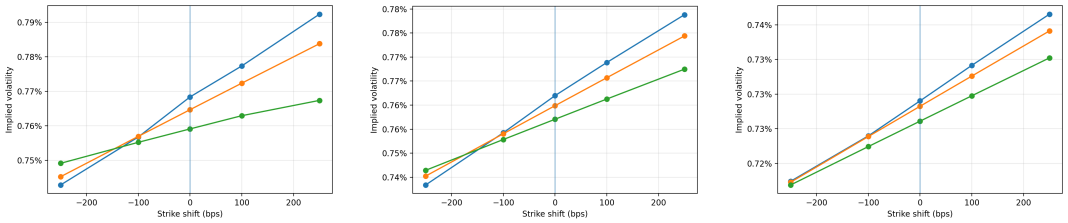
Figures 7–12 show model-implied volatilities for selected European swaptions spanning short, medium, and long expiries from the calibration set. Figures 7, 8, 9 plot model-implied volatilities at strikes ATM, ATM ± 100bps, and ATM ± 250bps, while Figures 10, 11, and 12 display volatilities at strikes 0.5ATM, 0.75ATM, ATM, 1.25ATM and 1.5ATM.



(a) 2Y × 15Y swaption **(b)** 6Y × 11Y swaption **(c)** 10Y × 7Y swaption

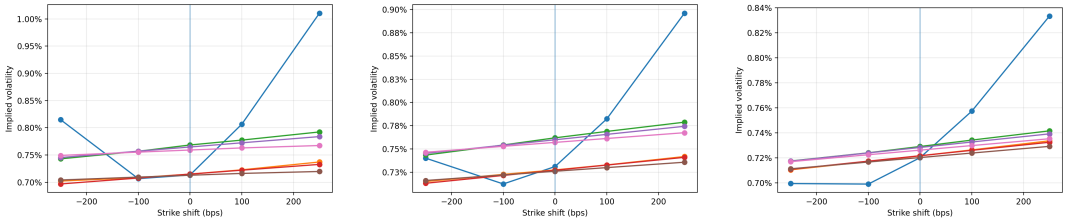
Figure 7: Volatility smile comparison across models at strike offsets, ATM calibration. Cheyette blue, HW1F orange, HW2F green.

Looking at the volatility skews across models, the Cheyette model consistently produces the steepest skew, followed closely by HW1F, while HW2F deviates more noticeably from both. This is consistent with the more flexible local volatility specification of the Cheyette model, which allows it to better differentiate implied volatilities across strikes. Furthermore, as the 2Y × 15Y swaptions especially show, the Cheyette model produces implied volatilities with some curvature, whereas the HW1F and HW2F model-implied volatilities are perfectly linear, highlighting the flexibility of the local volatility specification. However, when all three models are compared directly against market quotes in Figure 9, none of them adequately replicate the curvature of the market smile — a known structural limitation of affine interest



(a) 2Y × 15Y swaption **(b)** 6Y × 11Y swaption **(c)** 10Y × 7Y swaption

Figure 8: Volatility smile comparison across models at strike offsets, strike-targeted calibration. Cheyette blue, HW1F orange, HW2F green.



(a) 2Y × 15Y swaption **(b)** 6Y × 11Y swaption **(c)** 10Y × 7Y swaption

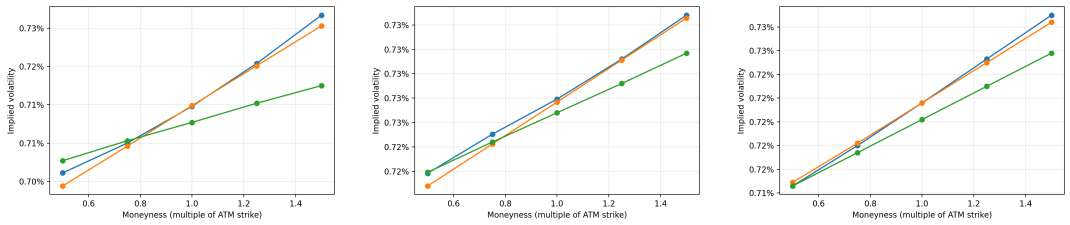
Figure 9: Volatility smile comparison across models at strike offsets, both calibration strategies and market quotes included. Market quotes blue; Cheyette strike-targeted green, ATM orange; HW1F strike-targeted blue, ATM red; HW2F strike-targeted pink, ATM brown.

rate models, which by construction can only produce linear skews rather than the convex smile observed in markets. Even in the case of the Cheyette model, this is to be expected since the model is initialised using HW1F, and calibrated to only one instrument/strike per expiry on a long-dated instrument. Under this setup, the optimiser has little information and reason to curve the local volatility function extensively, and the local volatility mechanism is therefore not fully utilized by the calibration scheme, resulting in nearly indistinguishable smile shapes.

The choice of calibration strategy appears to have a meaningful impact on the shape and level of the model-implied volatility skew. ATM calibration (Figures 7 and 10) anchors the models at the ATM volatility level, which constrains the skew and results in less variation in the volatilities across models and strikes. In contrast, strike-targeted calibration shifts the calibration anchor to match the strike of the Bermudan swaption, producing steeper and better differentiated volatility skews and an overall improved fit to the general level of market volatilities. This translates directly into more accurate Bermudan pricing, as was demonstrated in Section 5.2. Interestingly, the differences between the two calibration strategies appear to be more pronounced for Cheyette and HW1F than for HW2F, suggesting that HW2F is less responsive to the choice of calibration instruments, which is potentially another reason for its relatively weaker pricing performance.

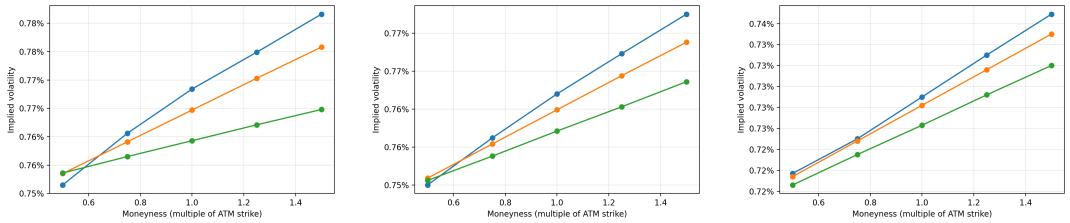
Finally, across all models and calibration strategies, the differences in model-implied volatilities are more pronounced for swaptions with a shorter expiry such

as the 2Y × 15Y than for longer expiries such as the 10Y × 7Y. This is consistent with the well-known empirical property that volatility skews tend to flatten with expiry. Swaptions with a longer expiry are less sensitive to short-rate movements, and averaging volatility over a longer time horizon tends to level out the difference of volatilities across strikes.



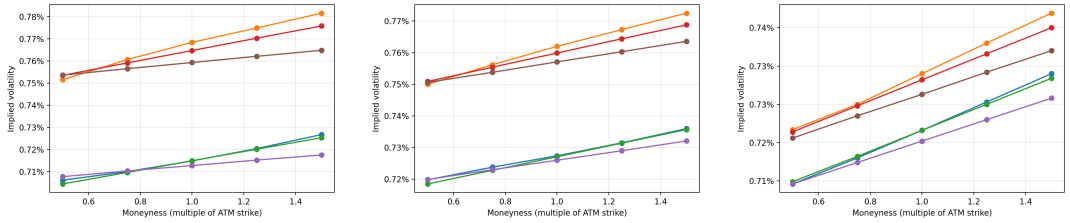
(a) 2Y × 15Y swaption **(b)** 6Y × 11Y swaption **(c)** 10Y × 7Y swaption

Figure 10: Volatility smile comparison across models at strike multiples of the ATM rate, ATM calibration. Cheyette blue, HW1F orange, HW2F green.



(a) 2Y × 15Y swaption **(b)** 6Y × 11Y swaption **(c)** 10Y × 7Y swaption

Figure 11: Volatility smile comparison across models at strike multiples of the ATM rate, strike-targeted calibration. Cheyette blue, HW1F orange, HW2F green.



(a) 2Y × 15Y swaption **(b)** 6Y × 11Y swaption **(c)** 10Y × 7Y swaption

Figure 12: Volatility smile comparison across models at strike multiples of the ATM rate, both calibration strategies. Cheyette strike-targeted orange, ATM blue; HW1F strike-targeted red, ATM green; HW2F strike-targeted brown, ATM purple.

6 Conclusion

This thesis has explored how the choice of interest rate model and calibration strategy affects the pricing of Bermudan swaptions. We began by introducing the problem and its relevance in risk management and quantitative finance, followed by the necessary mathematical and financial preliminaries. We then provided an overview of the methodology used in the numerical parts, most notably volatility and interest rate modeling, calibration methodology, and the overall pricing process of Bermudan swaptions. The numerical part first introduced the experimental setup, covering the target Bermudan swaption, the discounting and projection curves, market volatility interpolation via the SABR model, and the specific models and calibration strategies employed - HW1F, HW2F, and the Cheyette model, each paired with ATM and strike-targeted calibration strategies. These combinations were assessed with three research questions in mind: how accurately each replicates market volatility quotes, how well each prices the Bermudan swaption, and whether the additional factor in HW2F delivers meaningful value.

The calibration procedure succeeded well across all models, with no notable differences in calibration errors or call probabilities. In terms of the pricing of the Bermudan swaption, the calibration strategy proved to be the more significant choice compared to the model, with strike-targeted calibration consistently producing higher prices. The HW2F model exhibited a major theoretical inconsistency, pricing the Bermudan swaption with annual exercise dates below the biennial Bermudan, which should not occur. This could be fixed with better parameterization and calibration of the HW2F model. The HW1F and Cheyette models produced fairly similar prices under both calibration strategies, with the Cheyette requiring slightly more computation time but not excessively. The model-implied volatilities reinforce this, as HW2F produces the flattest skews, indicating limited ability to fit the shape of the market volatility surface. Cheyette produced the steepest skews and greatest variation across strikes, and exhibits some flexibility in the volatility structure as well, whereas HW1F and HW2F are perfectly linear. That said, none of the models are particularly effective in capturing the full curvature of market volatility smiles. This is partly a structural limitation of affine short-rate models, which by construction produce only approximately linear skews, and partly a consequence of the single-strike calibration scheme employed. With only one instrument per expiry, the local volatility flexibility of the Cheyette model is not fully realised. A richer calibration set spanning multiple strikes would better constrain the local volatility surface and likely yield more differentiated smile results. Nevertheless, the single-strike setup is well-suited for the purposes of this thesis. It ensures a consistent and practically motivated calibration framework across all models, keeping the focus on the comparative analysis of pricing and calibration strategies rather than maximising smile-fitting accuracy.

Based on these findings, the suggestion for financial institutions is to price Bermudan swaptions using either the HW1F or Cheyette model with strike-targeted calibration. The choice between the two models depends on the priorities of the institution: Cheyette offers greater flexibility and adaptability across rate environments, while HW1F offers simplicity and computational efficiency. HW2F is not recommended

with this configuration, although it is possible that better parameterization or a richer calibration set that better considers the second factor could improve its performance. However, this introduces notable additional complexity for what is a potential minor improvement, meaning that HW1F and Cheyette are likely still better options.

There are several avenues for future research on the topic. Within the current framework used, results might be improved by adjusting the calibration strategies to include a richer calibration set with multiple strikes considered, the SABR model specifications, or the model parameterization. The conclusions are also based on a single Bermudan swaption priced on a single date in a specific rate environment, which inherently limits their generalization. Repeating the experiments on multiple dates and different rate environments would make the analysis more robust and help assess whether the findings hold more broadly. The specific contract considered here has a long lifetime with a low strike and relatively infrequent call dates for what a Bermudan swaption can have, and a differently specified contract might also yield greater distinctions between the models. On the modeling side, incorporating stochastic volatility would be a natural extension, as stochastic volatility models are theoretically better at capturing the market volatility smiles and skews. A local stochastic volatility extension of the Cheyette model exists, which Andersen and Piterbarg (2010b) claim to be the best suited model for Bermudan swaptions, though the added complexity and computational demand need to be considered carefully in a practical setting. Finally, more advanced numerical methods, such as the machine learning model approach of Ech-Chafiq et al. (2023), who use regression trees and random forests, could be used to compute the continuation values to improve pricing accuracy.

References

Leif B. G. Andersen and Vladimir V. Piterbarg. *Interest Rate Modeling. Volume II: Term Structure Models*. Atlantic Financial Press, 2010a. ISBN 978-0-9841343-1-4.

Leif B. G. Andersen and Vladimir V. Piterbarg. *Interest Rate Modeling. Volume III: Products and Risk Management*. Atlantic Financial Press, 2010b. ISBN 978-0-9841343-2-1.

Alexander Antonov, Michael Konikov, and Michael Spector. SABR Spreads Its Wings. *Risk*, 26(8):58, 2013.

Alexandre Antonov, Michael Konikov, and Michael Spector. The Free Boundary SABR: Natural Extension to Negative Rates. *Available at SSRN 2557046*, 2015.

Louis Bachelier. Théorie de la Spéculation. In *Annales Scientifiques de L'École Normale Supérieure*, volume 17, pages 21–86, 1900.

Paolo Baldi. *Stochastic Calculus: An Introduction Through Theory and Exercises*. Springer, 2017. ISBN 978-3-319-62225-5. doi: 10.1007/978-3-319-62226-2.

Tomas Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 2009.

Fischer Black. The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3(1-2):167–179, 1976.

Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

Fischer Black, Emanuel Derman, and William Toy. A One-factor Model of Interest Rates and Its Application to Treasury Bond Options. *Financial Analysts Journal*, 46(1):33–39, 1990.

Damiano Brigo and Fabio Mercurio. *Interest Rate Models - Theory and Practice*. Springer, 2001.

Robert Brown. XXVII. A Brief Account of Microscopical Observations Made in the Months of June, July and August 1827, On the Particles Contained in the Pollen of Plants; and On the General Existence of Active Molecules in Organic and Inorganic Bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

Markus K Brunnermeier. Deciphering the Liquidity and Credit Crunch 2007–2008. *Journal of Economic perspectives*, 23(1):77–100, 2009.

Lin Chen. *Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management*. Blackwell publishers, 1996.

Oren Cheyette. Markov Representation of the Heath-Jarrow-Morton model. *SSRN Electronic Journal*, 2001.

Jaehyuk Choi, Minsuk Kwak, Chyng Wen Tee, and Yumeng Wang. A Black–Scholes User's Guide to the Bachelier Model. *Journal of Futures Markets*, 42(5):959–980, 2022.

Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC, 2003.

John C Cox. The Constant Elasticity of Variance Option Pricing Model. *Journal of Portfolio Management*, page 15, 1996.

John C Cox, Jonathan E Ingersoll, Stephen A Ross, et al. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407, 1985.

John Crank and Phyllis Nicolson. A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat-Conduction Type. *Mathematical Proceedings of the Cambridge Philosophical Society*, 43(1):50–67, 1947. doi: 10.1017/S0305004100023197.

Christian Crispoldi, Gérald Wigger, and Peter Larkin. *SABR and SABR LIBOR Market Models in Practice: With Examples Implemented in Python*. Springer, 2016.

Freddy Delbaen and Walter Schachermayer. A General Version of the Fundamental Theorem of Asset Pricing. *Mathematische Annalen*, 300(3):463–520, 1994.

Emanuel Derman and Iraj Kani. Riding on a Smile. *Risk*, 7(2):32–39, 1994.

L Uri Dothan. On the Term Structure of Interest Rates. *Journal of Financial Economics*, 6(1):59–69, 1978.

Daniel J Duffy. *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. John Wiley & Sons, 2006.

Bruno Dupire. Pricing With a Smile. *Risk*, 7(1):18–20, 1994.

Zineb El Filali Ech-Chafiq, Pierre Henry-Labordere, and Jérôme Lelong. Pricing Bermudan Options Using Regression Trees/Random Forests, 2023. URL https://arxiv.org/abs/2201.02587.

Damir Filipovic. *Term-Structure Models: A Graduate Course*. Springer Science & Business Media, 2009.

Jim Gatheral. *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, 2011.

Paul Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53. Springer, 2004.

John W Goodell. COVID-19 and Finance: Agendas for Future Research. *Finance Research Letters*, 35:101–512, 2020.

Patrick Hagan, Andrew Lesniewski, and Diana Woodward. Probability Distribution in the SABR model of Stochastic Volatility. In *Large deviations and asymptotic methods in finance*, pages 1–35. Springer, 2015.

Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing Smile Risk. *The Best of Wilmott*, 1:249–296, 2002.

Espen Haug. *The complete guide to option pricing formulas / by Espen Gaarder Haug*. McGraw-Hill, 2nd ed. edition, 2007. ISBN 9780071389976.

David Heath, Robert Jarrow, and Andrew Morton. Bond Pricing and the Term Structure of Interest Rates: A New Methodology For Contingent Claims Valuation. *Econometrica: Journal of the Econometric Society*, pages 77–105, 1992.

Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2): 327–343, 1993.

Thomas SY Ho and Sang-Bin Lee. Term Structure Movements and Pricing Interest Rate Contingent Claims. *The Journal of Finance*, 41(5):1011–1029, 1986.

John Hull. Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models. *The Journal of Derivatives*, 2:37–48, 1994.

John Hull and Alan White. Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies*, 3(4):573–592, 1990.

John C Hull and Sankarshan Basu. *Options, Futures, and Other Derivatives*. Pearson Education India, 2016.

International Swaps and Derivatives Association (ISDA). Key Trends in the Size and Composition of OTC Derivatives Markets in the First Half of 2024. December 2024. URL https://www.isda.org/a/GpbgE/Key-Trends-in-the-Size-and-Composition-of-OTC-Derivatives-Markets-in-the-First-Half-of-2024.pdf.

Farshid Jamshidian. An Exact Bond Option Formula. *The Journal of Finance*, 44(1): 205–209, 1989.

Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.

Ioannis Karatzas and Steven Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1988.

Ioannis Karatzas and Steven E Shreve. *Methods of Mathematical Finance*, volume 39. Springer, 1998.

Sven Klingler and Olav Syrstad. Life After LIBOR. *Journal of Financial Economics*, 141(2):783–801, 2021.

Francis A Longstaff and Eduardo S Schwartz. Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. *The Journal of Finance*, 47 (4):1259–1282, 1992.

Francis A Longstaff and Eduardo S Schwartz. Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies*, 14(1):113–147, 2001.

Donald W Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2): 431–441, 1963.

Robert Merton. Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

Paul Miron and Philip Swannell. Pricing and Hedging Swaps. *Euromoney Books*, 1991.

Peter Mörters and Yuval Peres. *Brownian Motion*, volume 30. Cambridge University Press, 2010.

Jorge Nocedal. Numerical Optimization. *Springer*, 2006.

Bernt Øksendal. Stochastic Differential Equations. In *Stochastic Differential Equations: An Introduction with Applications*, pages 38–50. Springer, 2003.

William H Press. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

Steven E Shreve. *Stochastic Calculus for Finance II: Continuous-time Models*. Springer, 2004.

Lai Van Vo and Huong Thi Thu Le. From Hero to Zero: The Case of Silicon Valley Bank. *Journal of Economics and Business*, 127:106–138, 2023.

Oldrich Vasicek. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

Norbert Wiener. Differential-Space. *Journal of Mathematics and Physics*, 2(1-4): 131–174, 1923.

Siyu Ye. Hull white practical guidance, 2023.

A Yield Curve Construction

Maturity Date	Discount Factor	Maturity Date	Discount Factor
9.7.2025	0.9996	2.7.2027	0.9652
16.7.2025	0.9993	3.1.2028	0.9551
4.8.2025	0.9982	3.7.2028	0.9454
2.9.2025	0.9967	2.7.2029	0.9241
2.10.2025	0.9952	2.7.2030	0.9018
3.11.2025	0.9936	2.7.2031	0.8785
2.12.2025	0.9922	2.7.2032	0.8546
2.1.2026	0.9907	4.7.2033	0.8303
2.2.2026	0.9893	3.7.2034	0.8059
2.3.2026	0.9880	2.7.2035	0.7817
2.4.2026	0.9866	2.7.2036	0.7577
4.5.2026	0.9851	2.7.2037	0.7343
2.6.2026	0.9838	2.7.2040	0.6689
2.7.2026	0.9825	3.7.2045	0.5799
3.8.2026	0.9810	4.7.2050	0.5108
2.9.2026	0.9796	2.7.2055	0.4517
1.10.2026	0.9782	2.7.2065	0.3546
4.1.2027	0.9739	2.7.2075	0.2813
2.4.2027	0.9696		

Table A1: Discount factors $P(0,T)$ used in the construction of the discounting curve.

Curve	Day Count Convention	Interpolation Method
ESTR OIS Discount Curve	ACT/360	LogLinear
Euribor 6M Forward Curve	ACT/360	LogLinear

Table A2: Curve construction parameters used, required to replicate the yield curves employed in the pricing and calibration procedures.

B SABR Model Specification

Parameter	Value	Description
Quote Type	Mid	Calibration uses the mid-market volatility, average of bid and ask quotes. Ensures the model reflects fair market value without bid-ask spread bias.
Strike Type	Relative	Strikes are defined relative to ATM (e.g., $K \pm 100$bps) rather than as fixed absolute rates. This ensures the smile moves dynamically with the forward rate.
Swaption Quote Type	Normal vol	Input volatilities quoted in normal volatilities, required to account for negative rates.
Flavor	Multiple strikes	The volatility cube is constructed using a range of strikes rather than a single strike, allowing the model to capture the skew and convexity of the market.
Strike Interpolation	Free-boundary AKS	Uses the free-boundary Antonov-Konikov-Spector expansion of the SABR model (Antonov et al., 2015), see Section 3.1. Allows for negative rates.
Surface Interpolation	Linear	Volatility between known tenor/expiry grid points is interpolated linearly. Helps with instability that might occur with higher-order splines.
Strike Extrapolation	Linear	For strikes outside the calibration strike range, volatility is extrapolated linearly to avoid unrealistic behavior in deep ITM/OTM zones.

Continued on next page...

Table B1 – continued from previous page

Parameter	Value	Description
Surface Extrapolation	Flat	For time horizons beyond the last liquid market point (> 50Y), volatility is held constant.
Missing Values Method	Free-boundary AKS	Missing data points in the input volatility surface are filled by calibrating the AKS SABR model to available data and interpolating the missing coordinates.
Precalibrate SABR	True	SABR parameters are precalibrated prior to interpolation to ensure stability and convergence of the volatility surface.
Beta (β)	0.5 (Fixed)	The elasticity parameter β is fixed at 0.5. Fixing β is standard practice to stabilize the calibration of the remaining parameters. A value of 0.5 represents a hybrid between Normal ($\beta = 0$) and Lognormal ($\beta = 1$) dynamics, allowing for negative rates.
Target Vol Basis	ACT/365	Annualization of volatility is calculated using the Actual/365 day count convention, consistent with standard swaption market practices.
SABR Parameters	$\alpha_{\text{SABR}}, \beta, \rho, \nu$	The four stochastic parameters governing the model: α (initial volatility), β (elasticity), ρ (correlation), and ν (volatility of volatility).

Table B1: Swaption volatility cube interpolation configuration of SABR model.

C SABR Calibration Results

	Swap Tenor												
Expiry	**1Y**	**2Y**	**3Y**	**4Y**	**5Y**	**6Y**	**7Y**	**8Y**	**9Y**	**10Y**	**12Y**	**15Y**	**20Y**
1M	1.16E-03	1.77E-03	2.85E-04	2.38E-04	3.02E-04	2.84E-04	3.21E-04	4.01E-04	4.35E-04	3.90E-04	4.84E-04	5.22E-04	4.50E-04
3M	2.23E-03	2.81E-03	6.59E-04	6.03E-04	5.63E-04	5.58E-04	5.51E-04	5.59E-04	5.23E-04	5.11E-04	5.22E-04	5.24E-04	5.54E-04
6M	5.30E-04	4.83E-04	4.55E-04	4.27E-04	4.05E-04	3.90E-04	3.82E-04	3.80E-04	3.80E-04	3.77E-04	3.39E-04	3.11E-04	3.10E-04
1Y	5.53E-04	4.55E-04	4.02E-04	3.70E-04	3.55E-04	3.44E-04	3.20E-04	2.96E-04	2.75E-04	2.68E-04	2.69E-04	2.74E-04	3.01E-04
2Y	4.13E-04	4.00E-04	4.04E-04	3.75E-04	3.56E-04	3.38E-04	3.40E-04	3.56E-04	3.79E-04	4.04E-04	4.19E-04	4.22E-04	4.25E-04
3Y	4.92E-04	4.36E-04	4.07E-04	3.81E-04	3.79E-04	3.87E-04	4.01E-04	4.19E-04	4.33E-04	4.36E-04	4.34E-04	4.41E-04	4.55E-04
4Y	4.50E-04	4.10E-04	3.97E-04	3.97E-04	4.05E-04	4.16E-04	4.12E-04	4.08E-04	4.06E-04	4.06E-04	4.12E-04	4.25E-04	4.38E-04
5Y	4.12E-04	4.04E-04	4.06E-04	4.05E-04	3.95E-04	3.87E-04	3.82E-04	3.80E-04	3.78E-04	3.77E-04	3.89E-04	4.04E-04	4.17E-04
6Y	4.01E-04	3.93E-04	3.82E-04	3.73E-04	3.66E-04	3.63E-04	3.61E-04	3.61E-04	3.59E-04	3.63E-04	3.71E-04	3.84E-04	3.94E-04
7Y	3.68E-04	3.57E-04	3.49E-04	3.44E-04	3.37E-04	3.36E-04	3.35E-04	3.35E-04	3.41E-04	3.46E-04	3.52E-04	3.63E-04	3.73E-04
8Y	3.33E-04	3.26E-04	3.17E-04	3.13E-04	3.13E-04	3.14E-04	3.15E-04	3.21E-04	3.27E-04	3.32E-04	3.36E-04	3.47E-04	3.57E-04
9Y	3.01E-04	2.98E-04	2.95E-04	2.96E-04	2.96E-04	2.98E-04	3.05E-04	3.13E-04	3.19E-04	3.23E-04	3.29E-04	3.37E-04	3.46E-04
10Y	2.78E-04	2.82E-04	2.83E-04	2.84E-04	2.85E-04	2.95E-04	3.05E-04	3.12E-04	3.16E-04	3.20E-04	3.28E-04	3.34E-04	3.41E-04
12Y	2.61E-04	2.68E-04	2.67E-04	2.84E-04	2.94E-04	2.98E-04	3.00E-04	3.01E-04	3.09E-04	3.15E-04	3.22E-04	3.22E-04	3.32E-04

Table C1: SABR global fit errors across swap tenors (1Y–20Y) and option expiries (1M–12Y).

Tenor	α_{SABR}
1Y	0.0287
2Y	0.0375
3Y	0.0391
4Y	0.0397
5Y	0.0399
6Y	0.0398
7Y	0.0396
8Y	0.0394
9Y	0.0393
10Y	0.0393
12Y	0.0391
15Y	0.0393
20Y	0.0400

Table C2: Calibrated α_{SABR} parameters by swap tenor.

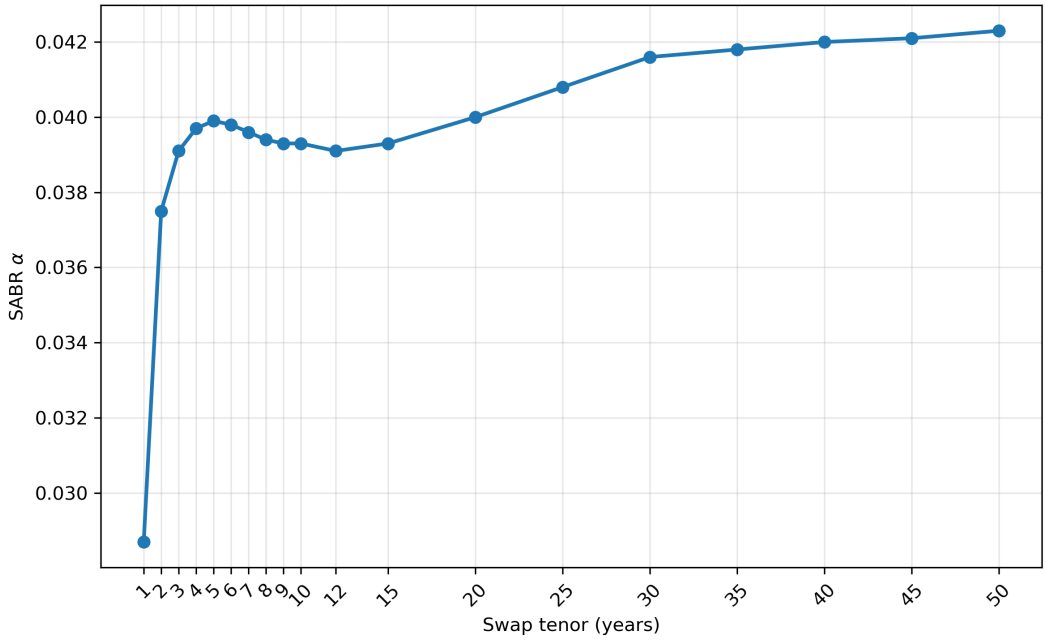


Figure C1: Term structure of α_{SABR}.

	Swap Tenor												
Expiry	**1Y**	**2Y**	**3Y**	**4Y**	**5Y**	**6Y**	**7Y**	**8Y**	**9Y**	**10Y**	**12Y**	**15Y**	**20Y**
1M	1.0468	0.9629	1.2839	1.2909	1.3518	1.3351	1.3607	1.4587	1.5016	1.5139	1.4723	1.5069	1.6072
3M	1.0751	0.9629	0.9948	0.9483	0.9211	0.9225	0.9344	0.9654	0.9653	0.9723	1.0042	1.0448	1.1186
6M	0.9297	0.8366	0.8220	0.8044	0.7756	0.7694	0.7686	0.7637	0.7579	0.7537	0.7646	0.7808	0.8133
1Y	0.6514	0.6071	0.5936	0.5784	0.5632	0.5788	0.5915	0.6026	0.6131	0.6249	0.6275	0.6333	0.6413
2Y	0.4814	0.4030	0.4361	0.4664	0.4954	0.5155	0.5362	0.5583	0.5807	0.6034	0.6028	0.6020	0.6007
3Y	0.5073	0.4467	0.4659	0.4841	0.5048	0.5220	0.5397	0.5591	0.5775	0.5951	0.5910	0.5898	0.5944
4Y	0.5081	0.4751	0.4834	0.4935	0.5046	0.5179	0.5283	0.5398	0.5523	0.5657	0.5643	0.5625	0.5698
5Y	0.4958	0.4965	0.4944	0.4912	0.4867	0.4911	0.4963	0.5031	0.5109	0.5191	0.5219	0.5211	0.5291
6Y	0.5116	0.5078	0.5018	0.4961	0.4915	0.4961	0.5009	0.5073	0.5140	0.5235	0.5245	0.5232	0.5303
7Y	0.5132	0.5026	0.4973	0.4937	0.4908	0.4946	0.4991	0.5042	0.5126	0.5203	0.5202	0.5194	0.5246
8Y	0.5036	0.4902	0.4859	0.4844	0.4838	0.4878	0.4909	0.4971	0.5039	0.5096	0.5083	0.5075	0.5126
9Y	0.4912	0.4751	0.4732	0.4733	0.4743	0.4770	0.4823	0.4866	0.4914	0.4952	0.4940	0.4912	0.4948
10Y	0.4795	0.4598	0.4594	0.4608	0.4624	0.4666	0.4702	0.4745	0.4772	0.4799	0.4781	0.4740	0.4771
12Y	0.4703	0.4545	0.4534	0.4590	0.4630	0.4646	0.4658	0.4681	0.4712	0.4747	0.4720	0.4664	0.4680

Table C3: SABR ν (vol-of-vol) parameters across swap tenors (1Y–20Y) and option expiries (1M–12Y).

	Swap Tenor												
Expiry	**1Y**	**2Y**	**3Y**	**4Y**	**5Y**	**6Y**	**7Y**	**8Y**	**9Y**	**10Y**	**12Y**	**15Y**	**20Y**
1M	-0.8024	-0.7614	-0.2523	-0.1299	-0.0143	-0.0118	0.0080	0.0215	0.0364	0.0670	0.0263	0.0016	-0.0307
3M	-0.7712	-0.7614	-0.0869	-0.0415	0.0154	0.0338	0.0406	0.0469	0.0542	0.0544	0.0274	-0.0006	-0.0720
6M	-0.4000	-0.2994	-0.2745	-0.2482	-0.2151	-0.2007	-0.1918	-0.1668	-0.1516	-0.1325	-0.1490	-0.1820	-0.2347
1Y	-0.3650	-0.3189	-0.2833	-0.2486	-0.2143	-0.1967	-0.1788	-0.1628	-0.1489	-0.1360	-0.1491	-0.1706	-0.2360
2Y	-0.2765	-0.2595	-0.2541	-0.2488	-0.2480	-0.2422	-0.2375	-0.2343	-0.2316	-0.2290	-0.2416	-0.2617	-0.3021
3Y	-0.3484	-0.3276	-0.3211	-0.3152	-0.3133	-0.3110	-0.3095	-0.3097	-0.3089	-0.3074	-0.3119	-0.3265	-0.3650
4Y	-0.3900	-0.3760	-0.3689	-0.3643	-0.3614	-0.3626	-0.3612	-0.3619	-0.3635	-0.3661	-0.3718	-0.3826	-0.4137
5Y	-0.4256	-0.4124	-0.4061	-0.3993	-0.3945	-0.3962	-0.3988	-0.4033	-0.4087	-0.4144	-0.4241	-0.4354	-0.4573
6Y	-0.4459	-0.4382	-0.4309	-0.4243	-0.4190	-0.4207	-0.4233	-0.4269	-0.4307	-0.4373	-0.4457	-0.4567	-0.4757
7Y	-0.4550	-0.4531	-0.4482	-0.4443	-0.4412	-0.4434	-0.4458	-0.4488	-0.4544	-0.4593	-0.4661	-0.4765	-0.4921
8Y	-0.4595	-0.4645	-0.4618	-0.4605	-0.4600	-0.4630	-0.4654	-0.4704	-0.4753	-0.4791	-0.4834	-0.4917	-0.5040
9Y	-0.4617	-0.4743	-0.4746	-0.4758	-0.4770	-0.4805	-0.4864	-0.4914	-0.4959	-0.4999	-0.5017	-0.5050	-0.5146
10Y	-0.4647	-0.4863	-0.4887	-0.4911	-0.4930	-0.5008	-0.5077	-0.5130	-0.5185	-0.5245	-0.5240	-0.5211	-0.5262
12Y	-0.4674	-0.4838	-0.4901	-0.4977	-0.5018	-0.5057	-0.5062	-0.5092	-0.5128	-0.5155	-0.5142	-0.5115	-0.5140

Table C4: SABR ρ (correlation) parameters across swap tenors (1Y–20Y) and option expiries (1M–12Y).