Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Johan Hagenbjörk
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Johan Hagenbjörk
Abstract

The global fixed income market is an enormous financial market whose value by far exceeds that of the public stock markets. The interbank market consists of interest rate derivatives, whose primary purpose is to manage interest rate risk. The credit market primarily consists of the bond market, which links investors to companies, institutions, and governments with borrowing needs. This dissertation takes an optimization perspective upon modeling both these areas of the fixed-income market. Legislators on the national markets require financial actors to value their financial assets in accordance with market prices. Thus, prices of many assets, which are not publicly traded, must be determined mathematically. The financial quantities needed for pricing are not directly observable but must be measured through solving inverse optimization problems. These measurements are based on the available market prices, which are observed with various degrees of measurement noise. For the interbank market, the relevant financial quantities consist of term structures of interest rates, which are curves displaying the market rates for different maturities. For the bond market, credit risk is an additional factor that can be modeled through default intensity curves and term structures of recovery rates in case of default. By formulating suitable optimization models, the different underlying financial quantities can be measured in accordance with observable market prices, while conditions for economic realism are imposed.

Measuring and managing risk is closely connected to the measurement of the underlying financial quantities. Through a data-driven method, we can show that six systematic risk factors can be used to explain almost all variance in the interest rate curves. By modeling the dynamics of these six risk factors, possible outcomes can be simulated in the form of term structure scenarios. For short-term simulation horizons, this results in a representation of the portfolio value distribution that is consistent with the realized outcomes.
from historically observed term structures. This enables more accurate measurements of interest rate risk, where our proposed method exhibits both lower risk and lower pricing errors compared to traditional models.

We propose a method for decomposing changes in portfolio values for an arbitrary portfolio into the risk factors that affect the value of each instrument. By demonstrating the method for the six systematic risk factors identified for the interbank market, we show that almost all changes in portfolio value and portfolio variance can be attributed to these risk factors. Additional risk factors and approximation errors are gathered into two terms, which can be studied to ensure the quality of the performance attribution, and possibly improve it.

To eliminate undesired risk within trading books, banks use hedging. Traditional methods do not take transaction costs into account. We, therefore, propose a method for managing the risks in the interbank market through a stochastic optimization model that considers transaction costs. This method is based on a scenario approximation of the optimization problem where the six systematic risk factors are simulated, and the portfolio variance is weighted against the transaction costs. This results in a method that is preferred over the traditional methods for all risk-averse investors.

For the credit market, we use data from the bond market in combination with the interbank market to make accurate measurements of the financial quantities. We address the notoriously difficult problem of separating default risk from recovery risk. In addition to the previous identified six systematic risk factors for risk-free interests, we identify four risk factors that explain almost all variance in default intensities, while a single risk factor seems sufficient to model the recovery risk. Overall, this is a higher number of risk factors than is usually found in the literature. Through a simple model, we can measure the variance in bond prices in terms of these systematic risk factors, and through performance attribution, we relate these values to the empirically realized variances from the quoted bond prices.

Mätning och hantering av risker är nära kopplat till mätningen av de underliggande finansiella storheterna. Genom en datadriven metod kan vi visa att sex systematiska riskfaktorer kan användas för att förklara nästan all varians i räntekurvorerna. Genom att modellera dynamiken i dessa sex riskfaktorer kan tänkbara utfall för räntekurvor simuleras. För kortsiktiga simuleringshorisonter resulterar detta i en representation av fördelningen av
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

portföljvärden som väl överensstämmer med de realiserade utfallen från historiskt observerade räntekurvor. Detta möjliggör noggrannare mätningar av ränterisk där vår föreslagna metod uppvisar såväl lägre risk som mindre prissättningsfel jämfört med traditionella modeller.

Vi föreslår en metod för att dekomponera portföljutvecklingen för en godtycklig portfölj till de riskfaktorer som påverkar värden för respektive instrument. Genom att demonstrera metoden för de sex systematiska riskfaktorerna som identifierats för räntemarknaden visar vi att nästan all portföljutveckling och portföljvarians kan härledas till dessa riskfaktorer. Övriga riskfaktorer och approximationsfel samlas i två termer, vilka kan användas för att säkerställa och eventuellt förbättra kvaliteten i prestationshäradningen.

För att eliminera oönskad risk i sina tradingböcker använder banker sig av hedging. Traditionella metoder tar ingen hänsyn till transaktionskostnader. Vi föreslår därför en metod för att hantera riskerna på räntemarknaden genom en stokastisk optimeringsmodell som också tar hänsyn till transaktionskostnader. Denna metod bygger på en scenarioapproximation av optimeringsproblemet där de sex systematiska riskfaktorerna simuleras och portföljvariansen vägs mot transaktionskostnaderna. Detta resulterar i en metod som, för alla riskaverta investerare, är att föredra framför de traditionella metoderna.

På kreditmarknaden använder vi data från obligationsmarknaden i kombination räntemarknaden för att göra noggranna mätningar av de finansiella storheterna. Vi angriper det erkänt svåra problemet att separera fallissemangsgiving från återvinningsrisk. Förutom de tidigare sex systematiska riskfaktorerna för riskfri ränta, identifierar vi fyra riskfaktorer som förklarar nästan all varians i fallissemangsgansixiteter, medan en enda riskfaktor tycks räcka för att modellera återvinningsrisken. Sammanlagt är detta ett större antal riskfaktorer än vad som brukar användas i litteraturen. Via en enkel modell kan vi mäta variansen i obligationspriser i termer av dessa systematiska riskfaktorer och genom prestationshäradningen relatera dessa värden till de empiriskt realiserade varianserna från kvoterade obligationspriser.
For Annie
During my undergraduate studies, I was impressed by those students whose papers barely glanced over the basic concepts and jumped right into explaining the more complex theories briefly. To them, the basics seemed too trivial to justify an explanation and the complex theories seemed rather simple. Meanwhile, I struggled with understanding how the more complex theories were built upon the basic concepts and how everything was related. It took me several years to understand that this was not a sign of knowledge but rather ignorance and laziness. By skipping ahead, there is no need to understand the basics and how different concepts relate to each other.

I like to imagine the field of research as a jigsaw puzzle where the pieces are given to each researcher but without the box revealing the full picture. Each researcher thus spends time piecing together pieces from different parts of the puzzle. When someone made enough progress, this researcher tries to describe what their part of the puzzle looks like in a research paper. When someone gets enough experience, that researcher gives an attempt to depict the full picture or a part of the picture in a book. There are always slight variations in the pictures depending on the techniques used in what areas they highlight.

During my years as a doctoral student, I have spent lots of time studying rather few pieces in detail, always feeling that the bigger picture was missing. At the same time, it is hard to grasp the concept of a windmill, if you have never seen one for yourself. Therefore, I have not been able to see a full picture until a very late stage in this process. I probably should have read more books at an earlier stage, but I am not sure whether I had the comprehension to grasp the content at the time.

Now when I finally see a picture in front of me, this dissertation is my attempt to describe it. Although I am not writing a textbook, I feel obligated
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

to depict the full picture in accordance with my view of knowledge as described above. First and foremost, for my own sake, not necessarily in full detail but at least the broad strokes. In academia, however, there seems to be a consensus that “less is more” and I am rather sure that this is not due to ignorance. At the same time, I have read too many dissertations where I have been lost after only a few sentences. I have always felt that when space is not constrained, it is easier for the enlighten one to skip ahead, than for the uninformed to find that knowledge elsewhere. So, if someone finds parts of this dissertation to trivial, I truly salute you. Congratulations, you may skip ahead!

Johan Hagenbjörk
Granö, Västerbotten, 25 July 2019
In August 2006, my then-girlfriend, now wife, and I set course for Linköping where I started a 4.5-year engineering program in applied physics and electro-engineering, whereas she studied business law. I found the education interesting but challenging and ended up staying at the university longer than first anticipated. I now consider my studies to be completed.

First, I would like to thank my supervisor, Jörgen Blomvall, for giving me the opportunity to earn my doctorate at Linköping University. I have certainly learned a lot from you over the years, starting in 2010 when I enrolled in the mathematical finance profile. I would like to thank my colleague and friend, Pontus Söderbäck, for the journey we have shared as Ph.D. students. I honestly do not know if I would have made it to the end without your support, especially in the course “Probability”.

I would like to thank everyone living in “kollektivet” for letting me stay there for a bargain—especially the first-generation residents, Johan and Morgan, whom I had a blast living with. I would like to thank Benjamin for being a reliable gym-buddy, and your whole family for welcoming me into your home during the last six months in Linköping.

I would especially like to thank my family for their consistent support during my time in Linköping. A special thanks to my parents and in-laws for taking care of Annie while giving me the opportunity to write this dissertation.

The biggest thanks I owe to my wife, Maja. You have shown me endless patience and support during the 6.5 years I have been commuting to Linköping, often working seemingly constantly. For this, I am forever grateful.

Johan Hagenbjörk
Solna, 8 December 2019
List of Papers

The following papers are appended and will be referred to by their numerals.


Author Statements

The papers in this dissertations are all joint work with my supervisor, Jör- gen Blomvall, however, the heavy lifting (pun intended) has been done by Hagenbjörk. The simulations behind the papers are rather computationally demanding, especially for paper III. For this purpose, Hagenbjörk developed a sophisticated portfolio management system in C++ as an extension to the QuantLib framework (Ametrano and Ballabio, 2000–2019). This system has been used in the first three papers. For Paper IV, Hagenbjörk wrote a parser that takes a Reuters structure string and returns the proper bond object from QuantLib. All simulations in this dissertation are scripted in Python by Hagenbjörk. All single-curve term structures are estimated in the MATLAB framework written by Blomvall.

I Blomvall proposed the research idea and Hagenbjörk implemented it rather independently. Blomvall suggested the evaluation method, the corresponding statistical test and designed the kernel density estimator. Hagenbjörk was responsible for writing the paper.

II Blomvall developed all the mathematics behind the performance attribution framework and documented it thoroughly. In the paper, Hagenbjörk presented the framework in a different order. Blomvall gave frequent feedback on the paper and was responsible for the discussion part. Hagenbjörk implemented the framework and scripted the simulation.

III Blomvall proposed the research idea and the choice of method based on the master thesis by Uhrberg and Sjöholm (2013), who used another approach. Hagenbjörk implemented the idea with guidance from
Blomvall, who also provided help finding the problems that made the hedge unstable. Blomvall also wrote the solver. Hagenbjörk wrote the paper and independently developed the stochastic dominance tests. Besides giving feedback on the paper, Blomvall contributed by writing the discussion part.

IV Blomvall proposed the research topic, which was an extension of the master thesis by Hagenbjörk. Hagenbjörk found the suitable mathematical way to model credit risk, where Blomvall contributed with some critical pieces. Blomvall provided an AMPL implementation for measuring multiple term structures that Hagenbjörk modified and extended with the credit risk part. Hagenbjörk was responsible for setting up the data-retrieval from EIKON and constructed the MySQL database. Hagenbjörk was responsible for writing the paper.
# Contents

Abstract i

Sammanfattning iii

Preface vii

Acknowledgments ix

List of Papers xi

Author Statements xiii

1 Part 1 - Introduction 1

1 Introduction 3

1.1 Data 9

1.2 Methodology 10

1.3 Outline 13

2 Market background 15

xv
# Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

## Part 2 - Interest Rate Risk

### 3 Interest Rates

#### 4 Measuring Term Structures of Interest Rates

- 4.1 Measuring Realistic Term Structures
- 4.2 Measuring Multiple Term Structures

#### 5 Measuring the Systematic Risk Factors of Interest Rates

#### 6 Performance Attribution

#### 7 Measuring Interest Rate Risk

- 7.1 Historical Measures of Interest Rate Risk
- 7.2 Accurate Measurements of Interest Rate Risk
  - 7.2.1 Modeling Univariate Distributions for Risk Factors
  - 7.2.2 Modeling Joint Variation using Copulas
  - 7.2.3 Measuring Risk by Simulating Term Structures

#### 8 Hedging Interest Rate Risk

- 8.1 Traditional Hedging Methods
- 8.2 Flexible Hedging using Stochastic Programming

## Part 3 - Credit Risk

### 9 Credit Risk

- 9.1 Modeling Credit Risk
- 9.2 Defaults
  - 9.2.1 Liquidation
  - 9.2.2 Reconstruction
CONTENTS

9.2.3 Indentures, Covenants and Debt Acceleration . . . . 73
9.3 Recovery in Practice . . . . . . . . . . . . . . . . . . . . . . . 74
9.4 Credit Ratings . . . . . . . . . . . . . . . . . . . . . . . . . . 75
9.5 Scoring Models . . . . . . . . . . . . . . . . . . . . . . . . . . 77
9.6 Structural Models . . . . . . . . . . . . . . . . . . . . . . . . . 78

10 Reduced-Form Models . . . . . . . . . . . . . . . . . . . . . . . . 81
10.1 Poisson Processes . . . . . . . . . . . . . . . . . . . . . . . . 82
10.2 Technical Details . . . . . . . . . . . . . . . . . . . . . . . . 85
10.2.1 Building Blocks . . . . . . . . . . . . . . . . . . . . . . . 87
10.2.2 Defaultable Coupon Bond . . . . . . . . . . . . . . . . 90
10.2.3 Recovery Schemes . . . . . . . . . . . . . . . . . . . . . 91
10.3 Dynamics of Intensity-Based Models . . . . . . . . . . . . . . 92
10.3.1 A Two-Factor Gaussian model . . . . . . . . . . . . . 92
10.3.2 Cox-Ingersol-Ross Models . . . . . . . . . . . . . . . . 93
10.4 Risk Premia . . . . . . . . . . . . . . . . . . . . . . . . . . . . 94

11 Measuring the Financial Quantities on the Credit Risk Mar-
ket . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
11.1 Disentangling Recovery Rates from Hazard Rates . . . . . . 98
11.2 Pricing Constraints for Defaultable Bonds . . . . . . . . . . . 99
11.3 Measuring Hazard and Recovery Rates . . . . . . . . . . . . . 99
11.3.1 Measuring Default Intensity . . . . . . . . . . . . . . . 100
11.3.2 Measuring Recovery Rates . . . . . . . . . . . . . . . . 100
11.3.3 Setting up the Optimization Model . . . . . . . . . . . . 100
11.3.4 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102
4 Part 4 - Contribution 103

12 Contribution 105

12.1 Answers to the Research Questions 105

12.1.1 Answer to Research Question 1a 105

12.1.2 Answer to Research Question 1b 107

12.1.3 Answer to Research Question 1 107

12.1.4 Answer to Research Question 2a 108

12.1.5 Answer to Research Question 2 110

12.2 Fulfillment of the Purpose 111

13 Future Research 113

Appended Papers

Simulation and Estimation of Interest Rate Risk
- Paper I

A Generic Framework for Monetary Performance Attribution
- Paper II

Hedging of Interest Rate Risk Using a Stochastic Programing Approach
- Paper III

Identification of the Credit-Risk Related Systematic Risk Factors from Bond Prices
- Paper IV
1

Introduction
Introduction

Finance is a field that is concerned with the allocation (investment) of assets and liabilities over space and time, often under conditions of risk or uncertainty.

This definition is taken from the Finance article in Wikipedia (2019) and fits surprisingly well with the view presented in this dissertation. The allocation of assets can be viewed as a problem vastly complicated by the space of possible assets to invest in, the time horizon of the investment, risk, and uncertainty. When solving such an allocation problem, one would ideally like as good a solution as possible, with regard to some criterion. Finding such optimal solutions to a problem is the essence of the field of optimization. Approaching the field of finance via optimization is not a new idea, but traces back to the seminal paper on portfolio selection by Markowitz (1952). However, determining optimal solutions to practically useful theoretical models poses immense challenges that are hard to overcome. To explain why, we start by further breaking down and studying the definition of finance presented above.

Real-world financial problems always involve some degree of uncertainty, meaning that the probabilities of different outcomes for the investments are unknown. In a time where the field of probability theory still was in its infancy, Knight (1921) suggested the term risk to denote measurable uncertainty. According to Holton (2004), this definition was criticized as it ignored the exposure part and solely focused on the uncertainty part. Exposure implies that one cares about the outcome of the uncertainty, and some interpretations of risk instead focus solely on the exposure part and ignore
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

the uncertainty. Approaching uncertainty via risk requires\(^1\) the probabilities of the outcomes to be known or somehow estimated. If the multivariate probability distribution function of all financial assets would be known, the investment problem could theoretically be solved to optimality, given the utility function of the investor. The fact that the multivariate distribution function neither is observable nor stationary renders this exhaustive approach practically impossible. Instead, the distribution function often needs to be approximated, and different optimization methods pose different kinds of restrictions on this approximation. In practice, the utility function is often approximated by considering expected return and risk, which are weighed against each other in the objective function. Contrary to uncertainty, risk is something that can be measured by assigning probabilities to different outcomes. Holton (2004) argues that since the uncertainty not perceived cannot be defined, our perception of uncertainty is all that we can hope to define operationally. At best, we can operationally define our perception of risk through different risk measures. Instead of asking if a risk measure captures risk, one should, therefore, ask whether it is useful. Regardless of risk measure, mathematical modeling of the uncertainty is necessary to quantify risk. To obtain practically useful results, mathematical models need to be accurate descriptions of reality.

Allocating resources over space implies choosing the assets where to invest. The enormous number of assets available in the financial market severely complicates decision-making. On the bond market, each issuer may have several bonds of different maturities, leading to over a hundred thousand market traded bonds globally (World Federation of Exchanges, 2015). Considering the over-the-counter derivatives market, which consists of instruments that are not traded on any exchange, numerous new derivatives with different maturities are created each and every day.

Allocating resources over time often involves long investment horizons but also implies updating the solution to the investment problem as new information arrives. Since new information arrives seemingly continuously, this would imply solving the investment problem infinitely often. However, reallocating is never frictionless, as market liquidity and transaction costs, including bid-ask spreads, must be taken into account. These market frictions call for the use of multi-period models.

In a world where the returns, \(r\), of all assets are normally distributed,

\(^1\)Robust optimization instead handles uncertainty by regarding the worst possible (\(P \neq 0\)) outcome regardless of probability.
\( r \sim N(\mu, \sigma) \), the expected return, \( \mu \), and the volatility, \( \sigma \), fully specifies the distribution, where risk corresponds to volatility. The seminal mean-variance model by Markowitz (1952), lets a risk-averse investor find the optimal asset allocations in such a frictionless market. Here, returns are assumed to be multivariate normally distributed, \( r \sim N(\mu, C) \), where \( \mu \) is a vector containing expected returns and \( C \) is the covariance matrix of the assets’ returns. Studying normally distributed returns allows for elegant mathematical modeling, which facilitates financial analysis, especially for academic purposes. A problem with this approach is that it is not an accurate representation of reality since the observed asset returns are not normally distributed but exhibit heteroscedasticity, excess kurtosis, and skewness, as shown by Mandelbrot (1963). This means that the volatility is non-stationary and tends to be clustered; extreme events are more likely to occur than predicted by the normal distribution (fat tails); the distribution is slightly skewed, with more probability mass towards one of the tails. Another problem is that in order to obtain statistically significant estimates of the expected return, long time-series are needed. This results in extreme weights that fluctuate over time, which causes the model to perform poorly out of sample. In order to improve the performance of the mean-variance model, substantial effort has been devoted to handling estimation errors. DeMiguel et al. (2007) compare 14 extended mean-variance models and find that out of sample, none of them could outperform the naïve strategy of equally weighting all assets. They conclude that in order for the mean-variance models to outperform the naïve strategy, either the number of assets must be severely limited, or the estimation window must be extremely long\(^2\).

In the fixed income market, movements in the term structure of interest rates drive the asset prices. A term structure relates time points to interest rates and may consist of infinitely many points with interest rates that vary over time. To be able to handle risk, the joint variation of interest rates is studied to extract a reduced set of systematic risk factors that can explain the vast majority of the term structure variance. Contrary to stocks, fixed income instruments mature, which causes exposure to the systematic risk factors to gradually change as the time of maturity approaches. These fundamental differences render sample estimates of expected returns and covariance unsuitable to use as direct inputs to the mean-variance model. However, the basic conceptual idea of weighing expected returns against

---

\(^2\)In a market with 50 assets, a 6000 months long estimation window would be needed for monthly data to achieve a higher certainty-equivalent return than the naïve strategy.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

risk remains. In order to model risk, the systematic risk factors must first be identified, and to obtain a quantifiable risk measure, the dynamics of each risk factor must be modeled. These topics are covered in Part 2 of the dissertation. The systematic risk factors are entitled to time-varying risk premiums, which must be measured in order to obtain expected returns. This is a delicate problem where long time series are required to achieve statistical significance. However, by solely focusing on risk management, the need for modeling expected returns can be heavily reduced by suitable choices of risk measures. The bond market further involves uncertainty whether or not the issuer will be able to make future payments. To properly solve investment problems on the bond market, credit risk must be modeled. Credit risk is the topic of Part 3 of this dissertation and includes measuring the term structures of default intensities and recovery rates, together with their systematic risk factors. To further measure and manage risks, the dynamics and joint variation of each risk factor must be modeled. To be able to solve investment problems, their associated risk premia must also be measured. All this shows that accurate mathematical modeling of fixed income markets is a very complex task.

There are several optimization approaches to handle decision-making under uncertainty. Stochastic optimal control is a continuous-time approach from the field of engineering where stochastic differential equations are used to model the dynamics of a system. Although several financial applications exist, there has been limited success in dealing with realistic problems with high-dimensional state variables, inequality constraints, and non-Gaussian randomness. Stochastic dynamic programming is an approach where the optimal actions are determined for some predetermined times by computing the value of all possible states. This can be carried out recursively by assuming that the optimal action is taken in every state. The lattice is thus traversed from the leaf nodes to the root, called backward recursion. The possibility of tackling large-scale problems in time and space is limited due to the curse of dimensionality. Approximate dynamic programming overcomes this curse by making an approximation of the value function in every state. Stochastic programming (SP) approaches uncertainty by discretizing the dynamics and making a scenario-representation, forming a deterministic equivalent problem that can be solved by conventional optimization methods. The stochastic programming approach poses no limitations to the randomness of the dynamics and can deal with different types of constraints.

3 The number of states grows exponentially with respect to a tiny increase in the number of dimensions or parameters due to combinatory explosion.
Large scale problems with many assets can be solved in polynomial time\(^4\). For multi-period problems, the number of scenarios used in each node is limited by the fact that the number of scenarios grows exponentially with the number of periods. Too few scenarios may lead to an inaccurate approximation of the intended problem. However, the quality of the approximation can be evaluated by comparing the optimistic bound given by the solution to the SP problem to the pessimistic bound obtained by evaluating a feasible solution under a different scenario tree. In order to avoid the curse of dimensionality, two-stage models (without recourse) are often used, and these models have been studied extensively. These are one-period models where the decision is made in the first stage, based on scenario outcomes in the second stage. Even though two-stage models cannot properly handle market frictions, such as transaction costs, the models can be tuned to produce practically useful solutions.

The main advantage of stochastic programming is that it allows for the use of tools from mathematical programming to handle constraints. In addition, the realistic modeling of the dynamics of financial problems can be utilized through Monte Carlo simulation. These properties are essential when trying to obtain practically useful solutions. As explained above, realistic modeling of the fixed income market is a complex task. This dissertation is written with stochastic programming in mind, using a bottom-up approach. This implies starting by carefully measuring the financial quantities that drive the prices in each market. Accurate measurements of financial quantities are crucial in the identification of the systematic risk factors using data-driven methods.

To be able to perform stochastic programming, the dynamics of the systematic risk factors must also be realistically modeled. By using the bottom-up approach described above, this dissertation provides other contributions besides stochastic programming models. This includes accurate measurement of the financial quantities on the credit market, measuring risks, and portfolio analysis. One result of using this rather unique bottom-up approach to developing the models in this dissertation is that a lot of knowledge of how the market behaves can be extracted in the process. This information can also indirectly be used for improved decision making. This leads to the purpose of this dissertation.

\(^4\)The time complexity for solving the deterministic equivalent is \(O(mn^3)\), where \(m\) is the number of nodes in the scenario tree and \(n\) is the number of assets (variables). The cube is the result of the inversion of a dense matrix in each subproblem.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

**Purpose.** To develop accurate models that can be used to improve decision-making in the fixed income market.

In order to guide the research towards fulfilling the purpose, five research questions are posed. These questions are in line with the purpose, and their answers help to verify the fulfillment of the purpose. The main research question for Part 2 reads:

**Research question 1.** How should a stochastic programming-based model be formulated to improve hedging decisions in the interest rate market?

In order to answer this question, two additional questions needed to be addressed. The first question concerns modeling the dynamics of the systematic risk factors. As this is closely related to measuring risk, which also provides a means for evaluating and comparing the results of different models, the second research question is formulated in terms of measuring risk.

**Research question 1a.** How can interest rate risk be accurately measured?

In order to evaluate the hedging results, an accurate model for monetary performance attribution for the fixed income market is needed. The third research question covers this topic.

**Research question 1b.** How should accurate monetary performance attribution be carried out in the fixed income market?

For Part 3 of this dissertation, which concerns credit risk, the main research question reads:

**Research question 2.** Which data-driven systematic risk factors can be identified in the credit market?

To answer this question, the related financial quantities that drive the bond prices first must be accurately measured. This leads to the last research question.

**Research question 2a.** How can financial quantities in the credit market be accurately measured?

Next, a description of the data used followed by the research methodology for answering these questions is presented.
In accordance with the strive for modeling in a realistic setting, all data used for the studies in the appended papers consist of actual market data retrieved from Thomson Reuters EIKON. The data used in Papers I-III are essentially the same, whereas the data used in Paper IV differ substantially. Generally, fixed income prices are indicative, meaning that they are not binding for the market maker providing them. The price might be higher or lower than what the provider indicates when contacting a trading desk to perform a trade. This implies that quoted market prices generally cannot be fully trusted since they might be subject to change, outdated, or simply misprinted. Two types of problems with noise in market prices have been identified. The first problem is constant prices, in terms of quoted yields or clean prices. The second problem is temporary spikes in prices lasting for one day or even longer time frames. These obvious price errors affect the measurements of financial quantities if they are not addressed somehow. Throughout the papers in this dissertation, this has been dealt with by cleaning the price time series from constant prices and spikes using a statistical approach.

The prices for the U.S. OTC-derivatives used in Papers I-III are retrieved through two separate time-series requests for bid and yields for all available dates, and range back to the early 90:s for some instruments. The quality of the prices on this highly liquid U.S. market is high, but a few spikes removed as described in Paper II and III.

For the bonds in Paper IV, the data retrieval process is much more complicated. Since price history for matured bonds cannot be retrieved from EIKON, the prices must be stored in advance. In addition, time series for credit ratings and bond status for each bond must be stored as these are subject to change. These requirements led to the construction of a SQL database where bond information is stored together with daily snapshots of bid-asks prices. The prices in this less liquid market are, to a greater extent, affected by noise. These prices are cleaned with a slightly more general statistical method for removing spikes and constant prices. For Paper IV, the Swedish bond market is used for the study. This, since much data from the U.S. bond market turned out to be missing in the database. In addition, the U.S. data available contained more noise, which is problematic since the inverse problem used to measure the financial quantities is highly sensitive.

---

5Prices for matured bonds can be retrieved via Thomson Reuters Datastream, but information of the bonds and their issuers is lacking there.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

to noise. The Swedish bond market is large enough to perform accurate measurements and produces substantially lower price errors. Besides, it is beneficial to have a basic knowledge of the issuers that helps to make better assessments of the results and the errors. The results should be generalizable to all markets with a similar bankruptcy process, such as the U.S. market.

1.2 Methodology

Operations research is the primary method used to conduct the research in this dissertation. According to INFORMS (2019), operations research is a discipline that deals with the application of advanced analytical methods to help make better decisions. This involves employing techniques from mathematical sciences, such as mathematical modeling, statistical analysis, and optimization. When dealing with complex research questions such as those stated in the previous section, multiple modeling choices have to be made. Each broader choice usually leads to several more detailed choices, such as parameter settings. This may lead to an explosion of choices to be made when modeling complex problems, where some of the models are better to model reality than others. The rest of this section is dedicated to motivating some of the specific modeling choices. More detailed choices are motivated in Part 2 and Part 3 of the dissertation. To facilitate comprehension of how the dissertation is structured, and the connections to research questions and papers, a roadmap for the dissertation is presented in Figure 1.

The generalized framework for measuring term structures of interest rates, presented by Blomvall (2017), was initially developed much earlier and formed the basis of the dissertation of Ndengo Rugengamanzi (2013). Blomvall and Ndengo (2013a) show that the Blomvall (2017) framework outperforms all traditional methods for measuring term structures in terms of pricing errors versus variance. They also provide accurate measurements of the systematic risk factors through principal component analysis (PCA). The appearance of the risk factors is also verified through Kalman-filtering observed market prices and bootstrapping the term structure in a model-free setting. Blomvall and Ndengo (2013b) extend the framework to multiple yield curves. The above-mentioned research formed a solid foundation for this dissertation and is therefore described in Part 2. The method for measuring single and multiple term structures of interest rate is presented in Chapter 4, and how to measure systematic risk factors of interest rates is presented in Chapter 5.
Figure 1: Roadmap for the dissertation describing how the chapters of the theoretical parts of the dissertation are connected and their connections to the appended papers and the previous research from the group in Linköping.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Once the framework for measuring term structures is in place, this enables us to study ex-ante and ex-post realizations of the systematic risk factors. Studying ex-post realizations implies decomposing realized term structure movements into the different systematic risk factors, whereas studying their behavior ex-ante involves modeling their dynamics. Accurate ex-post decomposition of the term structure variance into each systematic risk factor enables accurate performance attribution if each risk factor can be related to price changes in a portfolio. This can be done by a Taylor approximation of the valuation function of each instrument making up the portfolio, and this is the topic of Chapter 6.

When modeling the dynamics of the systematic risk factors, one has to recognize the fact that the principal component time series are uncorrelated in sample. This indicates that it is of high importance to get an accurate model of the univariate distributions of each risk factor. For this purpose, different models of the distribution and its moments were tested and evaluated using objective statistical criteria (Bayesian information criterion). Once the univariate distributions are modeled, the joint variation can be modeled separately by the use of copulas. As the dynamics of the risk factors are modeled, they can be evaluated against historical realizations out of sample. To facilitate this evaluation and enable for construction of risk measures, we construct a suitable portfolio and simulate how its value would have evolved over time. How to model the dynamics of the systematic risk factors and how to measure risk is discussed in Chapter 7.

The choice of using stochastic programming for hedging interest rate risk was made by Blomvall. This decision was based on the successful results for hedging equity index options by Barkhagen and Blomvall (2016) and after trying another approach for hedging interest rate risk in the master thesis by Uhrberg and Sjöholm (2013). The choice can be justified by the fact that the stochastic programming problem is fast to solve given a set of scenarios and that stochastic programming poses no limitations to the distributions used to generate these scenarios. The use of minimizing variance is motivated by the fact that this risk measure is almost independent of the expected value for which the risk premia ideally should be modeled. This is not the case when using utility functions directly or other risk measures such as value at risk or expected shortfall. Hedging interest rate risk is the topic of Chapter 8.

When it comes to the credit-risk part of this dissertation, the method for measuring the financial quantities is an extension to the Blomvall (2017)
and Blomvall and Ndengo (2013b) method. This choice is motivated by the encouraging results on both the risk-free interest rate market (Ndengo Rugengamanzi, 2013) and the equity index options market (Barkhagen, 2015). Chapter 9 gives an introduction to credit risk, together with a legal and empirical approach to the topic. Researching this is important when determining how to model these properties mathematically. The different alternatives for modeling credit risk are presented in a historical context. The choice of using reduced-form models can be motivated by the fact that it provides a theoretically grounded model in which credit risk can be measured in terms of default probabilities and recovery rates. These two term structures can be measured via inverse problems with prices from bond and/or credit derivatives as input. Just as for interest rates, where forward rates proved to be the most local representation of information and thereby best suited for enforcing economical realism through regularization, hazard rates were found to be the corresponding financial quantity for measuring defaults. Chapter 10 provides a more in-depth mathematical background to how credit risk is modeled via the reduced-form approach. Chapter 11 presents how credit risk is measured through bond prices.

1.3 Outline

Chapter 2 provides a market background to give the reader some context and further motivation as to why this dissertation is important. The theory in this thesis is divided into Part 2 and Part 3, which can be read somewhat independently of each other. In these parts, my view of each topic is presented and related to the existing literature. Some of the chapters end by highlighting some of the results from the papers. Part 4 discusses the contribution of this dissertation. In Chapter 12, the research questions are answered, and the purpose is evaluated. Suggestions for future research are presented in Chapter 13, which concludes the dissertation.
For small investors, the global equity market is rather well-known and has become increasingly accessible since the introduction of online stockbrokers. The market capitalization of the publicly listed global equity market is close to USD 70 trillion\(^1\) and Figure 1 shows the distribution between the largest markets.

Despite the fact that the bond market is substantially larger than the publicly-traded equity market, it is relatively unknown to the common citizen. Bonds can be viewed as publicly traded securitized debt, and Figure 2 illustrates the global bond market size and growth since 1980. The countries with the largest bond markets are sorted based on the outstanding amount at the end of 2018 and the remaining 28 countries with data available are aggregated. It is evident that the U.S. bond market is by far the largest, accounting for over 45% of the global outstanding amount in 2018. Out of these USD 44 trillion worth of U.S. bonds, government-issued bonds constitute 42%, or USD 18.7 trillion. To be clear, this is the way the U.S. government and other countries borrow money. Financial corporations account for 41% of the outstanding amount, non-financial corporations for 16% and just over 1% of the bonds are issued by private banks. For non-financial corporations, the bond market is an alternative to borrowing from a bank.

Statistics for bondholders are harder to obtain, but Sveriges riksbank (2016) presents data for the Swedish bond market. Insurance companies are the largest investors, holding 40% of the outstanding amount. International

\(^1\)USD 1 trillion = 1,000 billion or 1,000,000,000,000 is roughly the market value of Microsoft, which is the largest company by market cap in 2019.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

![Global Equity Market Capitalization](chart)

**Figure 1:** Global equity market capitalization for listed companies in the largest markets (The World Bank, 2019) where 79 countries with available data has been aggregated into the remaining category.

Investors, mainly pension funds, hold 23% of the issued amount, whereas national pension funds hold 12%. Banks hold 14%, and the remaining 11% are held by companies and other investors. In general, bonds are attractive investments for pension funds and insurance companies with long-term liabilities, since bonds enable such institutions to match the duration of their liabilities, thereby reducing the risk. Central banks are other major bondholders, especially after the global financial crisis in 2008. Federal Reserve (2019) holds USD 3.8 trillion of debt on their balance sheet at the present date, July 2019. This is a decrease from the steady level of 4.5 trillion during 2015-2018, but an increase from the USD 0.9 trillion held before the global financial crisis. The reason for this is the quantitative easings, where central banks simply create money to buy government debt and sometimes other securities. This increased demand increases prices, which puts downward pressure on interest rates.

A much less well-known market is the interbank market. This is an over-the-counter (OTC) market for interest rate derivatives, meaning that the
derivatives are not traded on an exchange. According to Bank for International Settlements (2019a), this market is largely made up of interest rate swaps, which are contracts between two counterparts to exchange sequences of cash flows with each other over a specified period of time. The simplest and most common swap is the vanilla swap, which plays a central role in this dissertation. Fixed cash flows are exchanged for floating rates based on some underlying reference rate, generally LIBOR. The rates of the fixed leg are usually paid annually, while the frequency of the floating rate payments is determined by the maturity of the underlying rate. LIBOR rates exist for different currencies and maturities, but three-month is the most common underlying floating rate for swaps. The floating-rate cash flows are thus paid quarterly over subsequent three-months periods. The floating rates are fixed at the reset days, usually two business days before the start of each period and paid at the end of each period (Fabozzi and Mann, 2011). This means that the most imminent floating cash flow is known, while the future floating rates are stochastic. No payments are exchanged when entering an interest rate swap. Instead, they are valued at zero NPV by determining the
rate of the fixed leg accordingly. When entering an interest rate swap, the investor determines the notional amount upon which the interest of both legs is calculated. This amount can be matched with interest payments of another payment stream to swap fixed to floating rate or vice versa. Swaps can also be used to speculate on the movement of future rates.

Forward rate agreements (FRA) are the second most common interest rate derivative, according to Bank for International Settlements (2019a). This is another type of OTC-contract linked to LIBOR or any other underlying reference rate. The counterparties agree to exchange a single payment of a fixed rate for a future floating rate. The investor chooses the notional amount, and the fixed rate is determined so that the initial NPV is zero.

In the upper panel of Figure 3, the notional amounts of the global interest rate derivatives are shown. Notably, the total nominal amount of the market exceeds that of the bond market by several times. Contrary to the nominal value of bonds, the notional amount of interest rate derivatives never changes hands, which hinders direct comparisons. The lower panel of Figure 3 instead displays the gross market value of the outstanding derivatives, which measures the combined market value of all OTC-positions at the side with positive NPV. This value is much lower than the notional amount. The red and green areas include interest rate options such as caps/floors (upper/lower limits to interest rates) or swaptions (options where the underlying instruments are interest rate swaps).

Despite being substantially smaller than the interest rate derivatives market, the credit derivatives market has received more attention due to the global financial crisis. It is generally accepted that credit standards in U.S. mortgage lending were relaxed in early 2000, causing a housing bubble (Jickling, 2009). Securitization of mortgage loans and distribution via collateralized debt obligations (CDO), caused lenders to relax their requirements, especially due to the high demand of subprime loans packaged as AAA bonds, endorsed by the quality mark of all three main rating institutions². Unregulated credit derivatives, originally developed to manage financial risk more efficiently, can be blamed to have accelerated this process. Even though the mathematical field of credit risk to a large extent was fully developed, the derivatives were complex, and their underlying assumptions were not well understood. The fast-paced financial innovation created new complex financial products, and short-term bonus incentives encouraged high-risk strategies. As explained by Jickling (2009), there are many complex inter-

²Standard & Poor, Moody’s and Fitch account for roughly 95% of the market.
Figure 3: Global market size for the interbank market. The upper panel displays the notional values of the outstanding derivatives, upon which interest rates are paid and the lower panel displays the gross market value for those derivatives (Bank for International Settlements, 2019a).

connected factors behind the global financial crisis, and credit derivatives and the lack of understanding of credit risk modeling are certainly two of them. Even though this dissertation does not directly address credit derivatives, they are relevant for measuring and managing credit risk.
According to Augustin et al. (2014), the first credit derivative was traded in 1994, but the market did not come to grow until the ’00s, especially the second part of the decade. The upper panel of Figure 4 displays the notional values for the credit derivatives market, and we see that prior to the global financial crisis in 2007, the notional amount of the credit default swap market was larger than the entire bond market. A credit default swap (CDS) can be seen as an insurance against default. A fixed credit spread, acting as an insurance fee, is paid on a quarterly basis up until the default or maturity of the contract. In case of default, the face value of the bond is paid out in exchange for the delivery of the defaulted bond, or the corresponding difference may be settled in cash. Entering this type of insurance is not limited to bondholders. Any investor can either buy, or even sell, such an insurance to speculate in credit risk, and this enabled the notional value of the CDS market to exceed the value entire bond market. Since an investor could net a long position against a short position and thereby neutralize the exposure while earning the difference between the credit spreads, the total notional amount gives a rather vague impression of the outstanding risks. In the lower panel of Figure 4, we see that the total gross market value of the credit derivatives market peaks about a year later than the notional amounts. The total gross value exceeds USD 10 trillion in 2009, unveiling the substantial risks residing during the unraveling of the financial crisis. The notional amounts and especially the gross market values have decreased substantially since 2009.

Following the global financial crisis, the credit derivatives market has undergone severe regulation in accordance with the Dodd-Frank and Basel III act. A significant structural change according to Augustin et al. (2014), is the introduction of central counterparties, which accounted for 55% of the outstanding notional values at the end of 2018, as can be seen in Figure 5.

To summarize the current status of the fixed income market, the bond market continues to grow. For the OTC derivatives markets, however, the gross market values are shrinking. According to Bank for International Settlements (2019c), this can be explained by the structural changes in the OTC market in the form of new practices, central clearing, and increased trade compression, the elimination of economically redundant derivatives positions. Considering notional amounts, the interest rate derivative market continues to remain a large market while the credit derivatives market has shrunk considerably since the peak in 2008.
Figure 4: Global market size for credit derivatives. The upper panel displays the notional values of the outstanding derivatives and the lower panel displays the gross market value for those derivatives (Bank for International Settlements, 2019b).
Credit Default Swaps - Notional Amounts

Figure 5: Outstanding notional value for different counterparties on the CDS (Bank for International Settlements, 2019b).
Interest Rate Risk
Interest Rates

Interest is paid as compensation for delaying consumption. Usually, interest is expressed in terms of annual interest rates, which denote the percentage of the principal amount borrowed or lent.

There are several ways to express interest rates. First, the compounding must be specified. Under a simple rate, $r$, an investment pays the total amount $A = N(1 + rt)$ over the time $t$. For a compound interest, $r$, with compounding frequency, $m$, an investment pays $A = N (1 + \frac{r}{m})^{mt}$. When the compounding frequency tends to infinity, we obtain continuous compounding

$$\lim_{m \to \infty} N \left(1 + \frac{r}{m}\right)^{mt} = Ne^{rt}. \quad (3.1)$$

This way of expressing interest rates facilitates discounting i.e. computing the net present value (NPV) of the amount, and will be used throughout this thesis, unless otherwise stated. The time $t$ in these expressions has to be measured with respect to a day-count convention. A day-count convention determines the number of days between two dates and the total number of days in a year. This way, the time, $t$, can be computed as year fraction. There are numerous day-count conventions applicable to different markets, but we will not go further into details.

We start by describing the term yield to maturity, which denotes the internal rate of return on a bond if it is held to maturity. Yield to maturity is often used to quote prices for bonds and other interest-rate instruments because yields are easier to grasp than prices, which are functions of both yield and time. For a zero-coupon bond paying its face value, $N$, at maturity, $T$,
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

the (continuously compounded) yield-to-maturity, $y$, is the rate required to discount the face value to obtain the price, $P(t, T)$, at time $t$.

$$ P(t, T) = Ne^{-y(T-t)}. \quad (3.2) $$

From this it follows that

$$ y = -\frac{\ln P(t, T) - \ln N}{T-t}. \quad (3.3) $$

It is commonly assumed that $N = 1$, implying that the price of a zero-coupon bond, $P(t, T)$, can be used as a discount factor when to discount any cash flow from time $T$ to $t$. Since there is only one cash flow occurring for a zero-coupon bond, the yield to maturity is equivalent to the (continuously compounded) zero coupon rate, or spot rate which is usually denoted by $r$. The spot rate is the annual rate obtained by investing in a zero-coupon bond with maturity $T$. Hence, spot rates obtained at a time, $t$, usually differs between maturities $T$ as $r(t, T)$. Information of such a financial quantity as a function of time is called a term structure, or simply a curve. The spot rate is related to the discount factor as

$$ P(t, T) = e^{-r(t,T)(T-t)} \Leftrightarrow r(t, T) = -\frac{\ln P(t, T)}{T-t}. \quad (3.4) $$

By letting $T$ approach $t$, we obtain the short rate, $r(t) = \lim_{T \downarrow t} r(r, T)$. This represent the rate an investor would earn over an infinitesimally short period of time following $t$, but is important for modeling purposes.

For a straight coupon-bearing bond that, besides paying the face value at maturity, also pays an annual coupon rate of, $C$, with payment frequency, $m$, resulting in a total of $n$ coupons at times $t_1, \ldots, t_n$, the (continuously compounded) yield to maturity must be solved from

$$ P(0, T) = N \left( \sum_{i=1}^{n} \frac{C}{m} e^{-yT_i} + e^{-yT_n} \right), \quad (3.5) $$

where we set $t = 0$ to avoid problems with numbering of the coupons. To price a coupon-bearing bond, we discount all its cash flows by the corresponding spot rates from the term structures according to

$$ P(0, T) = N \left( \sum_{i=1}^{n} \frac{C}{m} e^{-r(T_i)T_i} + e^{-r(T_n)T_n} \right). \quad (3.6) $$

The slight but important difference between (3.5) and (3.6) is that the spot rates may differ between cash-flow times whereas the yield to maturity,
which is the internal rate of return for holding that bond until maturity, is unique for each bond.

Consider two risk-free zero-coupon bonds with maturities $T_1$ and $T_2$, where $t < T_1 < T_2$. From their prices, we can compute the forward rate between $T_1$ and $T_2$ from

$$P(t, T_1) = P(t, T_2)e^{f(t, T_1, T_2)(T_2 - T_1)}, \quad (3.7)$$

by solving for $f(t, T_1, T_2)$

$$f(t, T_1, T_2) = \frac{1}{T_2 - T_1} \ln \left( \frac{P(t, T_1)}{P(t, T_2)} \right). \quad (3.8)$$

This forward rate is the rate that is obtained between $T_1$ and $T_2$ if the investment is “locked in” at time $t$. Using the spot rate definition from (3.4), the relation between spot rates and forward rates can be written as

$$f(t, T_1, T_2) = \frac{r(t, T_2)(T_2 - t) - r(t, T_1)(T_1 - t)}{T_2 - T_1} \quad (3.9)$$

The instantaneous forward rate, $f(t, T)$, is obtained as the limit when $T_2$ approaches $T_1$

$$f(t, T) = \lim_{T_2 \searrow T} f(t, T, T_2) = r(t, T) + T \frac{d}{dT} r(t, T) = \frac{d}{dT} \ln P(t, T). \quad (3.10)$$

When using the notation $f(t)$, we are referring to the instantaneous forward rate obtained over an infinitesimally short time in following the future time $t$, agreed upon today. Integrating this last relationship in (3.10) between $t$ and $T$, gives us

$$\int_t^T f(t, u)du = -[\ln P(t, T) - \ln P(t, t)] = -\ln P(t, T). \quad (3.11)$$

Hence, for $0 \leq t \leq T$ it holds that the relationship between the discount factor and forward rate is

$$P(t, T) = \exp \left( -\int_t^T f(t, u)du \right), \quad (3.12)$$

and from (3.4) we get

$$r(t, T) = \frac{1}{T - t} \int_t^T f(t, u)du. \quad (3.13)$$
We have now defined the relationships between the three types of term structures: discount factors, spot rates, and forward rates. These are merely different ways of expressing the same information. We note that the spot rate with maturity $T$ is an average of the forward rates between $t$ and $T$. We, therefore, claim that forward rates are a more local representation of information. The discount factor is, in turn, a non-linear transformation of the spot rate via the exponential function.

Since there are not enough zero-coupon bonds traded in the market, the term structures are not directly observable. Hence, the term structure of interest rates must be measured from information available in the market. Assume that both zero-coupon bonds and straight coupon-bond prices are available, where coupon-bearing bonds usually are more common. Having access to the prices of the instruments, equations (3.4) and (3.6) can be utilized to measure the term structure. Since prices are available but the term structure is not, we thus have to solve an inverse problem. Measuring term structures is the topic of the next chapter.
Measuring Term Structures of Interest Rates

Making accurate measurements of the term structure from observations of market prices is a challenging inverse problem. Noise in market prices is a complicating factor for OTC markets where prices are indicative and may be changed when attempting to make a trade with a market maker. When considering a continuous term structure, there exists an infinite number of interest rates but only a finite set of interest rate instruments. Hence, there exists an infinite number of solutions to the inverse problem. As pointed out by Blomvall (2017), an important aspect of a well-posed inverse problem is that small changes in input data should lead to small changes in output data.

In the field of inverse problems, regularization is commonly used to ensure this property, especially for the inverse problems encountered when training models in machine learning. However, few methods found in the financial literature consider regularization for this type of problem. There are four different schools in the literature of measuring term structures: equilibrium models, bootstrapping combined with interpolation, parsimonious functions, and regularization.

The first school is to model the short rate, \( r(t) \), according to some stochastic process. These models can be further divided into equilibrium and no-arbitrage models. Equilibrium models usually start with assumptions about economic variables and derive a process for the short rate, \( r \) (Hull, 2018). This way, the dynamics of the term structure are provided as well, which is useful for studying the implications of \( r \) for bond prices but especially for option prices, where closed-form solutions often exist. The disadvantage
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

of equilibrium models is that even if the parameters are judiciously chosen, they can never give an accurate measurement of the term structure. No-arbitrage models, on the other hand, use a measured term structure as input and only specify the dynamics.

The second school uses parsimonious functions to describe the term structure. Nelson and Siegel (1987) use a function of four parameters to describe the term structure of forward rates as

\[ f(t) = \beta_0 + \beta_1 e^{-\frac{t}{\tau}} + \beta_2 \frac{t}{\tau} e^{-\frac{t}{\tau}} : \beta_0, \beta_1, \beta_2, \tau \in \mathbb{R}_+. \]  

(4.1)

Svensson (1994) extends this model with two additional parameters to allow for a second “hump” in the term structure,

\[ f(t) = \beta_0 + \beta_1 e^{-\frac{t}{\tau_1}} + \beta_2 \frac{t}{\tau_1} e^{-\frac{t}{\tau_1}} + \beta_3 \frac{t}{\tau_2} e^{-\frac{t}{\tau_2}} : \beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2 \in \mathbb{R}_+. \]  

(4.2)

The parameters in these parsimonious models are fitted to observed market prices by minimizing the sum of the quadratic errors in the observed market prices. This makes the term structures robust to noise, but the parameterization poses limitations to the form of the term structure which results in price errors. Blomvall (2017) points out that the optimization problem is non-convex. This may cause the solution to oscillate between different local optima by slightly changing the input prices. As stated above, this is an unwanted property for inverse problems and may result in large costs when hedging a portfolio.

The third school is interpolation between known spot rates. These could be yields from zero-coupon bonds, but more often from interest rates bootstrapped from forward rate agreements and interest rate swaps. Bootstrapping involves using known rates from earlier maturities to eliminate all but one unknown rate for each instrument, thereby building the term structure from the short end. Since all but one discount factor must be known, some instruments cannot be included in the bootstrapping process. The three-month LIBOR rate can be used as a starting point, followed by \( 3 \times 6, 6 \times 9, 9 \times 12 \) month forward rate agreements and continuing, with annually spaced swaps, \( T = 2, \ldots \). Simple interpolation techniques or cubic splines can be used to interpolate between these known points (Hagan and West, 2006). Simple interpolation techniques include piecewise constant forward rates, linear or log-linear interpolation of spot rates, or discount factors. More recent examples include tension splines (Andersen, 2007) and kriging splines (Cousin et al., 2016).
4.1 Measuring Realistic Term Structures

Although commonly used in the market, the techniques described in the previous section have several problems. Interpolation techniques are exact methods, implying that all included instruments must be repriced exactly. This makes such methods prone to over-fitting and thereby sensitive to noise, which manifests itself as oscillating forward rates. Due to the differentiating property described in (3.10), forward rates become even more sensitive as maturity increases. Oscillating forward rates are not economically realistic because forward rates are the expectation of the future interest rate under the $T$-forward measure, technically $f_T = E^T[r(T) | F_t]$ (Blomvall, 2017). This implies that rapid changes in forward rates should stem from current information about the future, $F_t$. However, the market rarely has reliable information about the economy several years into the future. Therefore, a smooth forward rate curve is economically realistic. Inexact methods mitigate this problem by allowing for price deviation while penalizing the deviations, usually though a least-square method. However, the parsimonious functions proposed by Nelson and Siegel (1987) and Svensson (1994) are not flexible enough to perform accurate measurements.

Blomvall (2017) presents a framework for measuring term structures that generalizes the above-mentioned methods. The method is based on a trade-off between price errors and smooth curves, which is accomplished by a least-square method and regularization. The problem is convex when formulated in terms of discount factors. Since this framework is central to the dissertation, it will be described in some detail.

We allow the observed prices to either be bounded by the bid and ask quotes or equal to unique prices. This corresponds to inequality or equality constraints in the optimization model. Let $\mathcal{E}$ denote the set of instruments bounded by unique prices and equality constraints and let $\mathcal{B}$ denote the set of instruments bounded by bid and ask quotes. We present the framework in terms of forward rates since they contain the most local information and therefore is most suitable for regularization. Unfortunately, as Blomvall (2017) proves, the resulting optimization is non-convex. We hence use different valuation functions in terms of forward rates, $g(f)$, to price different types of instruments. For instruments in $\mathcal{E}$ with unique prices, $P_i$, we write

$$P_i = g_i(f), \quad i \in \mathcal{E},$$  \hspace{1cm} (4.3)
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

or equivalently stated on vector form,

\[ \mathbf{P} = \mathbf{g}_e(f), \]  

(4.4)

where \( \mathbf{P} = \{P_i\}_{i \in \mathcal{E}} \) and \( \mathbf{g}_e(f) = \{g_i(f)\}_{i \in \mathcal{E}} \).

For instruments bounded by bid and ask prices, these act as lower and upper bounds for the theoretical prices

\[ P^l_i \leq g_i(f) \leq P^u_i, \quad i \in \mathcal{B}, \]  

(4.5)

or equivalently stated

\[ \mathbf{P}_l \leq \mathbf{g}_b(f) \leq \mathbf{P}_u, \]  

(4.6)

where \( \mathbf{P}_l^l = \{P^l_i\}_{i \in \mathcal{B}}, \mathbf{P}_u^l = \{P^u_i\}_{i \in \mathcal{B}} \) and \( \mathbf{g}_b(f) = \{g_i(f)\}_{i \in \mathcal{B}} \). We allow for the theoretical prices \( \mathbf{g}(f) \) to deviate from the observed market quotes by introducing \( \mathbf{z}_e \), the deviations from unique prices, and \( \mathbf{z}_b \), the deviations from bounded prices. Let the matrices \( \mathbf{E}_e \) and \( \mathbf{E}_b \) contain the penalties for deviating from the observed unique and bounded market prices. Let \( \mathbf{F}_e \) and \( \mathbf{F}_b \) be scaling matrices for the market price deviations. When \( \mathbf{E} \) and \( \mathbf{F} \) are diagonal matrices, the penalties and scaling for each instrument is determined by its corresponding diagonal elements. The optimization problem is formulated as

\[
\begin{align*}
\min_{f, \mathbf{z}_e, \mathbf{z}_b} & \quad \frac{1}{2} \mathbf{z}_e^T \mathbf{E}_e \mathbf{z}_e + \frac{1}{2} \mathbf{z}_b^T \mathbf{E}_b \mathbf{z}_b + h(f) \\
\text{s.t.} & \quad \mathbf{P}_l - \mathbf{g}_b(f) + \mathbf{F}_b \mathbf{z}_b \leq \mathbf{P}_u \quad (4.7a) \\
& \quad \mathbf{g}_e(f) + \mathbf{F}_e \mathbf{z}_e = \mathbf{P}_e \quad (4.7b) \\
& \quad f \in \mathcal{F}. \quad (4.7c)
\end{align*}
\]

Both squared marked-price deviations and term structure roughness are penalized in the objective function, where \( h(f) \) represents the roughness function. The constraints (4.7a) and (4.7b) ensures that the theoretical prices plus the scaled price errors stays within the bid-ask spread for instruments \( i \in \mathcal{B} \), or becomes equal to the unique observed prices for instruments \( i \in \mathcal{E} \). The last constraint, (4.7c), enables further restrictions on the term structure and can be used to formulate the traditional methods within this framework. The shape of the term structure should however not be handled by such constraints but rather in the regularization function \( h(f) \).

The roughness measure is implemented by penalizing the term structure of forward rates, \( f(t) \), and its derivatives, \( f^{(k)}(t) \), up to order \( K \), using the
weighting functions $w_k(t)$. As stated by Blomvall (2017), term structures naturally exhibit large second-order derivatives. In the short end of the curve, this is due to expected changes of short rates and in the long end of the curve, the convexity adjustment causes the hump in the term structure. Therefore, the commonly used regularization obtained by penalizing the second-order derivative, used by Adams and Van Deventer (1994) among others, is probably not ideal. As a generalization, one can instead compute the roughness of the difference between the forward rates $f(t)$ and a reference curve $\bar{f}(t)$ and its derivatives, where $\bar{f}(t) = 0$ is the special case of penalizing the curve and its derivatives directly. Since term structures naturally exhibit level, slope, and curvature, we ideally want to penalize unrealistic levels of these naturally occurring stylized facts. The continuous version of the roughness function is

$$h(f) = \int_{T_0}^{T_n} w^0(t) \left(f(t) - \bar{f}(t)\right)^2 dt + \sum_{k=1}^{K} \int_{T_0}^{T_n} w^k(t) \left(f^{(k)}(t) - \bar{f}^{(k)}(t)\right)^2 dt.$$  

(4.8)

Measuring continuous term structures with this regularization function is an infinite-dimensional optimization problem. To be able to solve it, we have to discretize the term structure and use a discrete version of the roughness function. Discretizing the term structure does not necessarily result in any loss of information since one day is the shortest duration over which interest accrues. By approximating the derivatives by the finite difference method, we can obtain a discretized roughness function. When considering the special case when only the first and second derivative is used, we obtain the roughness function for $n$ discrete forward rate points as

$$h(f) = \sum_{t=0}^{n-2} w^1_t \left(\frac{f_{t+1} - f_t}{\Delta_t}\right)^2 \Delta_t + \sum_{t=1}^{n-2} w^2_t \left(\frac{2}{\Delta_{t-1} + \Delta_t} \left(\frac{f_{t+1} - f_t}{\Delta_t} - \frac{f_t - f_{t-1}}{\Delta_{t-1}}\right)\right)^2 \Delta_{t-1} + \Delta_t.$$  

(4.9)

Using daily discretization implies a high dimension of the optimization model (4.7). Using the standard optimization software IPOPT, the solution times may be as low as 15 seconds per term structure (Blomvall, 2017). Specialized solvers can be developed by utilizing the structure in the optimization problem as was done by Manzano and Blomvall (2004). Thereby, solution times of 50 ms can be reached.

Blomvall and Ndengo (2013a) show that the framework proposed by Blomvall (2017) dominates, or is close to dominating, all traditional methods.
when comparing out-of-sample pricing errors under first-order stochastic dominance. They also show that for the same level of average price error norms, the Blomvall (2017) method provides lower risk for both the U.S. (4513 days) and the Swedish (3414 days) swap market. In this study, $F$ and $E$ are set to the identity matrix, where the penalty matrix $E$ is scaled by a parameter to obtain an efficient frontier. The roughness function consists of penalizing the second-order derivative according to the weighting scheme

$$w_t^2 = \begin{cases} e^{\left(\frac{t}{365}-2\right)\ln 4}, & \text{if } t \leq 730 \text{ days} \\ 1, & \text{otherwise} \end{cases} \tag{4.10}$$

which Blomvall and Ndengo (2013a) refer to as the LSExp parameterization. The rationale behind this scheme is that information in the market is exponentially decaying and this is taken into account by increasing the weights exponentially in the roughness function up until a certain maturity, in this case 2 years. This parameterization has been used for measuring all term structures in the first three paper of this dissertation.

### 4.2 Measuring Multiple Term Structures

Before the global financial crisis in 2008, swap rates with the same maturity but different, tenors, i.e. lengths for the underlying floating rates, would be quoted at a non-zero but negligible spread. These spreads drastically widened during the credit crunch 2007 (Mercurio, 2009). Remarkably, this divergence did not create arbitrage opportunities when credit or liquidity issues are taken into account. As stated by Mercurio (2009), this suggests that consistent construction of a yield curve is possible only thanks to credit and liquidity theories justifying the simultaneous existence of different values for same-tenor market rates. After the credit crunch, the spreads tightened again but still remain significant as of today. To properly account for these spreads, multiple term structures should be used when discounting cash flows based on underlying rates with different tenors. Blomvall and Ndengo (2013b) extend the framework (4.7) into measuring multiple term structures for different tenors. This model assumes that there exists a single risk-free term structure of forward rates, $f_0$, which is based on overnight indexed swaps. Let $\tau \in \mathcal{T}$, be the set of different tenors longer than the shortest risk-free tenor, possibly overnight. The risk-free term structure and the
Measuring Multiple Term Structures

spreads of the different tenors, $\Delta f_\tau$, are measured simultaneously using

$$\min_{f_0, z_e, z_b, \Delta f_\tau} \frac{1}{2} z_e^T E_e z_e + \frac{1}{2} z_b^T E_b z_b + h(f_0) + \sum_{\tau \in T} h_\tau(\pi_\tau)$$

subject to

$$P_l \leq g_b(f) + F_b z_b \leq P_u$$

$$g_e(f) + F_e z_e = P_e$$

$$f_\tau = f_0 + \Delta f_\tau, \quad \forall \tau \in T$$

$$\Delta f_\tau \leq \Delta f_{\tau-1}, \quad \forall \tau \in T.$$

(4.11a) (4.11b) (4.11c) (4.11d) (4.11e)

Two new constraints are added. (4.11d) defines the relationship between the tenor spreads, $\Delta f_\tau$, and the risk-free forward rate curve, $f_0$. (4.11e) ensures that a curve with a longer tenor never crosses a curve with a shorter tenor. This framework forms the basis for the model used for measuring credit risk in the fourth paper of this dissertation. Here, the different curves are used to handle floating rate notes based on underlying rates of different tenors.

Accurate measurements of the term structures are key to accurately measure the systematic risk factors. This is the topic of the next chapter.
Measuring the Systematic Risk Factors of Interest Rates

Market risk results from movements in the systematic risk factors, and for the term structure of interest rates, this risk is called interest rate risk. Systematic risk is a risk that cannot be diversified away and thereby is entitled with a risk premium. If investors are risk neutral, the risk does not matter as long as the expected return is the same. Risk premiums exist because rational investors are risk-averse, meaning that their utility function is marginally decreasing. This implies that losing money decreases the utility more than the increase in utility obtained from gaining the same amount. Identifying systematic risk factors is by no means an easy task, but measuring the associated risk premium is even more complex. In the equity market, company risk is not systematic since it can be diversified away. In the interest-rate market, there are certain movements of the term structure that an investor cannot eliminate through diversification only. These systematic movements constitute risk factors of interest rate and are closely related to how the term structures are measured. In short-rate models, the risk factors are implicitly given by the model specification. For parsimonious models, the risk factors are given by the parameterization of the term structure. For other models, the risk factors must be measured by observing how term structures behave over time. Since each model suggests different risk factors, a central question is what the systematic risk factors really look like. There is no easy answer to this question, but measuring term structures and measuring interest rate risk are closely related subjects. It turns out that having accurate term structure measurements is of the utmost importance for being able to accurately measure the systematic risk factors, and by extension, the risk.
By measuring a long series of historical term structures, their behavior over time can be studied. Litterman and Scheinkman (1991) were the first who applied principal component analysis (PCA) to determine systematic risk factors for the term structure of interest rates. The analysis was applied to treasury spot rates, and the three most significant factors were found to explain almost all of the return variability across maturities. These factors proved especially useful for hedging and were named “Shift”, “Twist” and “Butterfly”. These factors account for changes in the level, slope, and curvature of the term structure.

PCA uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. This transformation is defined in such a way that the first loading points in the direction where data has the largest variance. By projecting the data onto the loading and subtract the resulting components, the process can be repeated to find additional loadings which thereby form an uncorrelated orthogonal basis set. PCA can be performed by an eigenvalue decomposition of the covariance matrix $C = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T$, where $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues and $\mathbf{E}$ is a matrix containing its corresponding eigenvectors. The eigenvalues correspond to the variance in each risk factor and the eigenvectors form the orthogonal base that constitutes the risk factors. Since the risk factors should reflect the movement in the term structures, daily innovations of term structures, $\Delta r_f^t = r_{f+1}^t - r_f^t$, should be used for computing $C$. Since the covariance matrix is needed to perform PCA, the term structures must be discretized into $n$ periods. PCA can be used as a compression technique where the most important features are kept, and the rest can be discarded to save space. The covariance matrix can thus be approximated by the $k$ eigenvectors, $\mathbf{E}_k \in \mathbb{R}^{n \times k}$, with the largest corresponding eigenvalues, $\mathbf{\Lambda}_k \in \mathbb{R}^{k \times k}$, as $C = \mathbf{E}_k\mathbf{\Lambda}_k\mathbf{E}_k^T$. This compression preserves

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i}$$

of the total variance, and by this analysis, we can compute the proportion of term structure variance explained by $k$ risk factors. If the movements in the term structure innovations are strongly correlated across maturities, a few risk factors are sufficient to explain almost all the term structure variance.

In the literature, PCA has been commonly used to analyze spot rates. Seeing that PCA maximizes variance and that noise tends to increase variance
by overfitting, PCA is sensitive to noise. Since spot rates are averages of forward rates, the noise present is also averaged and the results from the PCA may seem to improve. Laurini and Ohashi (2015) provide an excellent review of the problems caused by using forward rates when noise is present in market prices. They present a proposition that states that the number of risk factors should be the same regardless of whether spot- or forward rates are used. A difference in the number of risk factors needed to achieve the same explanatory power for forward rates as for spot rates in practice, is considered to be strong evidence of measurement noise in the market data.

The risk factor loadings for forward rate term structures measured with the Blomvall (2017) and the cubic spline method can be seen in Figure 1. We see that loadings for the Blomvall (2017) method are smooth, while the ones for cubic spline oscillates as a result of the term structures overfitting to measurements noise. For the Blomvall (2017) method, we see that the first four loadings all start at zero in the short end. As shown by Blomvall and Ndengo (2013a), the first three risk factors explain the majority of the variance in the long end of the term structure, while the following three risk factors explain the majority of the variance in the short end. The risk factors shown in Figure 1 are not just orthogonal factors that happens to explain the variance, they also have economic interpretation. As Blomvall (2012) show, a linear Shift is equivalent to changes in expected future short rates. A linear Twist is equivalent to changes in expected returns, and a Butterfly corresponds to uniform changes to volatility.

Blomvall and Hagenbjörk (2019) are able to show that performing the PCA on spot rates does not help when analyzing variance in terms of forward rates. Consider the integrating matrix $A \in \mathbb{R}^{M \times M}$, where

$$A_{i,j} = \begin{cases} 1/i, & \text{if } j \leq i \\ 0, & \text{otherwise} \end{cases} \quad (5.2)$$

which can be used to transform a vector of daily discretized forward rates, $\mathbf{r}^f$, into spot rates by $\mathbf{r}^s = A \mathbf{r}^f$. Furthermore, we can introduce the differentiating matrix $B$, which is the inverse of $A$, where

$$B_{i,j} = \begin{cases} 1 - i, & \text{if } j = i - 1 \\ i, & \text{if } j = i \\ 0, & \text{otherwise} \end{cases} \quad (5.3)$$

to transform a vector of daily discretized spot rates into forward rates, $\mathbf{r}^f = B \mathbf{r}^s$. 

39
We start by studying the in-sample volatility of the different term structure representations. Assuming that the PCA is performed on forward rates, the in-sample term structure variance is given by the diagonal elements of the covariance matrix \( C \). Using the approximation \( \tilde{C} \) gives us the forward rate volatility

\[
\sigma_f = \sqrt{\text{diag}(E_k \Lambda_k E_k^T)},
\]

where the square-root operates element wise. Using \( A \), the spot rate volatil-
ity can be computed as

$$\sigma_r = \sqrt{\text{diag} (A E_k \Lambda_k E_k^T A^T)}.$$  

(5.5)

Figure 2 shows the in-sample term structure volatilities for the Blomvall (2017) and cubic spline method for both forward and spot rates. The solid lines represent the term structure volatility from the covariance matrix, $C$, while the dashed and dotted lines represent the different approximations where six risk factors are used for the Blomvall (2017) method and sixteen risk factors for the cubic spline method, in order to explain the same proportion of the total variance. In the upper panel of Figure 2, spot rate volatility is illustrated, and we see that for the long end, the volatility does not differ between the methods. For the forward rate volatility displayed in the lower panel, we see that the cubic spline term structures contain more noise than the Blomvall (2017) term structures. We also see that it does not matter whether the PCA is carried out for spot rates or forward rates. The difference between the dashed and dotted lines is the result of spot rate risk factors failing to model forward rate variance properly.

We can also investigate how the different risk factors contribute to the total in-sample term structure variance when performing PCA on forward or spot rate innovations. The proportion of variance explained by different maturities can be computed as

$$\frac{\text{diag} (A E_k \Lambda_k E_k^T A^T)}{\text{diag} (C)},$$  

(5.6)

for increasing values of $k$. The proportion of the total variance explained by $k$ eigenvectors can be computed as

$$\frac{\text{Tr} (E_k \Lambda_k E_k^T)}{\text{Tr} (C)},$$  

(5.7)

where Tr denotes the trace of the matrix. Utilizing the equations (5.6) and (5.7) together with matrices $A$ and $B$, we can compute the explained proportions of variance and the proportion of the total variance explained according to the scheme in Table 5.1.

Performing this analysis for the Blomvall (2017) and cubic spline method gives us the results in Figure 3 and 4, where only six risk factors are used. Figure 3 shows the proportion of the total spot rate variance that can be explained by the first six risk factors when the PCA is performed on forward
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Volatility for Spot Rates

Volatility for Forward Rates

Figure 2: Term structure volatility for forward and spot rate measured by the Blomvall (2017) and cubic spline method. The dashed lines are indistinguishable from the solid lines since almost the entire covariance matrix can be explained in-sample using six and sixteen risk factors.

For the cubic spline method, the first six risk factors from spot-rate based PCA explain over 98% of the in-sample spot rate variance, while for the forward-rate based PCA, the risk factors only explain 82%. This might lead us to the false conclusion that measuring risk factors from spot rates is superior. Studying figure 4, we see that the risk factors from the spot rate based PCA only explain 43% of the forward rate variance compared to 82% for the forward-rate-based risk
Table 5.1: Scheme for computing explained proportions and the proportion of the total variance explained for term structures measured as forward or spot rates.

Factors. To explain forward rate variance, risk factors from spot rate based PCA are clearly inferior. For the Blomvall (2017) method, six forward rate risk factors explain almost the entire in-sample variance in both forward and spot rates.

Figure 3: The proportion of the total in-sample spot rate variance that can be explained by the Blomvall (2017) and cubic spline method when PCA is performed on forward rates (left panel) and spot rates (right panel).
Figure 4: The proportion of the total in-sample forward rate variance that can be explained by the Blomvall (2017) and cubic spline method when PCA is performed on forward rates (left panel) and spot rates (right panel).

As seen in Figure 2, 3 and 4, accurate measurements of term structures are essential to obtain accurate measurements of the systematic risk factors. As we show in Chapter 7, this also leads to lower risk measures. In order to measure risk, the dynamics of the risk factors must be modeled, and their distributions studied ex-ante. By observing the ex-post market realizations, performance attribution can be studied. This is the topic of the next chapter.
Performance Attribution

Performance attribution aims to decompose the portfolio return into different factors. Thereby it is possible to determine the origin of the returns and whether the returns stem from the intended investment strategy and decision-making process. For an investor, attributing the returns into components is a valuable tool for evaluating the investment process. Historically, it is used to explain why a portfolio’s performance differed from the benchmark. According to Blomvall and Hagenbjörk (2019), performance attribution and risk measurement can be viewed as dual aspects of accumulation of wealth, as both rely on an accurate description of the systematic risk factors. In risk measurement, which is studied in Chapter 7, the risk factors are seen as random variables for which dynamics are specified. Performance attribution, on the other hand, is carried out by studying ex-post realizations of these random variables.

The foundation of equity performance attribution was established by Brinson and Fachler (1985) and Brinson et al. (1995). They divide the excess portfolio returns into an asset allocation effect, a security selection effect, and an interaction term. Ankrim and Hensel (1994) extend this model to incorporate currency management effects. Even though fixed income prices are driven by changes in the term structure, fixed income attribution was originally not distinguished from equity attribution. Ankrim (1992) proposed a risk-adjusted performance attribution model, where beta is used as a proxy for risk. Kuberek (1995) used the risk factors obtained from PCA to build his framework. This idea was extended by the work of Colin (2006), Zambruno (2008) and Daul et al. (2016). The tradition of using returns for performance attribution complicates the problem. Arithmetic returns are
additive across assets but not over time. Geometric returns, on the other hand, are additive over time but not across assets. This creates problems when trying to link performance attribution results in a residual-free manner. For this purpose, several smoothing algorithms have been developed, see the work of (Carino, 1999; Frongello, 2002; Menchero, 2004, 2005).

Blomvall and Hagenbjörk (2019) present model-free a generic framework for monetary performance attribution, which mitigates the problem of linking returns. The framework is model-free and based on a second-order Taylor approximation of the analytical sensitivities to relevant risk factors. From the significant systematic risk factors loadings, \( E_k \), we can compute the movements in these risk factors as

\[
\Delta \xi = E_k^T \Delta r^f.
\]  

(6.1)

By differentiating the pricing expressions for each instrument, we can compute the first and second-order sensitivity with respect to the risk factors as

\[
g_{t-1,i} = \nabla \xi P_{t,i} \bigg|_{\Delta \xi_t = 0},
\]  

(6.2)

\[
H_{t-1,i} = \nabla^2 \xi P_{t,i} \bigg|_{\Delta \xi_t = 0}.
\]  

(6.3)

The difference in quoted price, \( \Delta \bar{P} \), for an instrument can be written as

\[
\Delta \bar{P}_{t,i} = \theta_{t-1,i} \Delta t + g_{t-1,i}^T \Delta \xi_t + \frac{1}{2} \Delta \xi^T t \Delta \xi_t + \varepsilon^I_{t,i} + \varepsilon^A_{t,i} + \Delta \varepsilon_{t,i}^P,
\]  

(6.4)

where \( \theta_{t-1,i} \Delta t \) is a carry term, \( \varepsilon^I_{t,i} \) is an error term caused by omitting insignificant risk factors, \( \varepsilon^A_{t,i} \) the error caused by the Taylor approximation, and \( \Delta \varepsilon_{t,i}^P \) is the difference between the theoretical valuation and the quoted price if available. By considering these three error terms, the performance attribution can be made completely residual free. By studying the size of the error terms \( \varepsilon^A_{t,i} \) and \( \varepsilon^I_{t,i} \), appropriate actions can be taken if they reach unsatisfactory high levels. The framework is also extended to incorporate foreign exchange rates and transaction costs.

Figure 1 shows the performance attribution for a portfolio of forward rate agreements in a foreign currency. The accurate measurement of forward rate risk factors allows us to explain almost the entire cumulative difference in portfolio value out of sample, using only six risk factors. As seen in the lower-left panel, the error terms are of negligible size, especially the
truncation error, $\varepsilon^I_t$, resulting from omitting all but the first six risk factors. When performing the same analysis for cubic splines, the truncation error instead accounts for the majority of the risk, as shown by Blomvall and Hagenbjörk (2019).

\[ \text{Accumulated Performance Attribution for all Factors from Blomvall (2017)} \]

\[ \begin{array}{c}
\text{Error terms} \\
\text{Risk factors}
\end{array} \]

Figure 1: Performance attribution for a portfolio of forward rate agreements in a foreign currency.
Measuring Interest Rate Risk

Accurate measurement of interest rate risk is tightly coupled to measuring term structures and systematic risk factors. Historically, interest risk has been about identifying risk factors and seeking ways to eliminate these. This view is related to hedging presented in the next chapter. A more modern view of risk involves specifying the dynamics of the risk factors in order to obtain an accurate representation of the distribution of the value a portfolio at a certain horizon. This can further be used to measure how much value is at risk, given that some extraordinary event occurs. This is in line with modern risk measures used in regulations specifying the capital requirements for such portfolios.

7.1 Historical Measures of Interest Rate Risk

Interest rate risk has been a topic of research for more than 80 years now since the introduction of bond duration by Frederick Macaulay (1938). According to Ingersoll et al. (1978), it was Hicks (1939) who independently introduced the duration concept as a proxy of interest rate risk, a risk measure that later would be rediscovered several times. Ingersoll et al. (1978) points out that the duration measure was originally defined in terms of discounting factors, but throughout history was tampered with to take the more well-known form based on yield to maturity. Consider a bond where the cash flows at time $t_i, i = 1, \ldots, n$ are of size $c_i$, where the last one, $c_n$, \ldots
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

also includes the redemption. The bond price at \( t = 0 \) is given by

\[
P = \sum_{i=1}^{n} c_i e^{-yt_i},
\] (7.1)

where \( y \) is the bond’s (continuously compounded) yield to maturity. The duration in terms of yield to maturity is defined as

\[
D = -\frac{1}{P} \frac{\partial P}{\partial y} = \sum_{i=1}^{n} c_i t_i e^{-yt_i}.
\] (7.2)

This definition implies that the price of the bond price drops by \( D\% \) when \( y \) increases by one unit, i.e. 100%. Another interpretation that is more closely related to its name is that the duration is the weighted average time of the cash flow payments. Viewing the cash flows, \( c_i \), as weights spread out on a mass-less teeter-totter, with distances according to their payment times, \( t_i \), the duration represents the equilibrium point.

Ingersoll et al. (1978) show that this duration measure cannot be used as a risk proxy unless the term structure is flat. Fisher and Weil (1971) provide a proof and empirical validation of how to construct an immunized bond portfolio by choosing the duration of the portfolio equal to the investment horizon. This proof was carried out through a constant parallel shift in the forward rate term structure. Cox et al. (1979) criticize the yield-to-maturity based duration measure since it is only valid under unrealistic circumstances that imply arbitrage opportunities. To resolve this, they introduced a stochastic duration measure based on what later became the famous Cox et al. (1985b) (CIR) equilibrium short-rate model.

After the introduction of the duration concept, it was later accompanied by convexity to improve bond immunization further.

\[
C = -\frac{1}{P} \frac{\partial^2 P}{\partial y^2} = -\sum_{i=1}^{n} c_i t_i^2 e^{-yt_i}.
\] (7.3)

The reason for this improvement is not related to convexity being the second derivative\(^1\) of the bond value under a parallel shift, as often described in the literature. Instead, it is the fact that convexity represents the sensitivity to a change in the slope of the term structure, as we show in the next paragraph.

\(^1\)Compare to gamma when delta hedging stocks movements using options. Neutralizing both delta and gamma will make a better approximation, much like introducing a quadratic term in a Taylor approximation.
Historical Measures of Interest Rate Risk

This was pointed out by Crack and Nawalkha (2000) and further stressed by Hodges and Parekh (2006). Despite lacking the proper understanding of the mechanics behind convexity, interest rate risk was thus already measured using two factors at an early stage. As pointed out by Hodges and Parekh (2006), the concept is still rather misunderstood.

To study convexity as the sensitivity to an increased twist of the term structure, we start by defining the concepts in terms of a general perturbation to the term structure of spot rates. Consider a bond priced at $t = 0$ and let $P(\xi(t))$ denote the price with respect to changes in an arbitrary risk factor with loadings $\xi(t)$,

$$ P(\xi(t)) = \sum_{i=1}^{n} c_i e^{-(r(t_i)+\xi(t_i))t_i}. \quad (7.4) $$

The duration and convexity measures in terms of spot rates can thus be defined as the price sensitivity with respect to a constant shift in the term structure, $\xi(t) = \delta$, i.e.

$$ D_r(\delta) = -\frac{1}{P} \frac{\partial P(\delta)}{\partial \delta} = \frac{1}{P} \sum_{i=1}^{n} c_i t_i e^{-(r(t_i)+\delta)t_i}, \quad (7.5) $$

$$ C_r(\delta) = \frac{1}{P} \frac{\partial^2 P(\delta)}{\partial \delta^2} = -\frac{1}{P} \sum_{i=1}^{n} c_i t_i^2 e^{-(r(t_i)+\delta)t_i}. \quad (7.6) $$

Evaluating the convexity at $\delta = 0$ implies

$$ C_r = \frac{\partial^2 P(\delta)}{\partial \delta^2} \bigg|_{\delta=0} = \sum_{i=1}^{n} c_i t_i^2 e^{-r(t_i)}. \quad (7.7) $$

Let us now investigate the price sensitivity with respect to a twist in the term structure by setting $\xi(t) = \delta t$. The pricing formula with respect to this risk factor is

$$ P(\xi(t)) = \sum_{i=1}^{n} c_i e^{-(r(t_i)+\delta t_i)t_i}. \quad (7.8) $$

The price sensitivity with respect to this risk factor hence becomes

$$ \frac{\partial P(\delta t)}{\partial \delta} = \sum_{i=1}^{n} c_i t_i^2 e^{-(r_i+\delta t_i)t_i}, \quad (7.9) $$

and evaluating this sensitivity for $\delta = 0$, gives us

$$ \frac{\partial P(\xi)}{\partial \delta} \bigg|_{\delta=0} = \sum_{i=1}^{n} c_i e^{-rt_i t_i^2} \quad (7.10) $$

51
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

We have thus proved that convexity is equivalent to a linear twist.

Fong and Vasicek (1984) consider a duration-matched portfolio and generalize the immunization concept to arbitrary non-parallel shocks to the forward rate curve. They introduce the $M^2$ measure, which is the weighted variance of time to payments around the horizon date $H$,

$$M^2 = \frac{1}{V_0} \sum_{i=1}^{n} (t_i - H)^2 c_i e^{-r_i t_i}, \quad (7.11)$$

where $V_0$ is the initial portfolio value. When the horizon, $H$, is equal to the duration, $D$, the $M^2$ model provides a lower bound for the change in terminal value of a portfolio. This bound is given by

$$\frac{V_H}{V_0} \geq -\frac{1}{2} K M^2, \quad (7.12)$$

where $K$ is a function of the forward rate changes such that

$$\frac{df(t)}{dt} \leq K, \quad \forall t \geq 0. \quad (7.13)$$

Nawalkha et al. (1990) showed that it exists a linear relationship between $M^2$ and convexity,

$$M^2 = C - 2DH + H^2. \quad (7.14)$$

As the $M^2$-model sought to minimize convexity, this created a paradox with the traditional view of convexity as a measure that contributes to bond returns. This effect can be seen if duration and convexity are used in a Taylor expansion to approximate bond return as

$$\frac{\Delta P}{P} \approx -D\Delta y + \frac{1}{2} C(\Delta y)^2. \quad (7.15)$$

The paradox was empirically investigated by Lacey and Nawalkha (1993) and was resolved via the modern view of convexity as a measure of sensitivity to slope risk mentioned above. Nawalkha and Chambers (1996) later derived a similar immunization model called $M$-absolute that reduces the risk inherent in the traditional duration model.

$$MA = \frac{1}{V_0} \sum_{i=1}^{n} |t_i - H| c_i e^{-r_i t_i}. \quad (7.16)$$

While $M^2$ is a two-factor risk model where both the $M^2$ measure and duration are considered, $M$-absolute is a single risk-factor model. $M$-absolute is
therefore easier to implement but cannot produce as tight a bound for the maximal change in portfolio value.

Upon the recognition of additional risk factors than shift and twist, multifactor models in terms of duration vector models, were introduced to extend the immunization concept. A common approach is to consider a parametric fitting of the term structure and derive the corresponding duration vector, Chambers et al. (1988) and Prisman and Shores (1988) consider polynomial functions to represent the term structure and construct duration measures. Willner (1996) and Diebold et al. (2006) use the exponential functions, proposed by Nelson and Siegel (1987), to define duration vector models. Nawalkha and Chambers (1997) generalize the $M^2$-model to a $Q$-factor duration vector model

$$M^m = \frac{1}{V_0} \sum_{i=1}^{n} (t-H)^m c_i e^{-r_i t_i}, \quad m = 1, \ldots, Q,$$  \hspace{1cm} (7.17)

where $m = 1$ reduces the model to the traditional duration model and $m = 2$ reduces it into the $M^2$ described earlier. They find that an $M^5$-model provides near-perfect hedging and eliminates 95% of the risk. Nawalkha et al. (2003) generalizes this model further by using a Taylor expansion of the bond returns with respect to functions of the cash flow maturities. The $m$:th order element thus becomes the weighted averages of $(g(t) - H(t))^m$. When tested for the polynomial functions, $g(t) = t^\alpha$, $0 < \alpha < 1$, the performance improves significantly for short planning horizons.

Key-rate duration was introduced by Ho (1992) and became a popular tool for handling non-parallel risk despite its limitations. Key-rate duration measures the portfolio sensitivity with respect to the change in several key rates for the term structure. The model has been criticized because the key rates are chosen arbitrarily, and the effects on forward rates become unrealistic.

Litterman and Scheinkman (1991) were the first who applied principal component analysis (PCA) to determine systematic risk factors for the term structure of interest rates. Knez et al. (1994) perform the same type of analysis to the short end of the term structure via the money market and propose a fourth factor named the private issuer factor. Bliss (1997) applies PCA on the Fama and Bliss (1987) term structures and concludes that since three factors are needed to explain a large proportion of bond price movements, single-factor models such as Cox et al. (1985b) and Vasicek (1977) are incomplete and will cause hedging errors. Carcano and Hakim (2011) compare hedging via PCA to the traditional key rate duration and duration
vector methods. They adjust the models for model errors and find PCA to outperform the traditional methods significantly.

For a more detailed review of the development of interest rate risk and hedging, see the book by Nawalkha and Soto (2009).

7.2 Accurate Measurements of Interest Rate Risk

Modern risk measures require not only the portfolio sensitivities to certain risk factors to be measured, but also the dynamics of these risk factors to be evaluated. Hagenbjörk and Blomvall (2019) establish the interconnection between the measurement of interest rate risk and measurement of term structures and identify three important properties. First, there should be no systematic price errors caused by the term structure measurement method. Second, the distribution of the risk factors must be consistent with their realizations. Third, a lower risk is preferable. There is a trade-off between these properties. An exact method, such as cubic splines, implies zero in-sample price errors. As seen in Chapter 5, overfitting to market price noise causes the risk (variance) in the term structures to increase. Modeling the risk factors consistently thus causes unnecessarily high risk. Using a parsimonious model such as Nelson and Siegel (1987) to model the term structure may create systematic price errors due to its limited shape. Modeling the risk factors given by the parameterization will most certainly lead to lower risk. An extreme example would be a constant term-structure, which implies zero risk while being consistent with its own realizations at the same time. This, of course, comes at the cost of extreme systematic price errors, which is not practically possible since mark-to-market valuation, i.e. valuing instruments consistent with market values, is required by both the IFRS and the U.S. GAAP accounting standards. Hagenbjörk and Blomvall (2019) show that the Blomvall (2017) model outperforms the cubic spline and dynamic Nelson and Siegel (1987) model in terms of consistency and risk levels. Blomvall and Ndengo (2013a) show that the Blomvall (2017) model dominates both of them in terms of out-of-sample price errors. Hence, using the Blomvall (2017) method for accurately measuring term structures results in more accurate risk factors, thereby enabling more accurate risk measurements.

Accurate measurement of interest risk revolves around accurate modeling of the dynamics of each risk factor and their co-dependence. In-sample, the
principal components for the risk factors, $\Delta \xi = \mathbf{E}_k \Delta \mathbf{r}^f$, are linearly uncorrelated by construction but out of sample, they are not. Considering the fact that the normal distribution is fully specified by the first two moments, $N(\mu, \sigma)$, then the risk factors would also be independent if they were normally distributed\(^2\). Hagenbjörk and Blomvall (2019) show that the principal component series of forward rates are not normally distributed. The fact that we must be able to use the risk factors out of sample, plus the non-Gaussian property, leads us to consider modeling their joint dependence. Copulas are suitable tools for modeling joint dependence since they allow for independent univariate modeling of each risk factor. The name “copula” stems from the Latin word for link, and a copula is a link between univariate marginal distribution to a multivariate distribution. According to Sklar’s theorem (Sklar, 1959), every multivariate cumulative distribution function can be expressed in terms of its marginal distributions and a copula.

### 7.2.1 Modeling Univariate Distributions for Risk Factors

Hagenbjörk and Blomvall (2019) show that the risk factors univariate distributions are non-normal, heteroskedastic, and skewed. To handle heteroskedasticity and skewness, the well-known GJR-GARCH($p$, $o$, $q$) model by Glosten et al. (1993) is used to model volatility according to

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i |\Delta x_{t-i}|^2 + \sum_{j=1}^{o} \gamma_i |\Delta x_{t-j}|^2 \mathbb{1}_{\{\Delta x_{t-j} < 0\}} + \sum_{k=1}^{q} \beta_i \sigma_{t-k}^2,$$

where $\Delta x_i$ is the risk factor. The simplest dynamic model for risk factors is to model a constant drift

$$\Delta \xi_t = \mu + \varepsilon,$$

where the residual variable, $\varepsilon$, is a random variable drawn from some distribution that is scaled by $\sigma_t$ obtained from (7.18). Autoregressive AR(1)-processes are popular for simulating term structures using the dynamic Nelson and Siegel (1987) model presented by Diebold and Li (2006). Using an AR($p$) process, the risk factor can be simulated by

$$\Delta \xi_t = \mu + \sum_{i=1}^{p} \Phi_{Li} \Delta \xi_{t-L_i} + \varepsilon,$$

\(^2\)It is sometimes wrongly assumed that PCA requires data to be normally distributed. However, if the data is normally distributed, so is the output, $\Delta \xi \sim N(\mathbf{0}, \Lambda)$, since every linear combination of the original features is normally distributed by definition.
where $L_i$ is the lag of order $i$. However, Hagenbjörk and Blomvall (2019) find no improvements by autoregressive processes.

In order to avoid overfitting and over-parameterization, Hagenbjörk and Blomvall (2019) use the Bayesian information criterion (BIC) to determine the order of the models, their parameters, and the distribution of the residual variable. We find that the well-known special case GARCH(1,1) with a Student’s $t$-distribution were preferred for all risk factors. The in-sample fit for the first three risk factors, shift, twist, and butterfly, can be seen in Figure 1. The out-of-sample fit for the first three risk factors, shift, twist, and butterfly, can be seen in Figure 2.

![Figure 1: In-sample histograms for the first three risk factors comparing GARCH(1,1)-models to constant-volatility models with both Gaussian and Student’s $t$-distributed residuals.](image)

7.2.2 Modeling Joint Variation using Copulas

Even though Sklar’s theorem is theoretically appealing, copulas do not eliminate the problem. Given the correct univariate marginal distributions, the
Accurate Measurements of Interest Rate Risk

Figure 2: Out-of-sample QQ-plot for the first three risk factors.

correct copula must still be found. While this has been an area of research, a few standard copulas have been used extensively in finance.\(^3\) Hagenbjörk and Blomvall (2019) find that the elliptic copulas, i.e. the Gaussian-copula and Student’s \(t\)-copula, are suitable to model the joint variation of risk factors. Loosely speaking, a Gaussian copula models the joint variation as if the underlying random variables were multivariate normally distributed. Using a Gaussian copula with univariate normally distributed variables is equivalent to modeling them using a multivariate normal distribution. Similarly, a Student’s \(t\)-copula models joint variation as if the underlying random variables were distributed according to a multivariate Student’s \(t\)-distribution. The Student’s \(t\)-copula has higher tail dependence than a Gaussian copula, but the Student’s \(t\)-copula converges to a Gaussian copula as the degrees of freedom increase (Aas, 2004). In practice, we use the copula to obtain correlated uniform random variables, \(u_i, i = 1, \ldots, k\), for the \(k\) risk factors. Correlated random numbers used in the models (7.19) or (7.20) are obtained via the inversion method, \(\varepsilon_i = F_i^{-1}(u_i)\), where \(F_i^{-1}\) is the inverse cumulative distribution function.

\(^3\)The Gaussian copula method for valuing CDO’s proposed by Li (2000) has been blamed for causing the global financial crisis. Salmon (2012) describes it as a “recipe for disaster” in his article “The Formula that Killed Wall Street”. However, Watts (2016) states that it was not the copula itself but the modest assumptions of the underlying default correlations that caused the underestimation of the risk in these instruments. He concludes that “in order to avoid a disaster; the cooking is more important than the recipe”.

57
7.2.3 Measuring Risk by Simulating Term Structures

Given a model for the risk factors, term structure scenarios for one time period ahead are easily obtained through Monte Carlo simulation,

$$r_{t+1} = r_t + E_k \xi_t.$$  
(7.21)

An example of forward rates for 1,000 one-day-ahead scenarios for the Blomvall (2017) and cubic spline method can be seen in Figure 3. As can be see, the term structures for the cubic spline method are oscillating more and are much more volatile than for the Blomvall (2017) method. In fact, the total volatility of the term structures is 52% higher for the cubic spline-based scenarios.

By valuing a portfolio for many different scenarios, a distribution of the portfolio value can be obtained. During the 1990s, Value at Risk (VaR) was widely adopted for measuring market risk in portfolios, but its origins can be traced back to capital requirements the New York Stock Exchange imposed on member firms in 1922 (Holton, 2002). VaR is defined as the loss (with sign), $X$, in portfolio value when the $\alpha \in (0, 1)$ worst-case occurs over a given time horizon. For $Y := -X$ VaR is the $(1-\alpha)$-quantile of $Y$,

$$\text{VaR}_\alpha(X) = - \inf \{ x \in \mathbb{R} : F_X(x) > \alpha \} = F_Y^{-1}(1 - \alpha).$$  
(7.22)

VaR has been criticized because it is not a coherent risk measure since it violates sub-additivity, $\text{VaR}(X+Y) \nleq \text{VaR}(X)+\text{VaR}(Y)$ for some choices of portfolios $X$ and $Y$. This means that VaR may be greater when combining two portfolios than when keeping them separated. While not all risks are diversifiable, the risk should not be able to increase. As a consequence, VaR has been criticized for providing a false sense of security and for ignoring tail risk. Another drawback of VaR is that it is non-convex, and therefore not suitable to use as objective or constraint in optimization problems.

Conditional Value at Risk (CVaR), also known as expected shortfall (ES), is closely related to VaR. CVaR is defined as the expected value of the loss in portfolio value when the $(1 - \alpha)$-quantile of $Y$ is exceeded. In terms of VaR, it can be defined as

$$\text{ES} = - \frac{1}{\alpha} \int_0^\alpha \text{VaR}_\gamma(X) d\gamma.$$  
(7.23)

CVaR is a coherent risk measure (Embretch and Wang, 2015), and Rockafellar et al. (2000) show that it is possible to transform the problem into
Accurate Measurements of Interest Rate Risk

![Graphs showing forward rates](image)

**Figure 3:** 1,000 one-day-ahead forward rate term structure scenarios for the Blomvall (2017) and cubic spline methods.

A linear programming (hence convex) problem. The Basel Committee on Banking Supervision (2013) recognizes the weaknesses of VaR and propose moving the quantitative risk metrics system from VaR$_{99\%}$ to ES$_{97.5\%}$, see Chang et al. (2019) for an evaluation of the two risk measures.

Hagenbjörk and Blomvall (2019) model the entire distribution using the Monte Carlo-simulation approach described by (7.19) and (7.21). The accuracy of Value at Risk (and implicitly ES) is studied out of sample by comparing realized portfolio outcomes to simulated values over 15 years by using a kernel density estimator. For a 1-day horizon, the Monte Carlo model
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

based on Blomvall (2017) term structures accurately measures $\text{VaR}_{0.05}$. Even though the corresponding cubic spline model underestimates the risk by breaching the VaR, the median risk is substantially higher. The dynamic Nelson and Siegel (1987) model overestimates the risk but still produces relatively low values. For the commonly used 10-day horizon, all methods overestimate the risk, while the Blomvall (2017) method produces the lowest median values. The overestimation suggests that noise is still present in the term structures to some degree, and this is confirmed by studying the autocorrelation for the principal components. Hence, even more accurate measurements of the term structure are needed to reduce the risk further. In order to mitigate the problem with auto-correlation, the price movements across time should be taken into account. However, this is outside the scope of this dissertation.

As seen in this chapter, specifying the dynamics for the systematic risk factors allows us to measure risk. It also makes risk management possible via stochastic programming. This is the topic for the next chapter.
Hedging Interest Rate Risk

Hedging risk involves taking market positions intended to reduce or completely eliminate some or all risk factors in a portfolio. As described in Chapter 7, measuring risk has traditionally been closely related to hedging risk since the approach when measuring risk has been to identify risk factors and immunize the portfolio to these. With the modern view of risk, which also involves modeling the dynamics of the risk factors, hedging can be carried out more cleverly. By knowing how each risk factor affects the portfolio risk, the cost of reducing risk can be taken into consideration. Hedging is an important task for the market makers who provide liquidity in the over-the-counter traded (OTC) derivatives market and, as seen in Chapter 2, these markets are huge. The market makers provide bid and ask quotes and earn the difference when traders cross the spread. However, market makers do not wish to be exposed to risk. With the upcoming “fundamental review of the trading book” (FRTB) directives by Basel Committee on Banking Supervision (2019), banks are required to hold capital based on the $\text{ES}_{0.975}$ risk measure.

8.1 Traditional Hedging Methods

Hagan and West (2006, 2008) present three methods for hedging interest rate risk named bumping, boxes, and waves. These methods rely on computing the portfolio sensitivity to different permutations of the term structure. Bumping implies bumping each quote of the hedging instruments with
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

a small amount, usually 1 basis point\(^1\). The term structure is recalibrated using the bumped yield, and the sensitivity is computed as a numerical derivative. Boxes and waves are two methods where the term structure of forward rates is modified directly using shapes that resemble boxes and triangular waves. Assuming that we can trade instruments of increasing maturities \(T_i, i = 1, \ldots, n\), we denote \(\mathcal{T} = 0 \cup \{T_i\}_{i=1}^n\). The boxes method implies piecewise parallel shifts of the forward rate term structure by height \(b\) according to

\[
a_{f,i}(t) = \begin{cases} 
0, & t < T_{i-1} \\
b, & T_{i-1} \leq t \leq T_i \\
0, & t > T_i,
\end{cases} \tag{8.1}
\]

where \(T_i \in \mathcal{T}\). The waves method uses the same technique, where the permutations form waves with peaks at each maturity point \(T_i \in \mathcal{T}\) according to

\[
a_{f,i}(t) = \begin{cases} 
0, & t \leq T_{i-1} \\
b\frac{t-T_{i-1}}{T_{i}-T_{i-1}}, & T_{i-1} \leq t \leq T_i \\
b\frac{T_{i+1}-t}{T_{i+1}-T_i}, & T_i \leq t \leq T_{i+1} \\
0, & t > T_{i+1},
\end{cases} \tag{8.2}
\]

Denote the initial portfolio value by \(V_0\) and let the gradient, \(g_0 = \nabla_\xi V_0|_{\xi=0}\), denote the sensitivities to movements in the \(m\) different risk factors \(\xi\). Similarly, let the matrix \(G\) contain gradients for the different tradable instruments stored column-wise. By performing a first-order Taylor approximation of the portfolio value, we obtain the portfolio sensitivity as a function of our holdings in the tradable instruments

\[
\Delta V(h, \Delta \xi) \approx \Delta \xi^\top g_0 + \Delta \xi^\top Gh. \tag{8.3}
\]

For the special case where \(m = n = \text{rank}(G)\), which holds for the three traditional Hagan and West (2008) methods, \(G\) is a square matrix which forms a linear system where the sensitivities can be eliminated by setting

\[
g_0 + Gh = 0 \iff h = -G^{-1}g_0. \tag{8.4}
\]

These three methods seek to eliminate all interest rate risk, which is possible in a setting without spreads. As Blomvall and Hagenbjörk (2019b) show, bumping and boxes also perform well over time in a setting with spreads, while the waves method is less stable. However, none of these methods are

\(^1\)1 basis point (bp) = 0.01\%.
Flexible Hedging using Stochastic Programming

capable of considering transaction costs. Figure 1 displays the performance attribution of a realistic backtest where bumping has been used to hedge a swap portfolio. It is evident that the extremely low risk seen in the lower right panel comes to the price of high transaction costs.

Figure 1: Performance attribution for the Hagan and West (2006) strategy bumping.

8.2 Flexible Hedging using Stochastic Programming

Stochastic programming is a technique where a scenario approximation of a stochastic optimization problem is solved. This technique enables flexible modeling of how the scenarios are generated and is thus highly compatible
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

with our view of modeling risk. In a two-stage stochastic programming problem, the decision is made at time $t_0$ and the stochastics affect the outcome in time $t_1$. These problems are generally solved faster than the scenarios can be created. For a multi-stage problem, the curse of dimensionality forces us to severely reduce the number of scenarios in order for the problem to be practically solvable. This, of course, reduces the accuracy of the approximation of the original problem intended to solve.

Blomvall and Hagenbjörk (2019b) present a stochastic programming method for hedging interest rate risk where transaction costs are taken into account. Portfolio variance is weighted against transaction costs using a hedging parameter in the objective function. The transaction costs are inferred from prices based on term structures from the Blomvall (2017) method. The general optimization problem is

$$\begin{align*}
\minimize_{\Delta x_P, \Delta x_R} & \quad \frac{1}{2} \text{Var}(\tilde{W}_{t+1}) + \alpha \left( c_{P,t}^T \Delta x_P + c_{R,t}^T \Delta x_R \right) \\
\text{subject to} & \quad \Delta x_P, \Delta x_R \geq 0,
\end{align*}$$

(8.5)

where $\alpha$ is the hedging parameter, $c_{P,t}$ and $c_{R,t}$ denotes the proportional transaction costs for entering the contracts. As shown by Blomvall and Hagenbjörk (2019b), a scenario approximation of (8.5) can be written as quadratic programming problem that can be solved efficiently. By using a very basic model to generate term structure scenarios, we obtain a flexible hedging method that is preferred over the traditional methods for any risk-averse investor.

This hedging application demonstrates what it possible to achieve in terms of improved decision-making, when all the pieces presented in the previous chapters of part 2 are in place. In the bottom-up approach taken in this dissertation, measurements of financial quantities form the basis for these results. As seen, it is of utmost importance that these measurements are accurate since they affect the results in all of the following steps.
Credit Risk
Credit Risk

Many financial assets are exposed to credit risk, implying that the claims might not be fully paid in case of a default. Derivatives are usually collateralized, meaning that the counterparty having a negative value must post securities or cash to cover the negative value in order to provide security in case of default. Bonds, on the other hand, are particularly vulnerable since all cash flows, including the redemption value, are fully exposed. The timing and effects of defaults are hard to predict, and therefore, credit risk is hard to model. Despite the fact that the mathematics behind credit risk to a large extent was already fully developed, only a few people successfully predicted the global financial crisis in 2008. However, risk modeling is generally not about predicting a certain outcome, but to assign the correct probabilities to the range of possible outcomes. This also turned out to be problematic as some risks were severely underpriced during the global financial crisis.

As Schönbucher (2003) breaks down the concept of credit risk, additional subcategories emerge. The most basic concept is called arrival risk and deals with whether default occurs or not. Arrival risk and is related to the probability of default, whether it occurs or not. Timing risk includes information of the arrival risk for all possible time horizons and involves information of the probability density function of the time of default, or its even more local representation default intensity or hazard rate, which is defined later\(^1\). Given that default occurs, it is uncertain to which extent the claim can be recovered. Measured as a proportion of the total claim

\(^1\)Some authors make a distinction between default intensity and hazard rate under certain conditions, see (Schönbucher, 2003, section 4.4) for more information. In this dissertation, the terms are used anonymously.
received upon default, the probability distribution of this recovery rate embodies recovery risk. Contagion risk describes the risk that several obligors default together. This can be further broken down into joint arrival risk and joint timing risk. The market expectations of recovery risk, and especially timing risk can largely affect the price of financial instruments without any default occurring. These type of risks arising from movements in different risk factors is called market risk. A relatively recent example of this is the European debt crisis, where several European sovereign states experienced financial stress. Greece was hit hardest, but other countries include Ireland, Portugal, Spain, Italy, and Cyprus. Despite sharing the same risk-free Euro rates, the market-expectations of the default probability of these countries rose tremendously. For a while, Greek 10-year government bond yields exceeded 30% around the turn of the year 2011-2012. At the same time, the German 10-year sovereign bond yield was merely 2%. At the time of completing this dissertation, the 10-year Greek sovereign bond trades around 2%, while the German 10-year yield has dropped into the negative territory.

Since market risk models are dynamic, they add an additional layer of complexity. To avoid arbitrage opportunities in the model, market risk must be modeled consistently with timing and recovery risk, the changes in these risks, and other market prices. For a model to capture all of the above-mentioned types of credit risk, the complexity of the model is severely increased. Depending on the application, one has to determine which aspects are important to model. For this thesis, which focuses on the pricing of corporate fixed- and floating rate bonds, a dynamic model is not necessary. In addition, many of the above-mentioned types of risk do not need to be modeled explicitly. Modeling usually starts with simple models, which can be further extended into more complex models. The aim of the credit-risk part of this dissertation is to measure the term structure of hazard rates accurately. The purpose of this is to be able to identify the systematic risk factors inherent in the default intensities, i.e. the market risk stemming from timing risk. Ideally, accurate measurements of the term structure of recovery rates can be carried out simultaneously. This would enable identification of the systematic risk factors of recovery risk. Obtaining accurate measurements also opens the door for the possibility to measure the risk premiums associated with these risk factors. However, this is outside the scope of this dissertation.
9.1 Modeling Credit Risk

It is evident that credit risk makes investing considerably more difficult. Investors will, of course, require an increased yield to maturity to be willing to take on the risks described earlier, and this yield will be determined by the market for those securities traded on the secondary market. A primitive but widely used measure for credit risk is *credit spreads*. This refers to the difference in yield to maturity between a corporate bond and a risk-free treasury bond. A crude approximate relation between the credit spread, $D$, (risk-neutral) probability of default $P_Q$ and recovery rate $R$ can be computed as

$$P_Q = \frac{D}{1 - R}.$$  \hspace{1cm} (9.1)

The simplicity of this formula makes it very appealing to use, but there are lots of unrealistic assumptions hiding behind it. Only two factors are assumed to drive credit spreads, namely default probability and recovery rate, and since recovery rates are limited to the range 0% to 100%, the credit spread should never be larger than the probability of default. van Deventer (2017) uses the sophisticated KRIS system from the Kamakura Corporation to extract default probabilities for over 5.9 million bond transactions from the TRACE system and found that this inequality was violated in 96.37% of all transactions.

It is evident that the simple credit spread model (9.1) is not rich enough to capture the essence of credit risk. Assuming that a bond cannot default, pricing it by discounting its cash flows by the risk-free term structure is a rather straightforward task. A natural question is whether it is possible to find a single defaultable term structure to value bonds of the same seniority and creditworthiness? Unfortunately, the answer is no. It is actually not the arrival risk that causes the problem but rather the recovery. Assuming the recovery rate is zero, it is possible to obtain a defaultable term structure by adding the term structure of hazard rates to the risk-free forward rates. This way, defaultable bonds can be valued using the same pricing formula as described earlier in equation (3.6). With a non-zero recovery rate, it is not possible to make this aggregation of interest rates and credit risk into a single term structure.

To model credit-risk in a more realistic setting, reduced-form models were introduced in the early '90s. The name stems from the fact that credit risk is reduced into a purely statistical model where the default is seen as an
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

unpredictable event occurring from some intensity process. The valuation of bonds is thereby disconnected from the issuer’s financial information. Instead, this information is implicitly taken into account when calibrating the model to market prices.

The structural approach from the ’70s instead takes its starting point from the balance sheet of the modeled firm. By assuming a stochastic process for the market value of the firm’s assets, default is modeled as the situation when the assets cannot be used to repay the debt in full. As opposed to reduced-form models, structural models output credit spreads endogenously. However, more recent studies\(^2\) show that structural models have problems matching the credit spreads observed in the market.

The earliest type of model to handle credit risk was scoring models, developed in the ’50s. Scoring models are used to calculate a score from performance indicators derived from the company’s accounting information. The performance indicators and their corresponding coefficients are computed through statistical analysis of samples of companies where some proportion have defaulted. The likelihood of default is given from the score, either directly or by comparison to benchmark values.

The three main approaches described above have different uses. Structural models are well-suited for analyzing questions from corporate finance, such as optimal capital structure or bankruptcy timings. For pricing and hedging, reduced-form models are preferred according to Jarrow and Protter (2004). In this dissertation, the need is for a model to price defaultable bonds. However, we want a model that easily can be further extended to risk management. Therefore, the reduced form modeling approach is presented separately in Chapter 10. The rest of this chapter will focus on defaults from a legal and practical view in section 9.2, empirical properties of recovery in section 9.3, credit ratings in section 9.4, score and structural models in section 9.5 and 9.6.

9.2 Defaults in Practice

Since this dissertation deals with the modeling of default probabilities and recovery, we start by describing these concepts from a more practical point of view. First of all, there is no standard definition of default, but according

\(^2\)See section 9.6 for a more in-depth review containing references.
Defaults

to the Basel Committee (2001), a default is considered to have taken place when one or several of the following events has occurred:

- It is determined that the debtor is unlikely to pay the debt obligations (principal, interest, or fees) in full.
- A credit loss event associated with any obligation of the debtor, such as charge-off, specific provision, or distressed restructuring involving the forgiveness or postponement of principal, interest, or fees.
- The debtor is past due more than 90 days on any credit obligation.
- The debtor has filed for bankruptcy or similar protection from creditors.

As pointed out by Schuermann (2004), this definition enables defaults that result in no loss to the owner of the claim in case a debtor makes a full loan payment more than 90 days past due. Even though far from all defaults are the result of bankruptcy, many still are. Schuermann (2004), therefore, examines the anatomy of U.S. bankruptcies and concludes that a defaulting company typically goes through four stages.

1. The last cash paid date is considered the starting point of the time-line of the bankruptcy process, even though this date naturally becomes known ex-post.

2. Default is considered to occur at some later point when the fulfillments are not met.

3. Bankruptcy can be declared from the time of default, but this does not need to be done immediately. Bankruptcy in the U.S. can be filed under two different chapters, where Chapter 11 is the most common one and permits reorganization. Chapter 7 instead governs the liquidation process.

4. Emergence from bankruptcy typically occurs from two to four years after the last-cash-paid date. This emergence can be either as a going concern or via liquidation.

The laws and practical applications of default, liquidation, and reconstruction may differ between countries. According to our findings, Sweden and
the U.S. have very similar regulations. Since the goal was to study the U.S. bond market, we give a short description of the procedure according to the U.S. bankruptcy law and practice.

### 9.2.1 Liquidation According to Chapter 7 of the United States Bankruptcy Code

As reported by Eberhart et al. (1990), liquidation according to Chapter 7 involves the appointment of a trustee who collects the non-exempt property of the debtor, sells it, and distributes the proceeds to the creditors. The value is distributed to the suppliers of capital in such a way that senior creditors are fully satisfied before any distribution is made to junior creditors and lastly, shareholders. However, there have been several studies showing that this, so-called absolute priority rule, often becomes violated. According to Schuermann (2004), it has been shown that in 65-80% of bankruptcies, even shareholders receive money before debtholders have been fully paid off. The main reason for this is that some creditors agree to violate the absolute priority rule in order to increase the speed of resolution of the bankruptcy process. The fact that expenses due to the administration of the bankruptcy are senior to all other claims is one reason for this. Senior claim holders thus have the incentive to reduce the junior claim holders’ interest in holding up the bankruptcy process by paying them. Eberhart et al. (1990) show that bondholders, on average, benefit slightly from violations of the absolute priority rule.

### 9.2.2 Company Reconstruction According to Chapter 11 of the United States Bankruptcy Code

Chapter 11 bankruptcies are more common but also more complex than Chapter 11 liquidations. The reconstruction allows the operations of the company to continue while the company is shielded from its creditors. Meanwhile, all major business decisions must be approved by the bankruptcy court. Bris et al. (2006) examine 286 bankruptcies handled by courts in Arizona and New York between 1995 and 2001. Their data show that bankruptcies were filed under Chapter 11 in 221 (77%) of the cases and that the companies filing for Chapter 11 bankruptcy were substantially larger and also had more creditors, on average, than those filed under Chapter 7 (159
compared to 27). They also conclude that contrary to earlier beliefs, liquidation under Chapter 7 is neither faster nor cheaper than reorganizations under Chapter 11. In addition, liquidation under Chapter 7 provides lower recovery rates to the creditors. Bharath et al. (2010) points out that Chapter 11 has become more creditor friendly over the years as the bankruptcy time has been reduced, from 23 months prior to 1990, to 16 months after the year 2000. Even though Chapter 11 does not require the absolute priority rule to be followed, the frequency of violations has been reduced dramatically over the last decades, according to Bharath et al. (2010). From a violation rate of 75% prior to 1990, the rate dropped to 22% during 1991-2005, while becoming as low as 9% in the cases between 2000 and 2005.

9.2.3 Indentures, Covenants and Debt Acceleration

Bond indentures are legally binding agreements between an issuer and the bondholders. According to Guha (2002), most indentures are based on a model indenture form which regulates “Events of Default and Remedies” among other things. The indentures usually include covenants, which can be either affirmative or negative, as reported by Hahn (2009). Affirmative covenants require the issuer to meet its obligations and to report to the creditor on a regular basis. Negative covenants prohibit the debtor from initiating any action that may increase the default risk of the debt. This may involve restrictions on the dividend payouts or the capital structure of the company. It may also include protection against events that substantially changes the borrowers’ risk profile. This includes mergers, consolidations, and credit rating downgrades.

A debt acceleration clause is commonly specified in the bond indenture, and becomes available if the debtor fails in any of the covenants specified in the indenture. This implies that the holder, or a trustee representing holders of a large share of the total outstanding principal amount (usually at least 25%), may declare the principal amount, and in many cases, also the accrued interest, immediately payable. According to Guha (2002), this provides bondholders with an absolute and unconditional right of the principle, plus the accrued interest if not stated otherwise. Hahn (2009) describes the prime role of debt acceleration clauses as enhancing the efficiency of corporate governance. First, it reduces the debt agency costs, which are senior to all other payments. Secondly, acceleration is an effective means to initiate bankruptcy and prevent future mismanagement of the issuing company.
Finally, it is a mean of encouraging creditors with conflicting interests to negotiate collectively. Acceleration clauses thereby serve the creditors both prior to and after a default.

### 9.3 Recovery in Practice

According to the absolute priority rule, the company’s assets must be distributed to bondholders in order of seniority when a company is liquidated, according to Chapter 7. The fact that the absolute priority rule rarely is violated in Chapter 11 reconstructions makes seniority the most important determinant for recovery. *Senior* bonds prior to *junior*, also known as *subordinated* bonds. There is also an additional category named *senior subordinated*, which is senior to other subordinated debt but junior to senior debt. Apart from this, debt can be *secured*, meaning that the issuer is backing its value with some collateral, such as property or other financial securities. Secured bonds, therefore, have higher recovery rates associated with them.

Altman and Kishore (1996) provide a thorough investigation of recovery rates from observed defaults between 1978 and 1995. They compare recovery rates among different seniorities, industries, and credit ratings and find the average recovery rate to be 41.70 cents per dollar face value among the more than 700 defaults. The mean (median) recovery rate for senior secured bonds were 57.9% (51.0%), for senior unsecured 47.7% (40.7%), for senior subordinated 34.4% (27.9%) and for subordinated 31.3% (32.0%). As expected, seniority is an important determinant of recovery rates. They also sorted companies into 61 industry sectors based on SIC codes and group them into 18 industry categories. They find that “public utilities” and “chemicals, petroleum, and plastics” experienced significantly higher recovery rates than other categories. They also compare how the bonds’ credit rating at the time of issue determines the recovery rate but did not observe any differences between investment-grade-rated bonds and speculative-grade-rated bonds after accounting for seniority.

Altman et al. (2005) empirically study the relationship between default and recovery over the period 1982-2002 and find a negative correlation. Under periods of financial distress, where defaults are more common, recovery rates tend to be lower. They also find defaults to be negatively correlated to interest rates. This is explained by interest rates generally being lower in
periods of financial distress. Schläfer and Uhrig-Homburg (2014) also find support for a pronounced negative relation between default probability and recovery rates when trying to disentangle recovery risk from credit risk. They find that, apart from seniority, both firm-specific and industry-specific factors matter for recovery ratings. Issuers with low leverage and strong liquidity experience higher recoveries, even more so if the business sector is in good financial condition. By relating implied recovery rates to historical realizations, they find evidence for a systematic component in recovery risk with an associated risk premium.

Modeling of recovery risk is an overlooked area of research, which can be explained by the problem of separating recovery risk from timing risk. In reduced-form models, recovery is usually modeled by some recovery scheme, where the separation is mathematically impossible under some schemes. Even if the separation is possible, it is still extremely complicated to infer implied recovery rates from bond prices since they seem to be affected by a large number of company- and industry-related factors apart from seniority. Finding a dynamic model is also complicated by the fact that recovery risk seems to be correlated to both timing risk and interest rate risk, at least under the physical measure.

### 9.4 Credit Ratings

A credit rating is a grade based on an evaluation of the credit risk of a prospective debtor, predicting their ability to repay the debt. There are three independent main rating agencies: Standard & Poor's (S&P), Moody’s, and Fitch. Although the notation of credit ratings differs somewhat among the agencies, equivalence is supposed to exist among different scales to facilitate comparison between their 21-24 different ratings. There are no explicit default probabilities associated with the different ratings, but historical data can be used to assess the (physical) probability of default.

A debt issuer may pay one or several of the rating institutions to rate their debt as a service to their investors. This leads to conflicts of interest since the companies paying for rating services want as high rating as possible to obtain lower borrowing costs.

Becker and Milbourn (2011) investigate whether the increased competition from a third rating agency (Fitch) entering the market affects the ratings. They find that as rating levels go up, the correlation between ratings and
market-implied yields fall, and the ability of ratings to predict default deteriorate.

Rating institutions also provide additional financial services apart from rating debt. Baghai and Becker (2018) investigate whether these services affect the credit ratings from the same agency and find clear evidence that an agency that receives non-rating revenue from an issuer, on average rates which issuers securities 0.3 notches above other agencies. The higher the revenue, the bigger the difference. Furthermore, they find that default rates are higher for firms that have paid for non-rating services.

Another important question is whether credit ratings imply the same probability of default between different asset classes, and in fact, this is enforced by the Dodd-Frank Act (United States, 2010). Cornaggia et al. (2017) investigate whether this holds over the most recent 30-year period and find that ratings exhibit persistent differences across asset classes. They find that the divergent rating behavior during the global financial crisis was extreme, but not an isolated incident. Therefore, they conclude that it will be difficult to enforce the comparability of ratings between asset classes. For structured products and especially collateralized debt obligations (CDOs), credit ratings contain little or almost no information about the historical default probabilities. The different bond issuer types municipal, corporate, sovereign, and financial, are also compared. There seem to be significant discrepancies in how well credit ratings reflect historical default probabilities among the different issuer types. The accuracy of ratings decreases in the presented order. Studying rating transitions from Moody’s, they also find evidence that corporate bonds receive more generous ratings at issuance than municipal or sovereign bonds.

Cursio and Baek (2016) find that companies are likely to make changes to their capital structure in order to avoid a credit rating downgrade, but not to achieve an upgrade. The rationale behind this may be that financial markets punish issuers that have a downgrade in credit ratings via covenant restrictions.

It is evident that the way the credit rating market operates is a source of conflicts of interest. However, credit ratings from the three leading credit agencies are the only measure of creditworthiness available to us. Since the credit risk part of this dissertation only concerns bonds, credit ratings seem to be an accurate descriptor of credit risk within the same issuer type. Nonetheless, comparing credit ratings between different issuer types, and especially financial corporations, seems to be more problematic.
9.5 Scoring Models

The first, and most well-known, scoring model is the Altman (1968) $Z$-score model. A sample of 33 bankrupt and 33 non-bankrupt manufacturing companies was used to extract five significant performance indicators, $X_1, \ldots, X_5$. In order to determine their weight coefficients, multiple discriminant analysis was used. By multiplying the weights by the number for each performance indicator, a $Z$-score is obtained according to

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5,$$

where

- $X_1 = \text{Working capital/Total assets}$
- $X_2 = \text{Retained Earnings/Total assets}$
- $X_3 = \text{Earnings before interest and taxes/Total assets}$
- $X_4 = \text{Market value equity/Book value of total debt}$
- $X_5 = \text{Sales/Total assets}$.

The calculated $Z$-score is compared to two benchmark values and if $Z < 1.81$ there is a high probability of bankruptcy while a score $Z > 2.99$ indicates that the company is likely to avoid financial distress.

Over the years, the $Z$-score model was extended to incorporate privately held companies ($Z'$-score) and non-manufacturing companies ($Z''$-score), both with different weights and benchmark values. According to Altman (2002), $Z$-score models attain a 68% to 80% level of accuracy when predicting defaults two-years prior and 82% to 96% one year prior to the bankruptcy, based on out-of-sample company bankrupts. Ohlson (1980) criticizes multiple discriminant analysis and instead uses a logistic regression model with nine factors to derive a scoring model based on data from over 2,000 companies. The main advantage of logistic regression models is that the probability of default is outputted directly. Bellovary et al. (2007) make a comprehensive review and comparison of different scoring models from 165 studies published between 1965 and 2007. Here, the Ohlson (1980)-model obtains a 96% accuracy in both one- and two-year span.

Despite being useful tools for predicting defaults, scoring models are not designed for pricing financial instruments. For this purpose, structural and especially reduced-form models, are better suited.
9.6 Structural Models

The structural approach is based on modeling the balance sheet of the issuer. Shareholders are seen to hold a call option on the firm’s total asset value, which they are free not to exercise when the assets cannot fully repay the debt in the case of a default. The seminal paper by Merton (1974) was the starting point for structural models and was published just one year after the groundbreaking option-pricing paper by Black and Scholes (1973). The company’s assets, $V$, is assumed to follow the process

$$dV_t = \mu dt + \sigma V_t dW_t,$$

and default may occur if $V_T$ is insufficient to pay back the outstanding debt, $D$, at its maturity $T$. Using the balance sheet identity, the equity at time $T$, can be written as $E_T = V_T - D$, and in case of default when $E_T < 0$, the shareholders are not obligated to supply additional capital. Since the equity value at maturity thus can be written as $E_T = \max(V_T - D, 0)$, the shareholders hold a European call option on the firm’s value with maturity $T$ and strike price equal to $D$. Bondholders receive $B_T = \min(D, V_T) = D - \max(D - V_T, 0)$, meaning that in case of default, bondholders receive a recovery of $V_T < D$. Applying the Black and Scholes (1973) pricing model, values for equity and debt at time $t$ can be obtained as

$$E_t = C^{BS}(V_t, \sigma, r, T-t),$$
$$B_t = D e^{r(T-t)} - P^{BS}(V_t, \sigma, r, T-t) = V_t - C^{BS}(V_t, \sigma, r, T-t),$$

where $C^{BS}(V_t)$ and $P^{BS}(V_t)$ are prices for European call and put options, $\sigma$ and $r$ are the (constant) asset volatility and risk-free rate and the last equality follows from the balance identity. The defaultable bond yield at time $t$ thus satisfies

$$B_t e^{y(t, T)(T-t)} = D,$$

or

$$y(t, T) = \frac{1}{T-t} \ln \left( \frac{D}{B_t} \right),$$

and the credit spread $s(t, T)$ is

$$s(t, T) = y(t, T) - r.$$

There are several drawbacks with the Merton (1974) model, many of them inherited from the assumptions behind the Black and Scholes (1973) model.
Other assumptions are that default only may occur at the maturity of the debt, $T$, and the fact that the market value\(^3\) of the firm’s assets is assumed to be continuously observable. Mathematically, the information available and the continuity of the process implies that the time of default is a predictable stopping time, which has zero default intensity at all times except the time of default (Schönbucher, 2000). This means that default does not come as a genuine surprise since it can be observed by watching the value process (9.3), which is assumed to follow a geometric Brownian motion according to (9.3). Since the increments of a Brownian motion are normally-distributed, it is “virtually impossible” for the value process, $V_t$, to reach $D$, as $T - t$ approaches zero. This implies that the credit spread, $s(t,T)$, tends towards zero as the time to maturity, $T - t$, approaches zero. This is not realistic since defaults may occur very unexpectedly, which is confirmed by the fact that credit spreads observed in the market cannot be explained under these assumptions (Eom et al., 2004). Several extensions to the original model have been proposed to overcome these weaknesses.

Black and Cox (1976) extend the Merton (1974) model by allowing for defaults before the maturity of the debt if the firm’s value hits a lower threshold, thus making the European option into a barrier option. Longstaff and Schwartz (1995) further extend this approach by considering the firm’s value model to be correlated to stochastic interest rates from the Vasicek (1977) model, reaching closed-form solutions for pricing fixed and floating rate debt. Zhou (2001) further extend this approach by introducing jumps to the firm’s asset value process, thereby making it a jump-diffusion process. This solves the problem with unrealistic short-term credit spreads. Collin-Dufresne and Goldstein (2001) allow firms to adjust outstanding debt levels in response to changes in firm value, thereby generating mean-reverting leverage ratios. Geske (1977) takes another approach to value risky securities with compound options using multivariate normal integrals, which have to be evaluated numerically. Eom et al. (2004) compare five different structural models by Merton (1974), Geske (1977), Longstaff and Schwartz (1995), Collin-Dufresne and Goldstein (2001), and Leland and Toft (1996). They conclude that all models have problems accurately predicting observed credit spreads. The Merton (1974) and Geske (1977) models severely underestimate the credit spreads, whereas Longstaff and Schwartz (1995) overestimate credit spreads for junk bonds while underestimating spreads for high-quality bonds.

---

\(^3\)The market value is usually different from the accounted value in the financial statements.
Duffie and Lando (2001) assumes that the accounting information is not perfectly observable, implying that default is a totally inaccessible stopping time and thereby has an intensity process associated with it. This links structural models to reduced-form models and enables realistic short-term credit spreads. Jarrow and Protter (2004) make the distinction between structural models and reduced-form models based on the assumptions of information. Structural models assume that the modeler has access to the same information as the firm’s management, whereas reduced-form models assume the same incomplete information as the market. In most cases, this imperfect knowledge leads to an inaccessible default time. They argue that the key distinction between structural and reduced-form models is not whether the default time is predictable or inaccessible, but whether the information set is observed by the market or not. Since pricing is done by the market, reduced-form models should be preferred when pricing or hedging according to Jarrow and Protter (2004).

For a more extensive description of the field of structural models, see the work of Schönbucher (2003); Lando (2004); Elizalde (2005b).
Reduced-Form Models

While structural models link the default to the value of the issuer, reduced-form models treat default as an unexpected event whose likelihood is governed by a default-intensity process. The analysis of default is most easily done using the continuous-time version of hazard rates. Following Lando (2004), we consider the time of default to be a non-negative random variable $\tau$. We assume that the distribution of $\tau$ has a continuous probability density function $f$, so that the probability function, $F$, and the survival function, $S(t)$, the probability of surviving up to time $t$, are linked as

$$P(\tau \leq t) = F(t) = 1 - S(t) = \int_0^t f(u)du. \quad (10.1)$$

The hazard rate $h$ is defined as

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} = -\frac{d}{dt} \ln S(t), \quad (10.2)$$

where the last equality follows from identifying that the the probability density function $f(t)$ is the derivative of the density function $F(t)$. Since $S(t) = 1 - F(t)$, we have

$$\frac{d}{dt}S(t) = -f(t), \quad (10.3)$$

and the quotient between the probability density function $f(t)$ and the survival function $S(t)$ can be interpreted as $-\frac{S'(t)}{S(t)}$, which is the derivative of $-\ln S(t)$. 

81
Using the definition, we can rewrite the survival function in terms of hazard rates by integrating both sides of (10.2),

\[ S(t) = \exp \left( - \int_0^t h(u) \, du \right). \quad (10.4) \]

The hazard function is of particular interest in default modeling because of its link to conditional probability of default, i.e. the probability of default given no previous defaults. Using Bayes rule and the definitions in (10.1), we obtain

\[
P(\tau \leq t + \Delta t | \tau > t) = \frac{P(\tau \leq t + \Delta t, \tau > t)}{P(\tau > t)} = 1 - \exp \left( - \int_t^{t+\Delta t} h(u) \, du \right), \quad (10.5)
\]

which implies that \( h(t) \Delta t \) is approximately equal to the conditional probability of a default in a small interval, after \( t \), given no default up to and including \( t \),

\[
\lim_{\Delta t \to 0} \frac{P(\tau \leq t + \Delta t | \tau > t)}{\Delta t} = h(t). \quad (10.6)
\]

Note that we can compute this conditional default probability for all \( t \) at time 0. In a more realistic modeling situation, there are factors other than the time affecting the default probability. If the firm has survived up until time \( t \), we have access to information that was not available at time 0. Therefore, we want to build models that conditional on a more general information set \( F_t \) so that, loosely speaking,

\[
P(\tau \leq t + \Delta t | F_t) = \mathbb{1}_{\{\tau > t\}} \lambda(t) \Delta t, \quad (10.7)
\]

where \( \mathbb{1}_{\{\cdot\}} \) is the indicator function, taking the value 1 if \( \{\cdot\} \) is satisfied and 0 otherwise. \( F_t \) contains information on the survival up to time \( t \) and \( \lambda \) is a process which is adapted to (in fact predictable with respect to) the filtration \( \mathbb{F} \). This desired way of modeling does not serve as a definition and, therefore, not possible to build models upon. Hence, we need a more technical definition, but first, we take a brief detour via Poisson processes, which are important pieces to this puzzle.

### 10.1 Poisson Processes

To model default-arrival risk within reduced-form models, Poisson processes are often used. They allow for the local implied default probability being
Poisson Processes

proportional to the length of the time interval, just as desired in (10.7). Poisson processes are usually used to model either rare or discretely countable events, and defaults are both rare and discrete. A Poisson process is an increasing process $0, 1, 2, 3, \ldots$ with jumps at time $\tau_1 \leq \tau_2 \leq \tau_3, \ldots$. Even though defaults may occur more than once for a company in the case of reconstruction, one usually only care about the first time of default. Therefore, it is common to model default as the time of the first jump, $\tau_1$. Following the description by Schönbucher (2003), we present the assumption that forms the basis for how default arrival risk is modeled in all intensity-based models:

**Assumption 1.** Let $N(t)$ be a counting process with (possibly stochastic) intensity $\lambda(t)$. The time of default $\tau$ is the time of the first jump of $N$, i.e.

$$\tau = \inf \{ t \in \mathbb{R}_+ | N(t) > 0 \}.$$  \hspace{1cm} (10.8)

The survival probability is thus the probability of receiving no jumps,

$$P(0,T) = P(N(T) = 0 | F_0).$$  \hspace{1cm} (10.9)

Next we present the definition of a Poisson process. For an intuitive construction of a Poisson process, see (Schönbucher, 2003, 5.1.2).

**Definition 1 (Poisson process).** A Poisson process with intensity $\lambda > 0$ is a non-decreasing, integer-valued process with initial value $N(0) = 0$, whose increments are independent and satisfy,

$$P[N(T) - N(t) = n] = \frac{1}{n!} (T - t)^n \lambda^n \exp\{- (T - t) \lambda\},$$  \hspace{1cm} (10.10)

for all $0 \leq t < T$.

A Poisson process has no memory, meaning that the jump probabilities of disjoint intervals are independent. Hence, it is not more likely for a jump to occur just because it was a long time since the last jump. The inter-arrival times, $\tau_{n+1} - \tau_n$, of a Poisson process are exponentially distributed with density $P[\tau_{n+1} - \tau_n \in t dt] = \lambda e^{-\lambda t} dt$ and the probability of two jumps occurring at exactly the same time is zero. When dealing with Poisson-jump type processes in stochastic calculus, the following rules apply:

- $E[dN] = \lambda dt$.
- $M(t) = N(t) - \lambda t$ is a martingale.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

- Covariations $dN dW = 0$, $E[dN^2] = \lambda dt$.
- For any Brownian motion, $W$, $\langle N, W \rangle = 0$.

Using Assumption 1 and the definition of a Poisson process (10.10) where $n = 0$, the survival probability is given by

$$S(t) = e^{-\lambda t}. \quad (10.11)$$

A constant intensity, $\lambda$, implies a flat term structures of credit spreads, which is not consistent with market observations. This can be avoided by using an inhomogeneous Poisson process for which the intensity is a function of time, $\lambda(t)$.

**Definition 2** (inhomogeneous Poisson process). An inhomogeneous Poisson process with intensity function $\lambda(t) > 0$ is a non-decreasing, integer-valued process with initial value $N(0) = 0$, whose increments are independent and satisfy

$$P[N(T) - N(t) = n] = \frac{1}{n!} \left( \int_t^T \lambda(u) du \right)^n \exp \left\{ - \int_t^T \lambda(u) du \right\}, \quad (10.12)$$

for all $0 \leq t < T$.

The properties of an inhomogeneous Poisson process resemble those of a homogeneous Poisson process:

- $E[dN | \mathcal{F}_t] = \lambda(t) dt$.
- $M(t) = N(t) - \int_0^t \lambda(u) du$ is a martingale.
- Jump martingales and diffusion martingales are uncorrelated, see (Schönbucher, 2003) for more details and a proof.

Using the definition (10.12) with $n = 0$, the survival probability is obtained as

$$S(t) = e^{-\int_0^t \lambda(u) du}. \quad (10.13)$$

Now, the hazard rates depend on the time horizon, $t$, implying that the term structure of credit spreads is not flat but given by $\lambda(t)$. By suitable choice of the function $\lambda(t)$, every hazard rate term structure can be modeled. For the purpose of this dissertation, an inhomogeneous Poisson process fulfills
all needs. However, there is one further generalization that is usually made when specifying intensity-models. A Cox-process or doubly stochastic Poisson process, allows for stochastic intensity dynamics. This is needed for measuring risks or pricing credit derivatives whose payoffs are directly affected by volatility. Roughly speaking, Cox processes are Poisson processes with stochastic intensity. The jump probability of the process $N(t)$ is

$$P[N(t + \Delta t) - N(t) = 1] = \lambda(t)dt,$$

where $\lambda(t)$ is a stochastic process. To avoid arbitrage, we must ensure that this process never becomes negative. The jumps are now correlated via the path of $\lambda(t)$, meaning that if a jump occurs, it is likely for $\lambda(t)$ to be high, hence making further jumps more likely. A Cox-process where the entire path of $\lambda(t)$ is known in advance is an inhomogeneous Poisson process. In order to properly define the Cox-process, we need to go into detail.

### 10.2 Technical Details

Assume that the economic uncertainty is modeled via the probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t, \mathbb{Q}))$, where $\Omega$ is the set of possible states of the economic world and $(\mathcal{G}_t)$ represents the flow of information over time. $\mathcal{G} = \sigma\left(\bigcup_{t \geq 0} \mathcal{G}_t\right)$ is a $\sigma$-field (or $\sigma$-algebra), a family of events to which we can assign probabilities by a probability measure. An event is thus a subset of $\Omega$, a $\sigma$-field models information and a filtration that models information flow. $\mathbb{Q}$ denotes the risk-neutral (or martingale) measure, which serves as a pricing measure and exists in an arbitrage-free economy. All securities can be priced as expected discounted values under this measure, but the particular choice of $\mathbb{Q}$ depends on the assumptions on risk premiums. We reserve $\mathbb{P}$ to denote the physical measure.

Assume that there exists a background process in the form of an $\mathbb{R}^J$-valued Markov process $X_t = (X_{1,t}, \ldots, X_{J,t})$, representing $J$ economy-wide variables that may be either observable (state variables) or non-observable (latent variables). Let $(\mathcal{F}_t)_{t \geq 0}$ denote the filtration generated by $X_t$, i.e. $(\mathcal{F}_t)_{t \geq 0} = \sigma\{X_s; 0 \leq s \leq t\}$. There also exists a counting process, $N_t$, initialized at 0, that represent the default process of the issuer, where default occurs when $N_t$ jumps from 0 to 1. Let $\mathcal{H}_t = \sigma\{N_u; 0 \leq u \leq t\}$, and let $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$ be the filtration containing information of both the market variables, $X_t$, and the jump process, $N_t$. This implies knowing both the
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

evolution of the observable market variables and whether or not default has occurred. We only focus on the first jump $\tau$, but it is possible to extend this setup to allow for repeated defaults. Let $\lambda : \mathbb{R}^J \to \mathbb{R}$ be a non-negative (measurable) function. The goal is to construct a jump process, $N_t$, with the property that $\lambda(X_t)$ is the $\mathcal{G}_t$-intensity of $N$. Let $E_1$ be an exponential random variable with mean 1, which is independent of $(\mathcal{F}_t)_{t\geq 0}$. Define the default time, $\tau$, as the first jump of a Cox-process with intensity process $\lambda(X_u)$,

$$\tau = \inf \left\{ t : \int_0^t \lambda(X_u)du \geq E_1 \right\}. \quad (10.15)$$

When $\lambda(X_u)$ is large, the integral grows faster and reaches the (stochastic) level of $E_1$ faster. Therefore, the probability of default becomes higher. By ensuring that the integrated intensity stays finite over finite intervals, the model can be extended to incorporate repeated defaults.

From the definition (10.15), the survival probability can be computed as

$$S(t) = \mathbb{Q} \{ \tau > T | \mathcal{G}_T \} = \mathbb{E}^{\mathbb{Q}} \left[ \exp \left( -\int_0^T \lambda(X_u)du \right) \right]. \quad (10.16)$$

We also want to handle conditional expectations of relevant functions at arbitrary dates. If we stand at time $t$ and $\tau$ has not yet occurred, then the probability of default before $T > t$ is a function of the state variables $X_t$, and hence, $\mathcal{G}_t$ contains all the information needed. How to formally express this is addressed next.

Let $Z \in \mathcal{G}_t$ be a stochastic variable where $\mathbb{E}[|Z|] < \infty$, in order to make sure the conditional expectations exist. Then, there exists a $\mathcal{F}_t$-measurable random variable $Y_t$ such that

$$\mathbb{1}_{\{\tau > t\}} \mathbb{E}[Z|\mathcal{G}_t] = \mathbb{1}_{\{\tau > t\}} Y_t \quad (10.17)$$

for a proof, see (Bielecki and Rutkowski, 2004, lemma 5.1.2). Using this result, one can show that

$$\mathbb{1}_{\{\tau > t\}} E[Z|\mathcal{G}_t] = \mathbb{1}_{\{\tau > t\}} \frac{E[Z \mathbb{1}_{\{\tau > t\}}|\mathcal{F}_t]}{E[\mathbb{1}_{\{\tau > t\}}|\mathcal{F}_t]}, \quad (10.18)$$

or in particular

$$\mathbb{Q} \{ t < \tau \leq T | \mathcal{G}_t \} = \mathbb{1}_{\{\tau > t\}} \frac{\mathbb{Q} \{ t < \tau \leq T | \mathcal{F}_t \}}{\mathbb{Q} \{ \tau > T | \mathcal{F}_t \}}, \quad (10.19)$$

86
thus replacing the history of the full filtration \( G_t \), with the history of the state variable process \( X_t \) given in \( F_t \). This result is essential to derive a dynamic version of the default probability. Start by noting that

\[
Q(\tau > T | G_t) = \mathbb{E}^Q[\mathbb{1}_{\{\tau > T\}} | G_t],
\]

and use (10.18) to obtain

\[
Q(\tau > T | G_t) = \mathbb{E}^Q[\mathbb{1}_{\{\tau > T\}} | F_t] \frac{\mathbb{E}^Q[\mathbb{1}_{\{\tau > T\}} | F_t]}{\mathbb{E}^Q[\mathbb{1}_{\{\tau > t\}} | F_t]}.
\]

(10.21)

Studying the numerator, we obtain

\[
\mathbb{E}^Q[\mathbb{1}_{\{\tau > T\}} | F_t] = \mathbb{E}^Q[\mathbb{E}^Q[\mathbb{1}_{\{\tau > T\}} | F_T] | F_t] = \mathbb{E}^Q[\exp \left(-\int_0^T \lambda(X_u) du\right) | F_t] = \exp \left(-\int_0^t \lambda(X_u) du\right) \mathbb{E}^Q[\exp \left(-\int_t^T \lambda(X_u) du\right) | F_T]
\]

where the last equality follows from the fact that the information up until time \( t \) is known. Performing the same analysis for the denominator, the deterministic terms will cancel out, and we arrive at the final expression for the dynamic survival probability at an arbitrary time \( t \),

\[
Q(\tau > T | G_t) = \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \exp \left(-\int_t^T \lambda(X_u) du\right) \right] | F_t).
\]

(10.23)

For a Cox process,

\[
M_t = N_t - \int_0^t \lambda_u \mathbb{1}_{\{\tau > u\}} du,
\]

(10.24)

is a martingale, i.e. \( \mathbb{E}^Q [M_u - M_t | F_t] = 0 \), see (Lando, 2004, ch 5.4) for a proof. This property is fundamental to ensure the existence of a martingale measure in order to ensure absence of arbitrage when pricing using models that are based on Cox-processes.

10.2.1 Building Blocks

A few building blocks often recur when pricing defaultable claims, and therefore, these usually receive extra attention. Lando (1998) presents these as
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

(i) \( X \mathbb{1}_{\{\tau > T\}} \): A stochastic payment \( X \in \mathcal{F}_t \) occurs at a fixed date, \( T \), if there has been no default before time \( T \). An example of this is a European option whose issuer may default before the option’s expiration.

(ii) \( Y \mathbb{1}_{\{\tau > s\}} \): A stream of payments at a rate specified by the \( \mathcal{F}_t \)-adapted process \( Y \) which stops when default occurs.

(iii) A recovery payment \( Z \) at the time of default where \( Z \) is a \( \mathcal{F}_t \)-adapted stochastic process.

We start by considering a defaultable zero-coupon bond issued at time 0, with maturity \( T \), where it pays the redemption value of 1 if no default occurs and zero otherwise. Assume that, under the risk-neutral probability measure \( Q \), the default time \( \tau \) has an intensity \( \lambda(X_t) \), according to the setup described above. Assume that there is a short-rate process \( r(X_s) \) such that the default-free zero-coupon bond price at time \( t \) can be computed as

\[
P(t, T) = \mathbb{E}^Q \left[ \exp \left( -\int_t^T r(X_u) du \right) \bigg| \mathcal{G}_t \right]. \tag{10.25}
\]

Assuming zero recovery, the price of the defaultable bond is given by

\[
\tilde{P}(t, T) = \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \exp \left( -\int_t^T r(X_u) du \right) \mathbb{1}_{\{\tau > T\}} \bigg| \mathcal{G}_t \right]
\]

\[
= \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \mathbb{E}^Q \left[ \exp \left( -\int_t^T r(X_u) du \right) \mathbb{1}_{\{\tau > T\}} \bigg| \mathcal{F}_T \lor \mathcal{H}_t \bigg] \bigg| \mathcal{G}_t \right]
\]

\[
= \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \exp \left( -\int_t^T r(X_u) du \right) \mathbb{E}^Q \left[ \mathbb{1}_{\{\tau > T\}} \bigg| \mathcal{F}_T \lor \mathcal{H}_t \bigg] \bigg| \mathcal{G}_t \right]
\]

\[
= \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \exp \left( -\int_t^T r(X_u) du \right) e^{-\int_t^T \lambda(X_u) du} \bigg| \mathcal{G}_t \right]
\]

\[
\quad \quad \quad = \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \exp \left( -\int_t^T (r + \lambda) du \right) \bigg| \mathcal{G}_t \right]. \tag{10.26}
\]

We now want to replace the conditioning on \( \mathcal{F}_t \) with conditioning on the smaller filtration \( \mathcal{F}_t \). According to Lando (1998), the unit exponential random variable, \( E_1 \), which governs default, is independent of the \( \sigma \)-field \( \mathcal{F}_t \). In particular, \( E_1 \) is independent of the \( \sigma \)-field \( \sigma \left( \exp \left( -\int_t^T (r + \lambda) ds \right) X \right) \lor \mathcal{F}_t \),
and therefore

\[
\hat{P}(t, T) = 1_{\{\tau > t\}} E^Q \left[ \exp \left( -\int_t^T (r + \lambda(X_u))du \right) \bigg| G_t \right]
\]

\[
= 1_{\{\tau > t\}} E^Q \left[ \exp \left( -\int_t^T (r + \lambda(X_u))du \right) \bigg| F_t \right].
\] (10.27)

The short rate \( r(X_t) \) of the risk-free bond thus has been replaced with \( r(X_t) + \lambda(X_t) \). As stated in the chapter introduction, it is thus possible to obtain a (deterministic) defaultable-term structure by adding the hazard rate to the risk-free term structure. By replacing the payoff 1 with an \( F_t \)-adapted payoff \( X \), this example generalizes to the case of building block (i).

For building block (ii), we obtain

\[
E^Q \left[ \int_t^T Y_s 1_{\{\tau > s\}} \exp \left( -\int_s^t r_u du \right) ds \bigg| G_t \right]
\]

\[
= 1_{\{\tau > t\}} E^Q \left[ \int_t^T Y_s \exp \left( -\int_t^s (r_u + \lambda_s)du \right) ds \bigg| F_t \right],
\] (10.28)

and for building block (iii),

\[
E^Q \left[ \exp \left( -\int_t^\tau r_u du \right) Z_\tau \bigg| G_t \right]
\]

\[
= E^Q \left[ \int_t^\tau Z_s \lambda_s \exp \left( -\int_t^s (r_u + \lambda_u)du \right) ds \bigg| F_t \right].
\] (10.29)

The proofs for the second and third building blocks are carried out the same way as for the first. For the third building block, we note that \( Z \) is paid out at the time for default and that the density of the time of default is given by

\[
\frac{\partial}{\partial s} Q(\tau \leq s|\tau > t, F_T) = \lambda_s \exp \left( -\int_t^s \lambda_u du \right).
\] (10.30)

This section has been largely influenced by the work by Lando (1998, 2004); Elizalde (2005a). Other resources include the recent book by Jarrow (2018) and the more technical book by Bielecki and Rutkowski (2004).
10.2.2 Defaultable Coupon Bond

Using the building blocks described above, we can now derive an expression for a defaultable coupon bond with a stochastic recovery, $Z$, at the time of default. For the most general setting, we start by assuming that the bond pays a stochastic redemption value, $X \in \mathcal{F}_t$, at maturity and dividends according to a continuous process, $Y_s \mathbbm{1}_{\{\tau > s\}}$. The value at time $t$ is given by three terms

$$
P(t, T) = \mathbbm{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \int_t^T \exp \left( - \int_t^u (r_s + \lambda_s)du \right) \left( Y_u + \lambda_u Z_u du \right) \bigg| \mathcal{F}_t \right]$$

$$+ \mathbbm{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \exp \left( - \int_t^u (r_s + \lambda_s)du \right) X_u du \bigg| \mathcal{F}_t \right], \tag{10.31}
$$

where the first term handles dividends, the second term involves the recovery process and the last term the redemption. To arrive at this expression, all three of the building blocks are used. The type of bonds considered in this thesis are straight bonds and floating rate notes (FRN). A straight bond pays coupons of fixed size, $C/m$, based on some coupon rate, $C$, according to some schedule with some frequency, $m$, usually annually or semi-annually. The value of a coupon stream paying $n$ fixed-rate coupons at times $T_i > t, i = 1, \ldots, n$, can thus be expressed as

$$\mathbbm{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ \sum_{i=1}^{n} \frac{C}{m} \exp \left( - \int_t^{T_i} (r_u + \lambda_u)du \right) \bigg| \mathcal{F}_t \right] =$$

$$\mathbbm{1}_{\{\tau > t\}} \sum_{i=1}^{n} \frac{C}{m} \mathbb{E}^Q \left[ \exp \left( - \int_t^{T_i} (r_u + \lambda_u)du \right) \bigg| \mathcal{F}_t \right]. \tag{10.32}
$$

For an FRN, the coupons are linked to an underlying rate, usually LIBOR, whose tenor depends on the frequency of the coupon payments, usually quarterly or semi-annually. The LIBOR, $L_\tau(T_i, T_j)$, is an interest rate that is fixed at time point $T_i$ and paid at the later time $T_j$, where $\tau$ represents the tenor, and thus the time between $T_i$ and $T_j$. A floating rate note with credit spread $D$ and gearing $G$, linked to an underlying rate, the LIBOR $L$, pays the coupon $(G \cdot L(T_i, T_j) + D)\Delta_{i,j}$ at time $T_j$ for $(i, j) \in \mathcal{I}$, where $\Delta_{i,j}$ is the time between $T_i$ and $T_j$, measured with the appropriate day count convention. Note that all future coupons after the one following $t$, are stochastic. The value of the corresponding coupon stream of floating-rate
coupons at times $T_i, T_j > t, (i, j) \in \mathcal{I}(t, T), \ldots, n$, can thus be written as

$$1_{\{\tau > t\}} \sum_{(i,j) \in \mathcal{I}(t, T)} E^Q \left[ \exp \left( - \int_t^{T_i} (r_s + \lambda_s) ds \right) (G \cdot L(T_i, T_j) + D) \Delta_{i,j} \bigg| \mathcal{F}_t \right] =$$

$$1_{\{\tau > t\}} \sum_{(i,j) \in \mathcal{I}(t, T)} G \Delta_{i,j} E^Q \left[ \exp \left( - \int_t^{T_i} (r_s + \lambda_s) ds \right) L(T_i, T_j) \bigg| \mathcal{F}_t \right] +$$

$$1_{\{\tau > t\}} \sum_{(i,j) \in \mathcal{I}(t, T)} D E^Q \left[ \exp \left( - \int_t^{T_i} (r_s + \lambda_s) ds \right) \big| \mathcal{F}_t \right].$$

(10.33)

The problem here is the dependence between the interest rates and the underlying reference rate. However, it turns out that this is not a problem under the forward neutral measure $Q_T$. From Björk (2009, Proposition 26.8) it follows that

$$E^Q \left[ L(T_i, T_j) e^{-\int_t^{T_i} r_s ds} \big| \mathcal{F}_t \right] = P(t, T_j) E^T \left[ L(T_i, T_j) \big| \mathcal{F}_t \right],$$

(10.34)

which can be useful under the assumption of independence between interest rates and default intensities.

### 10.2.3 Recovery Schemes

The recovery process, $Z$, is usually modeled as the fraction, $\delta_t$, recovered of some value given, either exogenously or endogenously, via three primary recovery schemes. For a thorough review of recovery, see the book by Schönbucher (2003).

The recovery of face value (RFV) scheme considers the recovery rate to be an exogenous fraction of the face value of the defaultable bond. This is the closest to legal practice considering liquidation, as described earlier in section 9.2. Since the recovery is exogenously specified, there is nothing stopping us from including the accrued interest as well.

Jarrow et al. (1997) consider the recovery rate to be an exogenous fraction of the value of an equivalent default-free bond. This scheme is, therefore, called the recovery of treasury scheme (RT). As pointed out by Lando (2004), this scheme can be seen as a version of RFV where the recovered amount is discounted back to present value using the risk-free rate. As the majority of the bankruptcies are filed under Chapter 11, permitting reconstruction,
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Longstaff and Schwartz (1995) argues that this scheme better reflects this process.

Duffie and Singleton (1999) model the recovery as an exogenous fraction of the market value of the bond just before default. They show that this bond can be priced as if it was a default-free zero-coupon bond by replacing the usual short-term interest rate process, \( r_t \), with a default-adjusted short rate process \( \pi_t = r_t + \lambda_t L_t \), where \( L_t = 1 - R_t \) is the loss rate. This way, the timing risk, and the recovery risk are modeled together, which serve as both the advantages and disadvantages of the model according to Elizalde (2005a). This way, one can use well-known short-rate processes to model \( \pi_t \).

It is also easy to incorporate liquidity risk by introducing a liquidity spread component \( l_t \), such that \( \pi_t = r_t + \lambda_t L_t + l_t \). The main drawback is that it, due to the multiplicative nature of \( \lambda_t \) and \( L_t \), is impossible to separate \( \lambda_t \) from \( L_t \) by defaultable bonds alone.

10.3 Dynamics of Intensity-Based Models

In chapter 10.2, we derived expressions for pricing three common building blocks in terms of the short-term interest rate, \( r_t \), and default intensities, \( \lambda_t \).

In combination with the recovery schemes described above, these building blocks can be used for the valuation of defaultable securities and credit derivatives. However, we yet have to specify the dynamics of \( r_t \) and \( \lambda_t \), and there are two main approaches for doing this. The first is to model stochastic dynamics of \( r_t \) and \( \lambda_t \) in such a way that the necessary closed-form solutions for calibration can be derived analytically. These model specifications can help us understand the influence of various model parameters and their correlation. According to Lando (2004), the Kalman filter is a standard tool for estimating latent-factor models. A second approach is to extend the HJM-framework (Heath et al., 1992; Schönbucher, 1998) to incorporate a model of the credit spread and derive the arbitrage-free drift conditions.

10.3.1 A Two-Factor Gaussian model

By using two correlated extended Vasicek (1977) models for the short rate process, \( r_t \), and the short hazard rate process, \( \lambda_t \), we obtain the two-factor
Dynamics of Intensity-Based Models

Gaussian model

\[ \begin{align*}
    dr_t &= (k(t) - ar_t)dt + \sigma_t dW_t \\
    d\lambda_t &= (\tilde{k}(t) - \tilde{a}\lambda_t)dt + \tilde{\sigma}_t d\tilde{W}_t \\
    dWd\tilde{W} &= \rho dt.
\end{align*} \] (10.35)

(10.36)

(10.37)

Gaussian models are easy to work with. Any specification of the short rate and the intensity as affine functions of a Gaussian state variable diffusion gives expressions for both risk-free and defaultable bond prices which have closed-form solutions on the form (see (Schönbucher, 2003) for a derivation)

\[ \begin{align*}
    P(t, T) &= \exp \left( A(t, T; a, k, \sigma)B(t, T; \alpha) \right) \\
    \tilde{P}(t, T) &= P(t, T) \exp \left( A(t, T; \alpha, \tilde{k}, \tilde{\sigma})B(t, T; \alpha, \tilde{k}, \tilde{\sigma}) \lambda(t) \right)
\end{align*} \] (10.38)

where

\[ \tilde{k}(t) = k(t)\rho \tilde{\sigma}(t)\sigma(t)B(t, T). \] (10.39)

The simplicity of the two-factor Gaussian model is its strength. However, the simplicity leading to analytic tractability comes to the price of possibility for negative intensity spreads (and of course interest rates) since the Gaussian structure does not prevent this.

The two-factor Gaussian model can be extended to a multi-factor model. Duffie and Singleton (1997) and Duffee (1999) suggest that two-factor models perform rather poorly.

10.3.2 Cox-Ingersol-Ross Models

The Cox et al. (1985a) (CIR) short rate equilibrium model prevents interest rates from becoming negative by scaling the diffusion by the square root of the short rate

\[ \begin{align*}
    dr_t &= a(b - r_t)dt + \sigma \sqrt{r_t}dW_t. \\
    d\lambda_t &= \tilde{a}(\tilde{b} - \lambda_t)dt + \tilde{\sigma} \sqrt{\lambda_t}(\rho dW_t + \sqrt{1 - \rho^2}d\tilde{W}_t),
\end{align*} \] (10.40)

(10.41)

(10.42)

As Lando (2004) shows, a model specified by

is unfortunately not affine. Trying another similar specification results in a model where \( \lambda \) is not guaranteed to stay positive. Therefore, the typical
way to obtain a correlated CIR-process is to specify \( r_t \) as a CIR-process and letting

\[
\lambda(r_t) = \alpha + \beta r_t + s_t, \tag{10.43}
\]

where \( s_t \) is another CIR-process independent of \( r_t \). If \( \beta < 0 \), implying that the correlation between \( r_t \) and \( \lambda_t \) is negative, \( \lambda_t \) is unfortunately not guaranteed to stay positive. The potential of using CIR-processes is therefore rather limited.

Jamshidian (1996) was first to model default intensities using CIR-models. Duffie and Singleton (1997) uses a CIR-model to analyze the term structure of swap yields and Duffee (1999) uses an extended Kalman filter to calibrate a CIR-model to corporate bond prices.

### 10.4 Risk Premia

As explained in the chapter introduction, there are many different types of credit risk. A central question is whether the risks are undiversifiable, systematic risks entitled with a risk premium and if so, determining the size of the risk premium. The jump-to-default risk premium has been debated in the literature. Elton et al. (2001) show that the credit spread almost entirely can be decomposed into tax effects, expected losses from defaults, and a risk premium for bearing systematic risk. Making use of the Fama and French (1993) factors, they find strong evidence for the presence of a risk-premium that explains 85% of the remaining spread after accounting for expected losses and tax effects. Driessen (2004) uses a latent-factor reduced-form model with common and firm-specific factors plus a liquidity factor. He finds that tax and liquidity effects are important determinants of credit spreads but also that market-wide and firm-specific changes in credit spread are entitled to a risk premium. Huang and Huang (2012) use an analytically tractable jump-diffusion structural model to investigate the risk premia. They show that credit risk accounts for only a small fraction of yield spreads for investment-grade but for a much higher fraction of yield spreads for speculative-grade bonds. They find their conclusion to be robust under different reasonable parameter choices. Jarrow et al. (2005) show that, under some conditions resembling those of arbitrage pricing theory (Ross, 2013), the jump-to-default risk will be zero if the default process is assumed to follow a Cox-process. Bai et al. (2015) conclude that the jump-to-default risk is negligible compared to the contagion effect for an economy comprising
a large number of firms. They motivate this by the fact that in order for the jump-to-default risk to be priced, it cannot be conditionally diversifiable. This essentially means that defaults are independent, and there is a large amount of bonds trading in the market. Das et al. (2007) show that default clusters in time because of correlation among firms’ intensity processes and that, even after accounting for the intensities, there are contagion effects that cannot be explained by Cox-processes.

Taken together, there seem to be clear evidence for tax effects and a liquidity premium in the bond market. There also seems to exist a risk premium associated with the systematic risk inherent in the movements of the term structure of credit spreads, at least at a market-wide scale. It also exists a premium for bearing default risk, but evidence suggests that this is rather due to contagion risk than jump-to-default risk. In addition, Cox-processes may not be rich enough to model contagion risk.
Measuring the Financial Quantities on the Credit Risk Market

As seen in the previous chapters, credit risk is a topic that quickly becomes mathematically complex as there are many empirically observed properties that need to be modeled correctly. For the purpose of this dissertation, which is to identify the systematic risk factors on the credit market, the dynamics does not need to be modeled. This also eliminates the need for modeling contagion risk, as this risk affects the default probability immediately after the default of another issuer has occurred. In order to identify the systematic risk factors, the financial quantities on the credit market need to be accurately measured. In Chapter 9, we identified that default probabilities and recovery rates both were needed to price securities that are subject to credit risk. In Chapter 10, we defined hazard rates, which are more suitable to model since they contain more local information than default probabilities. In section 10.2 we derived expressions for valuing both straight bonds and floating rate notes, which will be used for formulating the inverse problem for measuring the term structures. The pricing equations (10.31), (10.32) and (10.33) contain expectations of forward-, hazard- and recovery-rates. As seen in 9.3, these have been proved to be dependent, at least under the physical measure. To be able to measure these curves separately, we need to assume independence.

As seen in the previous section, the defaultable discount factors,

\[ \exp \left( -\int_t^T (r_s + \lambda_s)ds \right), \]

(11.1)
contains risk-free rates and hazard rates additively. This makes accurate measurement of the risk-free term structure crucial to disentangle the hazard rates from the risk-free rates. After the rapid expansion of the CDS market prior to the global financial crisis, it has been more common to use CDS data to measure hazard and recovery rates. A benefit of this choice is that the cash flows of the instruments align with the cash flows from the interbank instruments. This reduces the importance of accurately measured interest rates. For bonds, however, cash flows may be encountered at any bank day. Another benefit of using CDS is the liquidity premium that is substantially higher in the bond market (Tang and Yan, 2006). For these reasons, the CDS market is a popular choice when measuring hazard and recovery rates. A major drawback of using CDS, however, is the presence of counterparty credit risk. Even if the CDS contracts are fully collateralized, or over-collateralized, as suggested by Arora et al. (2012), the jump-to-default is an unforeseeable event that typically cannot be fully collateralized. The contagion risk also creates a systemic risk that may affect the counterparty in case of the default of the underlying, which is the very situation the CDS is designed to cover.

11.1 Disentangling Recovery Rates from Hazard Rates

Early literature on the reduced-form model focused heavily on timing risk, whereas recovery risk received far less attention. Regardless of the choice of bond or CDS-market data, disentangling recovery rates from hazard rates has proved to be a complex task. Many early models used a constant recovery rate to measure the hazard rate (Elkamhi et al., 2014). Christensen (2005) use a four-factor model calibrated on bond data to measure hazard and recovery rates. He concludes that the noise present in bond prices complicates the separation of recovery rates from hazard rates. He also suggests that the separation is successful if bond yield movements can be decomposed into movements in hazard and recovery rates. In a subsequent paper, Christensen (2007) investigates CDS-data for Ford Motor Company and finds that it indeed is possible to disentangle recovery rates from hazard rates. Other studies of the disentanglement include (Pan and Singleton, 2008; Das and Hanouna, 2009; Sarbu et al., 2013; Le, 2015; Liu, 2017) for the CDS market, and (Bakshi et al., 2001) for the bond market. Given that we have access to an accurate method for measuring risk-free term structures of forward
Pricing Constraints for Defaultable Bonds

To simplify the valuation formulas for the defaultable bonds, we define the risk-free zero-coupon bond price,

\[ P(t, T) = \mathbb{E}^Q \left[ e^{-\int_t^T r_u du} \mid \mathcal{G}_t \right] = e^{-\int_t^T f(u) du}, \]  

(11.2)

and the survival function

\[ S(t, T) = \mathbb{Q}(\tau > T \mid \mathcal{G}_t) = \mathbb{1}_{\{\tau > t\}} \mathbb{E}^Q \left[ e^{-\int_t^T \gamma_u du} \mid \mathcal{G}_t \right] \equiv \mathbb{1}_{\{\tau > t\}} e^{-\int_t^T \lambda(u) du}. \]  

(11.3)

By assuming independence, the defaultable coupon-bond price can be written as

\[ \tilde{P}_{\text{STR}}(t, T) = \mathbb{1}_{\{\tau > t\}} \sum_{i=1}^{n} \frac{C_i}{m} P(t, T_n) S(t, T_n) \]

\[ + \mathbb{1}_{\{\tau > t\}} \int_t^T P(t, u) S(t, u) \lambda(u) \delta_u^Q Z_u du \]

\[ + \mathbb{1}_{\{\tau > t\}} P(t, T) S(t, T) X = g_{\text{STR}}(f, \lambda, \delta), \]  

(11.4)

where \( Z_u \) is the value to make a fractional recovery upon at time \( u \). Similarly, the defaultable floating rate note price can be written as

\[ \tilde{P}_{\text{FRN}}(t, T) = \mathbb{1}_{\{\tau > t\}} \sum_{i=1}^{n} P(t, T_n) S(t, T_n) \left( \mathbb{E}^T [L(T_i, T_j) \mid \mathcal{G}_t] + D \right) \Delta_{i,j} \]

\[ + \mathbb{1}_{\{\tau > t\}} \int_t^T P(t, u) S(t, u) Z_u \lambda(u) \delta_u^Q Z_u du \]

\[ + \mathbb{1}_{\{\tau > t\}} P(t, T) S(t, T) X = g_{\text{FRN}}(f, \lambda, \delta). \]  

(11.5)

Equations (11.4) and (11.5) represent the valuing constraints for the optimization problem presented in (11.6) below.

11.3 Measuring Hazard and Recovery Rates

As seen in the pricing equations (11.4) and (11.5), the risk-free rate is needed to price defaultable coupon bonds. For defaultable floating rate notes in
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

(11.5), the risk-free curves of matching tenors are needed to calculate a proxy for the expected LIBOR, $E^T [L(T_i, T_j)|G_t]$. These are measured using the framework described in (4.11), from section 4.2. This optimization model constitutes the first stage of the model, and the resulting curves serve as input to the second stage, where hazard and recovery rates are measured.

11.3.1 Measuring Default Intensity

When measuring default intensity, instruments are aggregated by credit rating. As seen in section 9.4, aggregating instruments by credit ratings is not without problems since credit ratings have proven to be inaccurate proxies for default probabilities historically. For bonds, the accuracy has been better than for structured products, but there are still discrepancies in historical default probabilities among bonds of different issuer types.

11.3.2 Measuring Recovery Rates

As seen in section 9.3, Altman and Kishore (1996) show that seniority is an important determinant for recovery. They further show that the issuers’ sector affects recovery. This led us to a model where each seniority receives a unique term structure of recovery rate. In addition, each sector is modeled as an offset, constant across maturities, that is added to the recovery term structure. This offset is added with enough penalty to center these shifts around zero. Despite only measuring one term structure per seniority and one offset per sector, the number of curves used for discounting will be the cartesian product between the sets of these seniorities and sectors. A constraint ensures that all recovery rates are bounded within $[0, 1]$.

11.3.3 Setting up the Optimization Model

Let $c \in C$ denote the set of credit risk classes, i.e. credit ratings and $s \in S$ denote the set of seniority classes. Let $i \in I$ denote industry sector, and define a recovery offset for each sector, $\Delta \delta_i$, $\forall i \in I$. We denote the difference between the hazard rate curve for a credit class and its nearest lower-grade credit class by $\Delta \lambda$. The continuous optimization problem for the second
Measuring Hazard and Recovery Rates

stage is formulated as

\[
\begin{align*}
\min_{\lambda, \delta, \Delta \delta, z_b, z_e} & \quad \frac{1}{2} z_e^T E_e z_e + \frac{1}{2} z_b^T E_b z_b + C \Delta \delta \sum_{o \in \mathcal{O}} \Delta \delta_o^2 + \\
& \quad \sum_{c \in \mathcal{C}} h_c(\lambda_c) + \sum_{c \in \mathcal{C} \setminus c_{\min}} h_\lambda(\Delta \lambda_c) + \sum_{s \in \mathcal{S}} h_s(\delta_s) \\
\text{s.t.} & \quad P_l \leq g_b(f, \lambda, \delta) + F_b z_b \leq P_u \\
& \quad g_e(f, \lambda, \delta) + F_e z_e = P_e \\
& \quad \Delta \lambda_c = \lambda_c - \lambda_{c-1}, \quad \forall c \in \mathcal{C} \setminus c_{\min} \\
& \quad \Delta \lambda_c \geq 0, \quad \forall c \in \mathcal{C} \setminus c_{\min} \\
& \quad 0 \leq \lambda_c \leq \lambda_{\max}, \quad \forall c \in \mathcal{C} \\
& \quad 0 \leq \delta_s + \Delta \delta_t \leq 1, \quad \forall s \in \mathcal{S}, o \in \mathcal{O}.
\end{align*}
\] (11.6a)

In this optimization problem, (11.6e) ensures that hazard rate curves for a lower grade credit class never cross curves for a higher grade class, where \( c_{\min} \) is the most creditworthy credit class\(^1\). (11.6f) ensures that the default intensity always stays positive. (11.6f) also poses an upper limit for the hazard rate which might have to be included to avoid highly unrealistic solutions\(^2\). (11.6g) simply ensures that the recovery for each seniority class \( s \) rate is bounded between zero and one.

\(^1\)Even though this is not always the case for the observed physical defaults (with relatively few observations), it should apply to the risk-neutral hazard rates assuming that credit ratings are accurate proxies for credit risk across all maturities. The rationale behind this is that it otherwise would be possible to construct a trade by entering one long position in a bond and one short position in another bond of slightly shorter maturity from the same issuer, leaving the investor exposed to \( \lambda_c(t) \), where \( c \) denotes the credit rating. By entering an identical trade with bonds from a worse rated issuer, with the same seniority and sector, a trader would otherwise be able to get exposure to \( \lambda_{\bar{c}}(t) \), where \( \bar{c} \) is a worse credit rating. Note that this does not constitute an arbitrage opportunity since any one of the two issuers may default.

\(^2\)Contrary to what one might intuitively believe, \( \lambda_c > 1 \) is not unrealistic since the (risk-neutral) default probability calculation involves the exponential function \( 1 - e^{-\int_t^T \lambda_u du} \).
11.3.4 Results

Together with the framework for measuring multiple risk-free term structures, (5), this model, (11.6), is used to make cross-sectional measurements of hazard rates and recovery rates. When repeating these measurements, a time series of term structures for different credit ratings and seniorities emerge. By using PCA, the systematic risk factors are extracted for each curve. Through a first-order Taylor approximation of the risk factors, we are also able to study the risk for both generic zero-coupon bonds of different maturities, as well as for market-quoted bonds. Through a performance attribution, we are also able to relate the empirically realized risks to the ex-ante simulations. By performing a regression analysis, we also show that the yield errors of the different bonds used to make the measurements are highly stable over time. This show that there are other bond and company-specific factors that affect bond prices, besides credit rating, seniority, and sector. Seeing this level of consistency opens up for the possibility to further study these factors. These results can be found in the appended Paper IV.
Contribution
Contribution

In this chapter, the research questions from Chapter 1 are answered alongside with an evaluation of the fulfillment of the purpose. The contribution of the thesis should be evaluated within the scope of the answers to the research questions and the assessment of the purpose.

12.1 Answers to the Research Questions

In this section, the research questions are answered. Answers to the research questions leading up to the main research question of each part are given first.

12.1.1 Answer to Research Question 1a

In order to answer the main research question for the interest rate part, regarding hedging of interest rate risk, two additional questions had to be answered. The first one was formulated in terms of measuring interest rate risk.

Research question 1a. *How can interest rate risk be accurately measured?*

Through the historical overview given in Chapter 7, it is clear that the historical risk measures have evolved from focusing solely on the exposure part of risk into taking the uncertainty part of risk into consideration as well. In order to handle the uncertainty part, the dynamics must be modeled
somehow. Value at Risk and expected shortfall are two related risk measures that focus on the tail with the largest portfolio losses. In some cases, the distribution function can be evaluated analytically, and in other cases, enough scenarios need to be generated to make an accurate representation of the tail of the distribution. Historical simulation is an alternative to Monte Carlo simulation. The advantage of historical simulation is that it captures all mathematical properties without requiring any explicit modeling. The drawback is that these properties are not stationary. In order to get enough scenarios, one may be forced to include scenarios from non-present market conditions. Another problem is that the number of historical scenarios might be insufficient.

To properly answer the research question, the term “accurately measured” need to be explained. In this context, “accurate” refers to the three properties identified in Paper I. First, there should be no systematic price errors caused by the term structure measurement method. Second, the distribution of the risk factors must be consistent with their ex-post realizations. Third, a lower risk is preferable. The second statement is most important when assessing accuracy. If the risk measure is not consistent with ex-post realizations, it cannot be called accurate. However, it is not a sufficient criterion since ex-post realizations are evaluated in terms of its own term structure measurement method. If this method is inaccurate in itself, the risk measure cannot be trusted. The third statement is a criterion for comparing methods that satisfy the first two, where a lower risk is preferable.

As elaborated in Chapter 7, Value at Risk is not a coherent risk measure, and regulations are moving towards the coherent measure expected shortfall, which in addition can be formulated as a linear programming problem. In Paper I, Hagenbjörk and Blomvall (2019) generate term structure scenarios by modeling the dynamics of each risk factor. As the risk factors are uncorrelated in sample, the non-stationary univariate distribution of each risk factor is modeled. On top of this, copulas are used to model joint variation. Since exposure is a prerequisite to measure risk, a portfolio of forward rate agreements is constructed. The results show that an accurate distribution of the portfolio value is obtained. The main contribution of this paper is related to this accurate representation, where accurate measurement of risk over short time frames is a special case. Hagenbjörk and Blomvall (2019) demonstrate that the risk is lower for all quantiles of Value at Risk using the Blomvall (2017) model compared to traditional term structure measurement models using the same modeling approach. This answers research question 1a.
12.1.2 Answer to Research Question 1b

In order to understand the realized portfolio changes and evaluate the performance of different hedging strategies, monetary performance attribution for the fixed income market was necessary to identify obstacles in the implementation.

Research question 1b. How should accurate monetary performance attribution be carried out in the fixed income market?

Historically, performance attribution has been carried out in terms of returns. The fact that returns cannot be aggregated across both assets and time severely complicates performance attribution in terms of returns. For some instruments, such as interest rate derivatives, the instrument values may fluctuate around zero. This renders performance attribution in terms of return meaningless. In Paper II, Blomvall and Hagenbjörk (2019) present a framework for monetary performance attribution with an illustration of the fixed income market. By performing a second-order Taylor approximation of the pricing function with respect to the systematic risk factors, the performance attribution can be made entirely residual free by introducing two error terms related to the Taylor approximation. The framework also supports foreign exchange rates and transaction costs. The contribution of this paper is, however, not limited to the results shown for the fixed income market. The framework is generic and can be used for any market where a suitable set of systematic risk factors is identified. However, as shown by Blomvall and Hagenbjörk (2019), accurate measurements of the systematic risk factors are essential to perform accurate performance attribution. This paper thus provides the answer to research question 1b, by defining the mathematical description of how risk factors contribute to portfolio earnings.

12.1.3 Answer to Research Question 1

The answers to research questions 1a and 1b are key components to answering the main research question of the interest rate part of this dissertation, which is formulated in terms of hedging.

Research question 1. How should a stochastic programming-based model be formulated to improve hedging decisions in the interest rate market?
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Neither the traditional hedging methods presented by Hagan and West (2006), nor the naïve back-to-back method takes transaction costs into account. In Paper III, Blomvall and Hagenbjörk (2019b) present a stochastic programming approach to hedging interest rate risk. The term structure measurement model by Blomvall (2017) is utilized to compute perceived transaction costs. The optimization model is formulated as a two-stage stochastic programming problem where transaction costs are weighed against portfolio variance. The risk will be hedged up until the point where the perceived transaction costs exceed the benefits of a reduction in variance, in which case the model instead will await a better opportunity later on. The hedging is carried out in an out-of-sample period between 2002 and 2018. We are able to show that any risk-averse investor would prefer to use the proposed stochastic programming model for hedging over a single-day time frame. This paper thus provides an answer to research question 1. However, the question of whether these results can be realized in practice is impossible to answer without incorporating it into real-world decision-making. The main contribution of paper III is, therefore, the demonstration of how decision-making can be improved when accurate models are available. As it turned out, the accurate model for the systematic risk factors was not necessary to obtain good hedging results in this setting where only variance is considered. However, it was necessary in order to rule out the possibility that the simple model caused instability.

12.1.4 Answer to Research Question 2a

Research question 2a. How can financial quantities in the credit market be accurately measured?

Credit risk mainly involves event risk, concerning the risk an issuer defaults over a certain time period and recovery risk, concerning how much can be recovered if a default occurs. The literature contains two main approaches to credit risk, where the structural approach is directly based on observing the financial statements of the issuer. The reduced-form approach does not have this direct link to the company but takes a statistical approach to the modeling of default probabilities. This link is instead provided through the calibration to market prices, which of course, reflects the financial situation of the issuer. The purpose of the study determines which approach to choose between these two. According to Jarrow and Protter (2004), reduced-form models should be preferred when pricing or hedging, which is the topic of this
dissertation. The reduced-form approach models term structures of default intensity (hazard rates), which by integration can be used for computing default probabilities in the same way as forward rates are used for discounting. This more local measure of default event risk is known as timing risk and implies modeling the event-risk over all possible time intervals. The recovery rate can be measured by assuming some recovery scheme and specifying the value that the bondholder is eligible to make a proportional recovery upon. Different recovery schemes have been proposed in the literature, often driven by mathematical tractability. The one that matches the legal aspects of the liquidation process is the recovery of face value that enables the bondholder to recover a proportion of the face value and any accrued interest. Considering that the majority of defaulting companies are reconstructed, which is an entirely different process, the modeling choice is not that simple. The literature finds different determinants of recovery rates. Empirically, Altman and Kishore (1996) find evidence that the seniority and sector are important determinants of recovery when studying 700 U.S. defaults. However, they find no support for credit rating affecting the recovery rate after adjusting for seniority. Schläfer and Uhlig-Homburg (2014) provide evidence that distress of the issuer’s sector and company-specific factors such as leverage ratio and liquidity affect recovery rates.

The early literature in the field of reduced-form modeling heavily focused on measuring the timing risk while ignoring the recovery risk. Disentangling the recovery risk from timing risk proved to be a difficult task. Disentanglement is usually attempted through assuming processes for the risk-free rates, hazard rates, and recovery rates that allow for various types of dependencies. After the rapid expansion of the CDS market in the early 2000s, CDS contracts became popular to use as they are more liquid due to their contractual nature and thereby contain less noise. The need for an accurate model of the risk-free term structure is also eliminated since the cash flow dates of CDS contracts align with those of the interest rate derivatives. On the other hand, CDS contracts are affected by counterparty default risk that is not fully taken into consideration on the market today (Jarrow, 2012; Arora et al., 2012). This does not result in any clear winner between bonds and CDS. The more recent literature has been focused on CDS data, but the CDS market size has severely diminished since the peak during 2008. This, plus the fact that we already have methods to accurately measure multiple risk-free interest rate term structures needed for valuing bonds, make bonds a suitable choice for our attempt to measure timing and recovery risk.

By assuming independence between risk-free rates, hazard rates, and recov-
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

... ery rates, under the risk-neutral measure, we are able to disentangle the effects of these three components on the bond prices for straights bonds and floating rate notes. By an extension to the Blomvall and Ndengo (2013b) framework for measuring multiple risk-free term structures, we are able to formulate a second optimization problem that uses the risk-free term structures as input to extract the hazard rates for different credit classes and recovery rates for different seniorities, where a constant shift depending on issuer sector also is taken into consideration. This enables joint measurements of the most local financial quantities that affect the pricing of defaultable bonds: forward rates, hazard rates, and recovery rates. However, as no consideration is taken to extract the liquidity premium or the co-variation between these terms, these effects will also be included in the measured hazard and recovery rates. As pointed out in the literature, the problem is highly sensitive to noise present in market data. Even though the measurements are stable over time, there are clear economical inconsistencies between the level and shape of recovery rates of different seniorities. At this point, no consideration is taken to company-specific data that has proven to be significant when estimating recovery rates. This is one step towards more accurate measurements of financial quantities in the credit market, which is the essence of research question 2a.

12.1.5 Answer to Research Question 2

Research question 2. Which data-driven systematic risk factors can be identified in the credit market?

Having accurate measurements of the underlying financial quantities is crucial for data-driven measurements of the systematic risk factors through principal component analysis (PCA). The problem of separating the timing and recovery risk complicates this matter as movements in both these term structures affect the prices in similar ways. As PCA is designed to capture variance, which is related to risk, the method is highly sensitive to noisy movements in the term structure resulting from noise in an ill-posed inverse problem. As a few unrealistic term structure movements may substantially affect the results of the PCA, one must make sure that the term structure innovations are economically realistic. This can be achieved through measuring the term structures using a well-posed inverse problem. Highly improbable term structure movements could also be eliminated by studying the probability of the movements.
Unfortunately, the inverse problem used to measure the term structures is non-convex. This manifested itself by converging to local optima with totally different levels of recovery rates during some of the days. Fortunately, the problem could be mitigated to a great extent. The effort spent to measure hazard rates and recovery rates over time enabled finding accurate measurements of these term structures. Few unrealistic principal components had to be eliminated in the end. Hence, we found that it indeed is possible to identify systematic risk factors for both the timing and recovery risk. We found that four components were needed to explain almost all of the in-sample variance in hazard rates across the 0-10 year range of maturity. These risk factors very much resemble the risk factors found for the risk-free interest rate market. The first two, “Shift” and “Twist”, explain the majority of the variance in the long end of the term structure while “Butterfly” and the 4:th risk factor primarily account for variance in the short end. How well these risk factors generalize to out-of-sample data is not studied, however. For the recovery rates, a single, completely flat, risk factor seems to be sufficient to explain the vast majority of the in-sample variance. The contribution of the identified “Twist” component is limited as it accounts for less than 1% of the in-sample variance. These findings are in line with the findings by Pan and Singleton (2008) in the CDS market.

These results could be used as motivation for how many risk factors that are needed to modeling the dynamics of the credit risk market through multifactor models using stochastic processes. We have identified a total number of 11 (6+4+1) latent systematic risk factors for the risk-free rate, hazard rate, and recovery rate. However, by reducing explanatory power for the short end of the term structures, we can remove two risk-factors for risk-free rates and one for hazard rates can be ignored. This leaves us with 8 (4+3+1) risk factors, but even this is more than what is commonly found in the literature. Christensen (2007) and Bühler and Trapp (2009) use four latent risk factors and Doshi et al. (2018) use five, but these seem to be the richest models in terms of the total number of risk factors.

12.2 Fulfillment of the Purpose

The purpose of this dissertation, stated in the introduction, is repeated below.

**Purpose.** *To develop accurate models that can be used to improve decision-making in the fixed income market.*
In order to address the fulfillment of this purpose, the terms “accurate models” and “improved decision-making” must be defined. A “model” refers to any mathematical description of a system. Such a model can be used for decision-making or as support for decision making, either directly or as a sub-component. The term “accurate” refers to the model being an accurate representation of the observed system. This property depends on the type of model but includes leaving no behavior unexplained, having no systematical errors and being consistent with ex-post realizations. To “improve” decision-making refers to the model’s ability to provide information that helps identifying better decisions according to some objective measure. The risk measures in Paper I and the performance attribution results in Paper II clearly provide information that helps identify better decisions. The hedging results in Paper III are themselves examples of improved decisions. The risk factors identified in Paper IV also provide information of the risk and how many risk factor that might be needed to model the yield movements in the bond market accurately.

All research presented in this dissertation is carried out in realistic settings that are aimed to reflect real-world conditions as accurately as possible. The data used is the same data available to financial practitioners. The simulated portfolios used are set up to mimic customer behavior or to reflect sensitive portfolio structures that may result from complex trading strategies. To properly function as decision systems or decision support systems, the time aspect becomes relevant. The decisions made in the fixed income market is still not automatized but carried out by contacts between human traders. This makes the speed of high-frequency trading systems seen in the equity market irrelevant. Even though some of the historical backtests performed within this dissertation are rather time-consuming, this is mainly the result of the decisions being repeated on a daily basis over a long history. The out-of-sample tests for Paper I-III range between the start of 2002 and the end of 2018, which consists of more than 4,000 trading days. To make a single decision, the models developed in these papers can be run in under a few seconds. This is a viable time since human traders still must process the information and interact with other traders. Since all research questions have been answered in such a realistic setting, the purpose is considered to be fulfilled.
Future Research

As demonstrated in Figure 1, the different parts build upon each other. Within the scope of the overall methodology used in the bottom-up approach described in Chapter 1, the theory is modularized. This means that different modules can be improved, reworked or replaced without affecting the structure of adjacent modules. The module forming the basis for all work is the optimization framework for measuring term structures. Even though this framework has laid the foundation for many rewarding theoretical results, the framework has not been overhauled its original form proposed in 2011. The work carried out since then has revealed several areas of improvement. Most importantly, no consideration is explicitly taken to information across time since the framework is designed to measure financial quantities using a cross-section of information available at a given time. In this dissertation, the shortcomings of this have been handled through basic forms of data cleaning. However, this only handles the worst cases of unrealistic data. All prices are observed with some measurement error, and the term structure will adapt to this noise if this uncertainty is not handled. Hagenbjörk and Blomvall (2019) show evidence of negative autocorrelation in risk factors, leading to overestimation of risk, especially for more extended time frames. This can be overcome by filtering the prices using a more sophisticated method, e.g. Kalman smoothing. Ideally, all term structures should be measured simultaneously by utilizing all available information. This way, regularization can be performed both across dates and maturities. Another area of improvement is the choice of parameters in the framework in order to reduce systematic price errors. Improving the accuracy of the measurements of financial quantities would help to improve the results across all interconnected modules.
Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets

Methods for generating long-term scenarios and thereby measuring risk for longer timeframes can be improved substantially. As mentioned above, more accurate terms structure measurements can be part of this solution. In addition, the modeling should be extended to capture properties that are relevant for extended time frames, such as mean-reversion of interest rates. For longer time frames, the evaluation should probably not be made in terms of a portfolio since this makes the risk path-dependent due to the fixing of interest rates for the coupons that end up in the portfolio. Instead, the quality of long-term scenarios could be evaluated in terms of quoted yields for instruments. The drawback of such an approach is that instead of evaluating scenarios in terms of one single portfolio that covers all maturities, the scenarios must be evaluated in terms of several different instruments whose yields correlate with the yields for instruments of shorter maturities.

The foundation of the theory for the monetary performance attribution framework presented by Blomvall and Hagenbjörk (2019) is solid. An area of improvement for this theory would be to determine how to link single-period attribution results in terms of returns within this framework. Another area of research would be to accurately measure financial quantities and its corresponding systematic risk factors for additional markets in order to enable performance attribution for these markets. This may also enable studies of how well the systematic risk factors from other financial quantities generalize out of sample.

A natural progression towards being able to solve investment problems in the interbank interest rate market would be to measure the risk premia for the systematic risk factors and study how the risk premia change over time.

For the credit market, the next step would be to further study performance attribution by investigating a bond portfolio and see how well our risk factors generalize out of sample. Compared to the risk-free interest rate market, the credit market would need several additional risk factors. For the model described in Paper IV, the systematic risk factors for hazard rates and recovery rates would be added to those of the risk-free interest rates. Since all term credit ratings, seniorities, and sectors are entitled to their respective term structures, and a carry term may also be derived from all term structures, the number of risk factors for a portfolio may be large. One must, therefore, find a meaningful way of presenting the performance attribution results.

Another natural extension to measuring credit risk would be to incorporate credit derivatives. This can be done either in combination with bonds
or separately. When mixing these types of instruments, the differences in liquidity and counterparty risk might be relevant to study further.

The modeling of the dynamics for the systematic risk factors for the term structure of hazard rates and ideally also for the recovery rates can also be studied further. This way, the risk in bond portfolios can be measured more accurately. The simulation of defaults across issuers can also be improved. To accurately simulate defaults, they should be correlated, and the contagion effect should be modeled. To be able to solve investment problems in the bond market properly, the risk premia for the systematic risk factors, as well as jump-to-default risk and contagion risk should be measured. It would be especially interesting to see how the risk premia of these risk factors have evolved as an effect of the quantitative easings and negative bond yields appearing in the aftermath of the global financial crisis 2008.
Bibliography


Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets


Blomvall, J. and Hagenbjörk, J. (2019b), Hedging of interest rate risk using a stochastic programming approach. unpublished.


Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets


Cursio, J. D. and Baek, S. (2016), The influence of credit ratings on capital structure. Available at SSRN 2552922.


Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets


Elizalde, A. (2005a), Credit risk models i: Default correlation in intensity models. unpublished.


122
BIBLIOGRAPHY


Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets


124


Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets


Optimization-Based Models for Measuring and Hedging Risk in Fixed Income Markets


United States (2010), ‘Dodd-frank wall street reform and consumer protection act: Conference report (to accompany h.r. 4173). washington: U.s. g.p.o’.


Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-162576
Nomenclature

A Integrating matrix to transform forward rates into spot rates.
AAA Highest rating on the S&P scale (“tripple-A”).
B Differentiating matrix to transform spot rates into forward rates.
Bond A securitized market-traded debt instrument.
bp Basis point = 0.01%.
C Annual coupon rate.
C Covariance matrix.
CDO Collateralized Debt Obligation.
CDS Credit Default Swap.
CIR Cox-Ingersoll-Ross.
CVaR Conditional Value at Risk.
D Credit spread.
δt Recovery process.
E Expected value.
E Penalty matrix for market price deviation.
Ek Matrix consisting of the k most significant eigenvectors.
EQU Expected value under the risk neutral measure.
εt,A Attribution error stemming from omission of higher-order Taylor terms and taking non-infinitesimal time steps.
εt,l Attribution error stemming from truncation of insignificant risk factors.
ES Expected Shortfall.
Identification of the Credit-Risk Related Systematic Risk Factors from Bond Prices

\( f \)  Instantaneous forward rate.
\( f_0 \)  Risk-free term structure.
\( F \)  Scaling matrix for pricing constraints.
\( F_t \)  The filtration containing the financial information from market prices.
FRA  Forward Rate Agreement.
FRN  Floating rate note paying variable coupons tied to a benchmark rate.
\( G \)  Gearing.
\( g \)  Risk-factor gradient.
GAAP  Generally Accepted Accounting Principles.
\( \gamma_t \)  Hazard process.
\( \mathcal{G}_t \)  The enlarged filtration containing all information, \( \mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t \).
\( H \)  Risk-factor hessian.
\( \mathcal{H}_t \)  The filtration containing the default information.
IFRS  International Financial Reporting Standards.
IRS  Interest Rate Swap.
\( \lambda \)  Hazard rate.
\( \Lambda_k \)  Matrix consisting of the \( k \) largest eigenvalues.
LIBOR  London Interbank Offered Rate.
\( \mu \)  Expected value, often expected return.
\( N \)  Nominal amount, face value.
NPV  Net Present Value.
OIS  Overnight Indexed Swap.
\( \Omega \)  Probability space.
OTC  Over the counter.
PCA  Principal Component Analysis.
\( \hat{P} \)  Quoted price.
\( P(t, T) \)  Price of a risk-free zero-coupon bond with maturity \( T \) at time \( t \)
NOMENCLATURE

\( \tilde{P}(t, T) \)  Price of a defaultable bond with maturity \( T \) at time \( t \)

\( \mathbb{Q} \)  Risk-neutral (martingale) probability measure

\( \mathbb{QQ} \)  Quantile-Quantile.

\( r \)  Continuously compounded interest rate.

\( RFV \)  Recovery of Face Value.

\( RMV \)  Recovery of Market Value.

\( RT \)  Recovery of Treasury.

\( r_t \)  Short rate.

\( S(t) \)  Survival function.

\( SEK \)  Swedish krona.

\( \sigma \)  Volatility.

\( SP \)  Stochastic Programming.

\( SQL \)  Structured Query Language.

\( STR \)  Straight bond paying fixed coupons.

\( \tau \)  Time of default or instrument tenor.

\( \theta_{t-1} \Delta_t \)  Carry earned from holding a position when term structures remain unchanged.

\( \top \)  Transpose.

\( USD \)  United States dollar.

\( VaR \)  Value at Risk.

\( Var \)  Variance.

\( X \)  Risky claim.

\( \xi \)  Risk factors.

\( \Delta \xi \)  Risk factor innovations.

\( y \)  Yield to maturity.

\( z \)  Penalized pricing errors.