Norwegian School of Economics
Bergen, Fall 2023

NHH

Expanding Horizons in Norwegian High-Yield Bond Spread Estimation

A Comparative Analysis of Structural Bond Pricing Models in the Period 2015 to 2023

Arne Hebnes & Marius Larsen Bang

Supervisor: Petter Bjerksund

Master Thesis in Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

ABSTRACT

Investigating the complex dynamics of credit risk pricing in corporate bonds has been a significant area of interest in financial research for many years. This thesis examines the effectiveness of various structural bond pricing models in predicting credit risk premiums for Norwegian high-yield bonds from 2015 to 2023. Building on the approach by Eom et al. (2004), we compare the predictive accuracy of the extended Merton model with that of the Longstaff & Schwartz (1995) and Leland & Toft (1996) models. Our findings indicate that the extended Merton model tends to underpredict spreads, consistent with prior studies. Additionally, the Leland & Toft model and Longstaff & Schwartz model underpredict spreads, a deviation from existing research. Among these, the Longstaff & Schwartz model emerges as the most precise in predicting spreads for Norwegian high-yield bonds, suggesting the utility of diverse structural bond pricing models beyond the commonly used Merton model. This research, in line with prior empirical studies, highlights a substantial deviation of modeled spreads from actual observed credit spreads, known as the "credit spread puzzle." The thesis attempts to explain the key drivers of this puzzle within the extended Merton model by examining a range of explanatory variables that serve as proxies for liquidity premium and risk premium. Contrary to previous research, our results suggest that the extended Merton model may inherently incorporate a liquidity premium, leaving the credit spread puzzle exclusively to risk premiums. Moreover, we find that the stock price momentum premium is significant in determining credit spreads, whereas neither the market factor size nor the value factor is. We find that the momentum variable may incorporate the effects typically associated with size and that the effect of value is captured by the capital distribution factor. Alongside comparable research, we also discover that high-yield bond investors demand a risk premium for being invested in the sectors of Oil and Gas E&P, Oil and Gas Services, and Shipping. Through this research, it becomes evident that accurately pricing high-yield bonds is a complex endeavor. Particularly challenging is isolating the true essence of the credit spread puzzle, given the potential biases present in both the extended Merton model and the benchmark market spreads.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGEMENTS

The completion of this research was made possible through the support and contributions of several key individuals and organizations. We would like to thank our supervisor, Petter Bjerksund, at the Norwegian School of Economics, whose guidance has been invaluable throughout these past couple of months. We are also thankful to Stamdata for granting us access to their comprehensive database, which has been crucial in collecting data on Norwegian high-yield bonds. Our appreciation goes to Nordic Bond Pricing as well, for supplying us with daily market prices essential to our sample analysis. The examination into structural bond pricing models and the complexity of the credit spread puzzle would not have been feasible without their valuable data. Lastly, we would like to thank Infront for providing access to their terminal, which streamlined our data collection process.

1. INTRODUCTION

1.1 Structural Bond Pricing

Structural bond pricing models are models that intertwine corporate finance with bond pricing. The main idea of structural models is to combine market information with a model where equity and debt are considered as derivatives, to price credit risk. The models presented by Black & Scholes (1973) and Merton (1974) are two examples of structural models. Multiple structural models have later been published in academia, each with different features for debt valuation. Geske (1977) uses a compound option technique to value a company's risky coupon bonds. Longstaff & Schwartz (1995) model both default risk and interest rate risk when valuing corporate debt. Leland & Toft (1996) price bonds by considering the firm's capital structure in an environment with apparent corporate tax benefits and bankruptcy costs. Collin-Dufresne & Goldstein (2001) is an extension of the Longstaff & Schwartz (1995) model where a constant leverage ratio is presumed, and where the firm is only allowed to deviate from its target leverage ratio over the short run. These models are just a few of many structural models developed after the initiation of Black & Scholes (1973).

A crucial and much-studied input variable in structural bond pricing models is the default probability of the bond issuer. Cremers et al. (2008), Chen et al. (2009), Zhang et al. (2009), Bhamra et al. (2009), Huang & Huang (2012), Chen et al. (2018), and Bai (2021), all calibrate structural models by using the historical default rate for different ratings and maturities. However, later research by Feldhütter & Schaefer (2018) find this approach statistically weak due to low levels of default frequency in the sample size, and that even with larger samples the firms are correlated, which mitigates the effect. They propose a different solution to calibrating default rates, extracting the default boundary by minimizing the difference between modeled and observed default rates in the market across maturities and ratings. They find high predictive power for the average modeled spreads, however only for investment grade bonds. Acknowledging the limitations and advancements in calibrating default rates, our analysis pivots towards structural bond pricing models that offer a more comprehensive assessment of credit spreads

The study that directed our interest in structural bond pricing models for comparative analysis, is the work of Eom et al. (2004). Their research employs different structural models to examine the spreads of corporate bonds issued by U.S. firms from 1986 to 1997. This approach does not calibrate historic default rates for the models. However, it establishes a framework for structural bond model comparison. As there exists no empirical research on structural bond pricing model comparison on Norwegian high-yield bonds, we use the approach of Eom et al. (2004) as a benchmark to contrast the outcomes of their pricing models with those obtained in this thesis. Eom et al. (2004) present an extended Merton, which we decided to employ as the main structural bond pricing model in this paper. This model is detailed in Chapter 3 and operates on a discretely shaped yield curve. In addition, we add the Leland & Toft (1996) model and the Longstaff & Schwartz (1995) model as comparative models for bond pricing predictions. Contradicting the extended Merton model, the Leland & Toft model assumes constant interest rates, and the Longstaff & Schwartz model assumes a stochastically mean-reverting yield curve, integrating the Vasicek (1977) approach.

These models provide us with a range of bond prices. While bond price comparison is interesting, we also want to address the spreads of the bonds. The spreads are derived as the difference between the modeled bond price and a comparable risk-free bond, adjusted for the maturity of the bonds. The spread tells us about the premium that is attached to the bond when compared to a risk-free bond. Due to the way we model the spreads, with risk-free bonds derived from a discretely shaped yield curve for the extended Merton model and the Leland & Toft model, and the Vasicek yield curve for the Longstaff & Schwartz model, we can directly compare them to the article of Eom et al. (2004).

By implementing the three structural bond pricing models on a sample of Norwegian high-yield bonds, we investigated the following research question for the thesis:

How does the extended Merton model compare to other structural bond pricing models on Norwegian-issued high-yield bonds?

We find the Longstaff & Schwartz model to be the model that achieves the lowest spread prediction error relative to spreads observed in the market. Moreover, it yields more precise spread predictions than those reported in our chosen benchmark study by Eom et al. (2004). The superior performance of the Longstaff & Schwartz model, especially when compared with the extended Merton model and the Leland & Toft model, underscores the viability of alternative modeling approaches for spread prediction. This finding substantiates the idea that models beyond the traditionally favored Merton model can also effectively predict reliable and accurate spreads.

1.2 The Credit Spread Puzzle

In recent times much research has been devoted to explaining the credit spread puzzle, which is the phenomenon of structural bond pricing models' tendency to predict inaccurate corporate bond spreads. Elton et al. (2001), Eom et al. (2004), Bao et al. (2011), Sæbø (2015), among others, propose different factors to define what might cause this puzzle. All of them find significant factors that explain the models' inability to accurately predict spreads. Later research by Feldhütter & Schaefer (2018) found no evidence of a credit spread puzzle when calibrating default rates with a minimization technique of the difference between the model and observed default rates. However, this technique still finds spreads that are too low on average for high-yield bonds, suggesting that there still is a credit spread puzzle for this category of bonds. Seeing as the estimation of default rates is time-consuming, technically challenging, and has yet to disprove the existence of a credit spread puzzle for high-yield bonds, we use the endogenously estimated default rate from the extended Merton model without further calibration. This allows us to focus our efforts on analyzing the consequence of different factors' effects on the extended Merton model. The second part of this thesis focuses on explaining the credit spread puzzle exclusively within the extended Merton model. Given that existing research predominantly concentrates on the Merton model for this purpose, this approach facilitates a direct comparison of our findings with established studies in this field.

By looking at the difference between actual spreads and modeled spreads using the extended Merton model, we examine the second research question:

What are the key determinants of high-yield bond mispricing in Norway, as predicted by the extended Merton model?

Our research confirms the existence of the credit spread puzzle, though its explanation remains complex. Diverging from prior studies, we find no evidence that a liquidity premium accounts for the puzzle. Instead, we discover that the momentum of stock prices significantly contributes to explaining it. Additionally, we observe that investors require a risk premium for bonds in specific industries, a factor not sufficiently incorporated by the extended Merton model. Contrary to earlier research, market risk premiums associated with the size and value of the issuing firms are not significant in our analysis. Chapter 6 discusses how certain assumptions in the extended Merton model may limit its spread prediction accuracy. This limitation could prevent the model from fully accounting for variables that are influential in actual market spreads.

1.3 Our Contribution To Existing Research

To address our research questions effectively, we need actual observed spreads to both test the accuracy of each implemented model and to analyze the impact of various explanatory variables on mispricing. Our methodology for obtaining actual market spreads diverges from previous research by utilizing quoted prices from Nordic Bond Pricing. This access is crucial, as it helps us avoid the potential bias introduced when relying solely on prices of publicly traded bonds, which has been common in earlier studies on structural bond pricing models. The comprehensive data from Nordic Bond Pricing allows us to mitigate sample selection bias more effectively, providing a stronger foundation for our analysis. This data enhances our capacity to thoroughly examine the predictive performance of the three structural models and to explore the nuances of the credit spread puzzle within the Norwegian high-yield market.

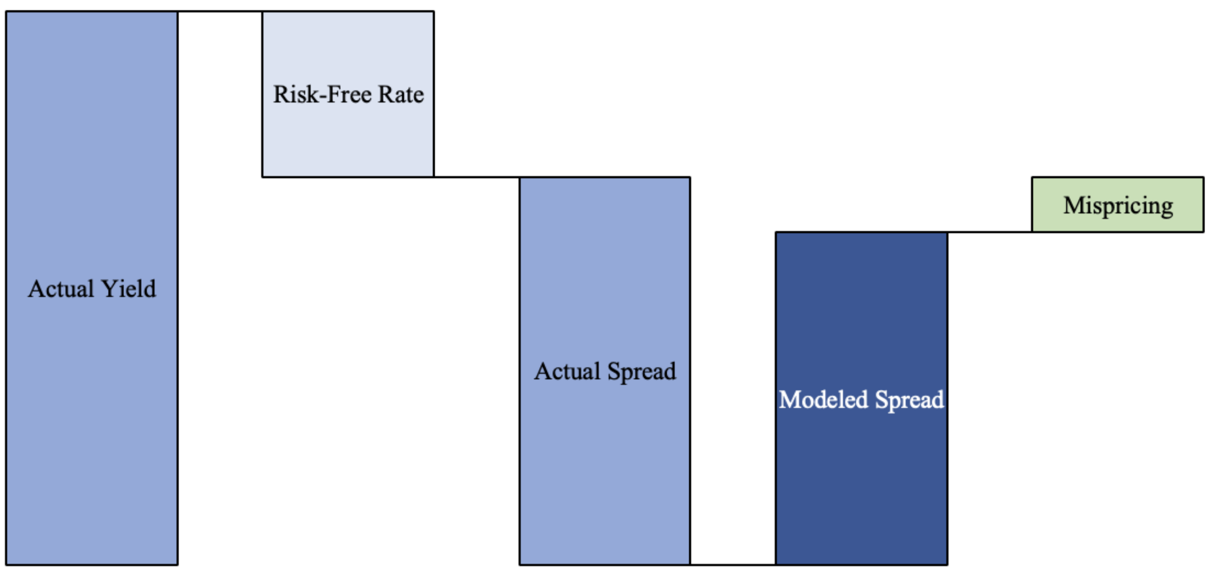
Building on the work of Ytterdal & Knappskog (2015) and Langdalen & Johansen (2016), our thesis aims to push the boundaries of current understanding. Ytterdal & Knappskog's study delves into spread predictions and the credit spread puzzle in Norway, utilizing a structural model and annual spread data from 2000 to 2012. Langdalen & Johansen expanded on this by applying an augmented Merton model to the Norwegian bond market. However, both studies had constraints: Ytterdal & Knappskog worked with limited sample size and estimated spreads only at bond issuance, resulting in 314 observations. Langdalen & Johansen use time series data on corporate bonds but rely exclusively on publicly traded bonds. Our research broadens their scope in several ways. We not only include a more comprehensive set of bonds, both listed and non-listed but also enhance the analytical framework by incorporating two-way fixed effects in our regression analysis, allowing for a more refined examination of the factors affecting credit spreads.

1.4 Analytical Framework

To address the first research question, we closely follow the framework of Eom et al. (2004) and implement different structural bond pricing models, and compare them to the extended Merton model based on their accuracy in estimating spreads. First, we visually compare the mean spreads on every date, with the benchmark spreads from Nordic Bond Pricing and compute summary statistics for the models. Second, we further use the analytical framework of Eom et al. (2004) by comparing the models based on four measures of pricing and spread percentage errors. We also compute the root mean squared error to address extremities in modeled spreads. This approach ensures a full comparison of the pricing and spread predicting abilities of the models and establishes a benchmark for further research on structural bond model spread comparison in the Norwegian high-yield market.

In addressing the second research question, we create a mispricing variable by subtracting the modeled spreads predicted by the extended Merton model from the actual observed credit spreads. To get to the mispricing variable, we follow a simple waterfall analogy.

Figure 1: Waterfall diagram displaying how the mispricing variable is calculated.



We then run both a two-way fixed effects regression and regular Ordinary Least Squares (OLS) multiple regression, using the mispricing variable as the dependent variable, and incorporating various explanatory variables to determine the key drivers of mispricing. We develop several regression models, each exclusively controlling for different categories of independent variables at a time. Finally, both regression tables are concluded with two models that integrate all the explanatory variables from the analysis. The regression analysis seeks to serve two main objectives: to identify whether there is a credit spread puzzle by looking at the explanatory power of the unbiased model and also to identify what contributes to the credit spread puzzle by looking at previously tested and untested variables.

1.5 Structure Of The Paper

This thesis is organized into several chapters, each focusing on a distinct aspect of bond theory and its practical application in the Norwegian high-yield bond market. Chapter 2 provides a basic understanding of bond theory, while Chapter 3 explains the chosen research methodology. In Chapter 4, we explore the data and discuss the characteristics of the final sample. Chapter 5 contains the results from the comparison of the extended Merton model with the chosen structural bond pricing models. Chapter 6 is an analysis of the regression on the mispricing variable based on the extended Merton model, and examines the credit spread puzzle and its containments. Chapter 7 is the conclusion, Chapter 8 is the criticism of the thesis, Chapter 9 is the appendix, and Chapter 10 is the references.

2. THEORY

2.1 The Fundamentals of Bonds

A bond is a security containing debt where the issuer loans money from an investor (the bondholder). The issuer is usually an entity that is either a government, a sub-government (municipality, state, or another form of juridical entity connected to the government), or a corporation. A bond can be issued with or without a predetermined payback day, called maturity date. However, bonds with a predetermined maturity date are the most common ones. The bond issuer makes periodical payments (coupon payments or interest payments) to the bond investor. These payments are usually done periodically and in line with each specific bond agreement (bond covenant), or as a full payment at the maturity date. An initial bond issuance usually happens through a primary market, which only specific investors have access to. After first being traded on a primary market, a bond can then be bought or sold on a secondary market.

2.1.1 Bonds in the priority hierarchy

Companies can choose from different types of financing when in need of additional capital. Investors deploying capital into a business expect to be compensated, but some investors are being prioritized above others. Debt investors (creditors) are placed higher in the priority hierarchy than equity investors, implying that interests and coupons are being paid off before a potential dividend or share repurchase can be initiated. Additionally, in the event of bankruptcy, debt is always paid off before equity.

2.1.2 Different bond types

Bonds are usually classified due to differences in maturity, convertibility, and the structure of return. Bonds are issued with a wide range of maturities; shorter than one year, longer than one year, or with no maturity at all. Bonds with a maturity of less than a year are called a certificate, while bonds with longer maturities are called bonds, and bonds without a maturity date are called perpetual bonds. The structure of interest payments can be either fixed interest rates or floating interest rates. Fixed interest payments imply that interest payments are the same size throughout the payment schedule. Floating interest payments means that the interest rate usually is linked to a benchmark interest rate like interbank rates such as the London Interbank Offered Rate (LIBOR), the Norwegian Interbank Offered Rate (NIBOR), or the Stockholm Interbank Offered Rate (STIBOR).

Bonds can also be issued with an embedded option, such as convertible bonds, callable bonds, and puttable bonds. Bond investors holding such bonds have the option to be repaid through a different settlement method instead of the regular principal repayment. Investors of convertible bonds can decide to receive payment in either the agreed-upon principal amount or an amount in equity shares. The potential upside gain in such bonds is larger than for regular bonds, leading these bonds to pay less interest over the bond's lifetime. A callable bond can be "called" back by its issuer before the expiry date, whereas a holder of a puttable bond can force the bond issuer to buy back the bond before the maturity date.

2.1.3 Credit rating

A credit rating is an evaluation of an issuer's creditworthiness or a specific issue of debt. Bonds are usually rated through rating agencies such as Moody's, S&P or Fitch. However, some bonds are also issued with something called a "shadow rating", which is a rating set by an intermediary bank. However, this rating is not stated to the public. The credit assessment done by either the rating agencies or the intermediary bank is based on an issuer's financial strength, or its ability to pay the debt's interest and principal at the right time.

Credit ratings are assessed on a scale from AAA to D (S&P and Fitch), or from Aaa to D (Moody's) depending on the relative measure of riskiness for the issuer or the issue itself. Although AAA and Aaa are the best possible ratings, these do not imply that the issuer is immune to default. Ratings imply that the probability of default is lower for a firm with a higher credit rating. As we will present in the next subchapter, the probability of default is an important part of the credit spread calculation.

As exhibited by Table 1, bonds are classified as either investment grade (often referred to as IG) or high-yield (frequently referred to as HY) based on the implied riskiness of the issued debt. Investment grade bonds are rated BBB- (S&P and Fitch) or Baa3 (Moody's), or higher. Such bonds have a relatively low probability of default, which usually leads to a lower credit spread. Bonds classified as investment grade are usually issued by governments, sub-governments, or corporations with a high degree of creditworthiness, hence making these low-risk investments. High-yield bonds (often referred to as “junk bonds”) are rated BB+ (S&P and Fitch) or Ba1 (Moody’s), or lower. The probability of default for these bonds is typically higher than for investment-grade bonds, accordingly, putting them in the category of being high-risk bonds. High-yield bonds usually generate a higher credit spread as we will demonstrate later.

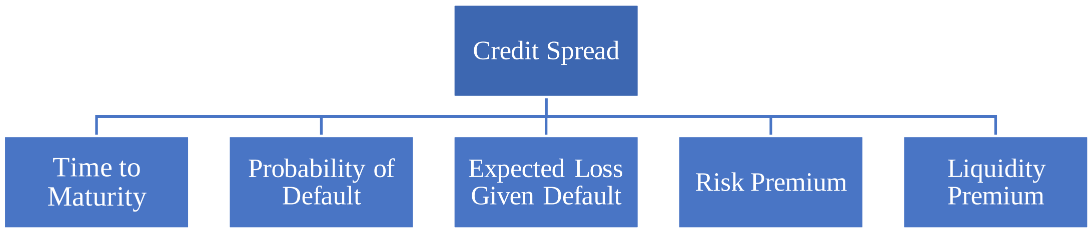
Table 1: Credit rating across rating agencies.

Risk class	Moody's	S&P/Fitch	Rating
INVESTMENT GRADE	Aaa	AAA	Best quality.
	Aa1 Aa2 Aa3	AA+ AA AA-	Strong ability for timely payments.
	A1 A2 A3	A+ A A-	Somewhat more exposed for negative changes.
	Baa1 Baa2 Baa3	BBB+ BBB BBB-	Adequate ability to meet payments. Some elements of protection missing.
SPECULATIVE GRADE / HIGH-YIELD	Ba1 Ba2 Ba3	BB+ BB BB-	Speculative risk. Future not well secured.
	B1 B2 B3	B+ B B-	Timely payment at the moment. Very exposed to negative changes.
	Caa1...	CCC+...	Default a likely option.
	Ca-C	CC-C	
	D	D	Default has occurred.

2.2 Credit Spread

Credit spread is the extra compensation an investor demands for investing in a bond with some degree of risk. The credit spread (in percentage terms) of a bond is determined by subtracting the yield-to-maturity (YTM) of a relevant government bond from the yield-to-maturity of the risky bond. This assumes that the government bond and the risky bond have identical characteristics except for the credit risk. The government bond then serves as a proxy for a risk-free rate, making the credit spread equal to the market risk premium of the bond. Andreassen & Semmen (2023) break down the credit spread in Figure 2, displaying that the credit spread consists of the bond's time to maturity, probability of default, expected loss given default, risk premium, and liquidity premium.

Figure 2: Break-down of the credit spread.



Source: Andreassen & Semmen (2023)

Bond pricing models usually calculate credit spreads using the three first variables reading from left to right of Figure 2, where the calculation of risk premium and liquidity premium is a more challenging endeavor. The credit spread of a risky bond can be determined by using the following formula from Feldhütter & Schaefer (2018):

$$Credit\ Spread = y - r = -\frac{1}{T}\log[1 - (1 - R)\pi^{Q}(T)] \tag{2.1}$$

where y is the YTM of a risky bond, r is the YTM of a government bond, R is the recovery rate, T is the bond's time to maturity, and $\pi^{Q}(T)$ is the risk-neutral probability of default. The formula illustrates the anticipated loss for a debt security investor in the event of default. It integrates both the likelihood of failing to meet contractual payment obligations and the recovery rate for the outstanding loan amount. The anticipated loss at the point of default should align with the investor's compensation for assuming credit risk. IG bonds are expected to have a lower expected default loss than HY bonds given that HY bonds presumably have a higher cumulative probability of default and a higher loss given default.

2.2.1 Risk premium and liquidity premium

Bonds are subject to credit risk because their future cash flows are uncertain. Equation 2.1 provides a method to assess this risk, yet it may overlook certain risk components. As depicted in Figure 2, risk and liquidity premiums are vital to a bond's credit spread, but these may not be fully captured by the equation. Since these premiums are part of actual market spreads, they should be included in bond pricing models for precise predictions. In Chapter 6, our regression analyses investigate how these market and liquidity premiums affect the disparity between actual market spreads and the model's predictions.

2.3 Option Pricing and Corporate Finance

Before delving into the bond pricing models utilized in this study, we first present the foundational concepts underlying these models, which are rooted in the Black-Scholes framework. The Black-Scholes model was originally devised for options pricing but has since its publication been adapted for various applications in corporate finance. In the following sections, we will discuss how the principles of option pricing are applied to concepts of corporate finance.

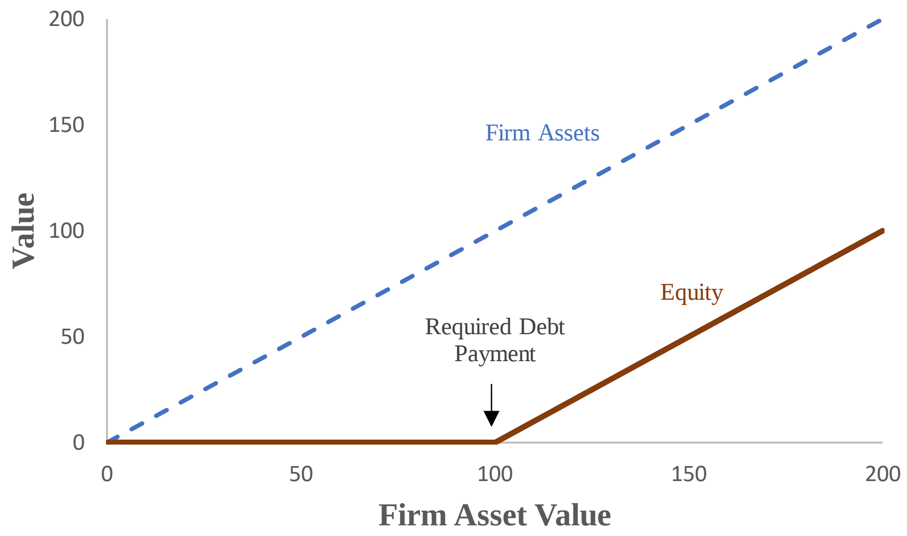
The value of a European call option at expiration is equal to:

$$C = max(S - K, 0) \tag{2.2}$$

where C is the value of the option, S is the stock price at expiration, and K is the exercise price of the option. The call option will only hold value if the stock price is higher than the exercise price at the expiration date. A stock can therefore be seen as a call option on the assets of a firm, with the value of debt outstanding as the exercise price.

The same idea can be used to understand the value of corporate equity in general. Figure 3 exhibits the relationship between a firm's equity value and the required debt payment. If the asset value is below the required debt payment at the debt's redemption date, then the firm will go bankrupt, and the equity will be worthless. However, if the asset value is above the required debt payment at maturity, then the equity holders will receive the value that remains after the debt has been paid.

Figure 3: Equity as a Call Option.



Source: Berk & DeMarzo (2019).

The idea of option pricing can also be used to value debt. When the equity holders are seen as holding a call option of the firm's assets, debt holders sit on the other side of the table owning the firm and having sold a call option with an exercise price that is equal to the required debt payment. If the value of the firm's assets is higher than the required debt payment at maturity, then the debt holders will receive all outstanding debt (equity holders will trigger the exercise price), and "give up" the firm. On the other side, if the asset value is below the required debt payment at maturity, then the debt holders will not receive the required debt payment (the equity holders will not trigger the exercise price) but will be entitled to what is left of the firm's assets. Figure 4 illustrates the relationship between the value of a firm's assets and debt.

Corporate debt can also be valued as a portfolio of risk-free debt and a short position in a put option on the firm's assets with an exercise price equal to the required debt payment. This portfolio valuation is presented in the following equation:

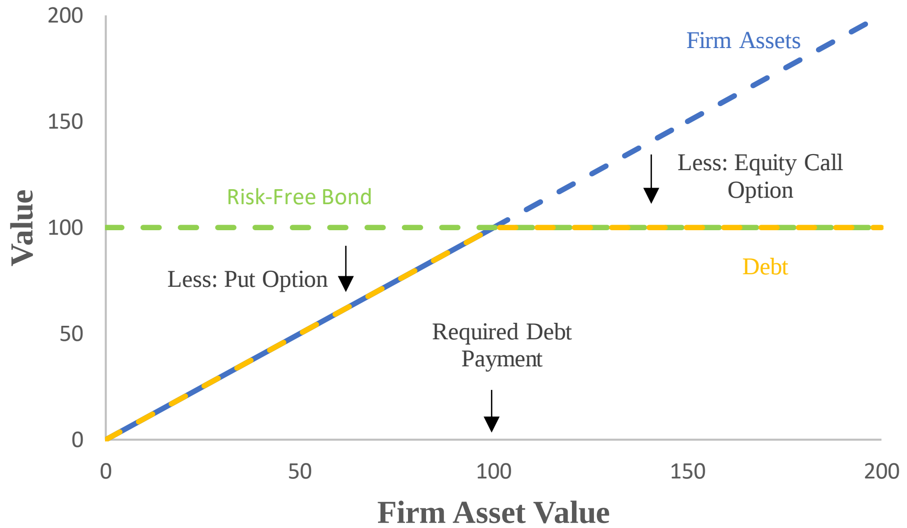
$$Risky\ debt = Risk\ free\ debt - Put\ option\ on\ the\ firm's\ assets \tag{2.3}$$

where the put price at the maturity date is equal to:

$$P = max(K - S, 0) \tag{2.4}$$

where the holder of a put will exercise the option if the stock price *S* is lower than the exercise price *K*. A short position in a put option is exercised when the value of a firm's assets falls below the debt payment. Figure 4 illustrates this relationship, showing how risky debt parallels the combination of risk-free debt and a short put option. If the firm's assets are valued less than the debt payment, the option holder exercises the option, receiving the shortfall amount, and the debt holder is left with the assets. In contrast, if the asset value exceeds the debt payment, the option will not be exercised, and the debt holder is paid in full.

Figure 4: Debt as an option portfolio.



Source: Berk & DeMarzo (2019).

2.4 The Merton Model

Fischer Black and Myron Scholes introduced a groundbreaking theory of option pricing in their Black & Scholes (1973) paper. While they recognized options as highly specialized financial instruments during that period, their importance was also underscored by their trading on the Chicago Board Options Exchange. Black and Scholes saw the wider applicability of their approach, realizing that the principles of the Black-Scholes model could extend beyond options, contributing significantly to the broader field of corporate finance.

The Black-Scholes formula made it possible to value options before their maturities and is also easy to implement. The model only requires five input variables to price a call option: stock price, exercise price, exercise date, risk-free interest rate, and the volatility of the stock price returns. All the models utilized in this thesis are based on the fundamentals of Black & Scholes (1973). However, Robert C. Merton was the first to implement the ideas of Black & Scholes in the setting of corporate finance in his Merton (1974) paper. Recognizing that Black & Scholes (1973) can have a broader application than just to calculating option prices, Merton develops an extended version of the Black-Scholes model to value equity for a company with only non-interest-bearing debt:

$$S = VN(d_1) - e^{-rT}FN(d_2) \tag{2.5}$$

where S is the value of equity, V is today's value of the firm's assets, F is the face value of non-interest-bearing debt, r is the risk-free rate, and T is time until maturity measured in years. Further, d_1 and d_2 are calculated in the equations:

$$d_1 = \frac{ln\left(\frac{V}{F}\right)+\left(r+\frac{1}{2}\sigma_V^2\right)T}{\sigma_V\sqrt{T}} \tag{2.6}$$

$$d_2 = d_1 - \sigma_V\sqrt{T} \tag{2.7}$$

where σ_V is the volatility on the firm's assets. Equation 2.5 can then be reorganized to find the value of the firm's non-current debt:

$$D = e^{-rT}F - \left(e^{-rT}F - V\frac{N(-d_1)}{N(-d_2)}\right)N(-d_2) \tag{2.8}$$

Equation 2.8 consists of three terms. The $e^{-rT}F$ is the present value of equivalent risk-free debt, $\left(e^{-rT}F - V\frac{N(-d_1)}{N(-d_2)}\right)$ is the expected present value of loss given default, and $N(-d_2)$ is the probability of default. Merton successfully creates a model that incorporates both the loss given default, the probability of default, and time to maturity, which are central elements of calculating the credit spread, as displayed by Figure 2 and Equation 2.1.

2.5 Modeling the Yield Curve

Models based on the Black & Scholes framework, require suitable discount rates to accurately compute present values. To correctly discount bonds of varying maturities, it is essential to apply specific zero-coupon rates. This necessitates constructing a yield curve that corresponds to each bond's maturity. Some models for yield curves are fixed, such as the model presented by Nelson & Siegel (1987). Other models adopt a stochastic mean-reverting approach, like the Vasicek (1977) model and the Cox-Ingersoll-Ross (known as CIR) model presented by Cox et al. (1985). Moreover, there are yield curve models that adjust their parameters to align with market data. For instance, the Black, Derman & Toy (BDT) model presented by Black et al. (1990) find the model parameters by calibrating market data much the same way as implied volatilities are computed in equity options. These models represent just a few of the many approaches to yield curve modeling, as various structural bond pricing models assume different methods for estimating yield curves. Chapter 3 delves into the choice of yield curve models for this paper and how they are linked to each specific structural bond pricing model.

3. METHODOLOGY

In this chapter, we present the three structural bond pricing models being used to predict spreads for the Norwegian high-yield bond sample. Chapter 3.1 starts by describing the fundamentals of the extended Merton model used, first presented by Eom et al. (2004). Then we move on to present the second model in Chapter 3.2, being the Leland and Toft model first presented in the article of Leland & Toft (1996). Thirdly, we present the Longstaff and Schwartz model in Chapter 3.3, originally detailed in the article of Longstaff & Schwartz (1995). Finally, we display how we compute spreads from modeled prices in Chapter 3.4.

All three structural bond pricing models incorporated in this thesis are collected from the paper of Eom et al. (2004). The extended Merton model used in this paper is based on the original Merton model first elaborated in the article of Merton (1974). The idea of this extended Merton model follows the same idea as the original Merton model and prices debt by assessing the relationship between an issuer's asset value and total liabilities at every period. However, this extended Merton model is adjusted to include coupon payments.

In the Leland & Toft (1996) model, firms issue a set amount of debt with continuous coupons. The model operates on a fixed interest rate and determines bankruptcy endogenously, providing insights into optimal capital structures. The Longstaff and Schwartz (1995) model prioritizes the combined analysis of default risk and interest rate risk in the valuation of corporate debt. The model introduces valuation techniques for both fixed and floating-rate debts, underscoring a thorough approach to risky debt evaluation.

The three structural models utilized for corporate bond valuation in this study have distinct assumptions. Our primary goal in incorporating these models is to assess the pricing accuracy across various yield curve forecasts. The extended Merton model is based on a discretely set yield curve, Leland & Toft's model assumes a constant interest rate, and Longstaff & Schwartz's model considers a stochastically mean-reverting yield curve. We hereby refer to the extended Merton model as M, the Leland & Toft model as LT, and the Longstaff and Schwartz model as LS. To determine the model with the most accuracy, we compare it with the actual prices and spreads of the bond sample.

3.1 The Extended Merton Model

The extended Merton model described by Eom et al. (2004) considers a defaultable bond with the maturity T and unit face value that pays semi-annual coupons at an annual rate c. The approach values coupons, and the final principal, as a portfolio of zero-coupon bonds. For simplicity, we use the assumption that $2T$ is an integer, such that we let $T_n,\ n = 1, \dots, 2T$, be the *n*th coupon date. In the model, $K_t = K\ \forall \in [0, T]$, where ∀ indicates that the default barrier K is constant for every period between the issue date and the maturity date. The default is triggered if the asset value falls below K on coupon dates. The price of a coupon bond can then be measured by the present value of the expected payoffs from coupons and the principle using the following equation:

$$p^M(0,T) = \sum_{i=1}^{2T-1} D(0,T_i) E^Q \left[\left(\frac{c}{2}\right) I_{\left\{V_{T_i} \geq K\right\}} + min\left(\frac{wc}{2}, V_{T_i}\right) I_{\left\{V_{T_i} < K\right\}} \right] +$$
$$D(0,T) E^Q \left[\left(1 + \frac{c}{2}\right) I_{\{V_T \geq K\}} + min\left(w\left(1 + \frac{c}{2}\right), V_T\right) I_{\{V_T < K\}} \right] \qquad (3.1)$$

$D(0, T_i)$ is the time 0 value of a default-free zero-coupon bond maturing at T_i, $I_{\{.\}}$ is the indicator function, $E_{[.]}^Q$ is the expectation at time 0 under the Q measure (the risk-neutral measure), and w is the recovery rate. The formula is divided into two parts. The first part is the sum of present values for all future expected coupons at a risk-neutral measure. The discount factor is given by $D(0, T_i)$, the value of the semi-annual coupon times the risk-neutral probability of the asset value V being above the default barrier at the maturity date is given by $\left(\frac{c}{2}\right) I_{\left\{V_{T_i} \geq K\right\}}$. The minimum asset value at maturity and the semiannual coupon times the recovery rate, at the risk-neutral probability of the asset value at maturity being below the default barrier, is given by $min\left(\frac{wc}{2}, V_{T_i}\right) I_{\left\{V_{T_i} < K\right\}}$. The second part measures the expected risk-neutral present value of the payment in the last period of the bond's life but follows the same calculation pattern as the first part. The last bond payment includes the principle and the last coupon put together.

It can further be shown that:

$$E^Q\left[I_{\{V_t \geq K\}}\right] = N\big(d_2(K,t)\big) \tag{3.2}$$

$$E^Q\left[I_{\{V_t<K\}} min(\Upsilon, V_t)\right] = V_0 D(0,t)^{-1} e^{-\delta t} N\big(-d_1(\Upsilon,t)\big) + \Upsilon\left[N\big(d_2(\Upsilon,t)\big) - N\big(d_2(K,t)\big)\right] \tag{3.3}$$

where $Y \in [0, K]$ and $N(\,.\,)$ denotes the cumulative standard normal function and:

$$d_1(x,t) = \frac{ln\left(\frac{V_0}{xD(0,t)}\right) + \left(-\delta + \frac{\sigma_v^2}{2}\right)t}{\sigma_v \sqrt{t}} \tag{3.4}$$

$$d_2(x,t) = d_1(x,t) - \sigma_v \sqrt{t} \tag{3.5}$$

The term structure $D(0,T)$ is measured by the Nelson-Siegel-Svensson model which is described in Chapter 3.1.2. Including the given term structure, the formulas above can be utilized to compute the price of a risky bond under the assumptions of Merton.[1]

3.1.1 Asset value and asset volatility

We compute asset value by adding the market capitalization to the book value of total liabilities in the sample period for every month. Both of these measures are gathered from the Infront Professional Terminal using their Excel Add-In. Asset volatility is an unobservable measure, but can be calculated using the historical equity return volatility from the equation:

$$\sigma_e = \sigma_V \frac{V_t}{S_t} \frac{\partial S_t}{\partial V_t} \tag{3.6}$$

The equation makes it possible to isolate asset volatility, named as σ_V, given that we have data for equity volatility σ_e, asset value at each month V_t, and market capitalization at each month S_t. We assume a constant measure for equity volatility for each issuer in the sample, computed as the standard deviation of equity returns over the whole sample period.

[1] See Appendix A1 for more a more in-depth breakdown of this extended Merton model.

Alternatively, we could have calculated asset value and asset volatility simultaneously through Merton's framework, linking the above equation to the general Merton model introduced in Chapter 2. Sundaresan (2009) presents a KMV-approach to find the implicit asset value and implicit asset volatility at the same time. This approach determines equity value as what remains after the firm's debt has been settled and is frequently used in structural models. In these models, the expected payoff is typically calculated relative to the face value of the bond under consideration. The value of equity begins to accumulate only after the bond's par value has been settled, assuming the company holds no other debt forms. However, the companies in the sample possess diverse types of debt, indicating that incorporating total liabilities is the optimal strategy, just like Eom et al. (2004) do in their paper. The equity value in the model then starts to accumulate once all liabilities are cleared.

3.1.2 Risk-free rate using Nelson-Siegel-Svensson

We gather information on Norwegian zero-coupon bonds (ZCB) to use as risk-free rates in line with the methodology of Eom et al. (2004). Although government bonds are not fully without risk, ZCBs issued by the Norwegian government are assumed to be one of the safest government bonds obtainable. The Norwegian central bank uses the Nelson-Siegel-Svensson (NSS) model to calculate missing zero-coupon rates on the yield curve. This method of yield curve modeling did first arise from Nelson & Siegel (1987), but was later extended by Svensson (1994). The Norwegian central bank has modeled a yield curve consisting of maturities for 6 months, 9 months, 12 months, and for yearly maturities between 1 year and 10 years. We extract daily yield curves from January the 1st 2015 until August 2023.[2] The coupons and principles of the high-yield bonds in the sample are then linked to the appropriate risk-free zero-coupon rates with the same maturity.

We could also model the yield curve using the framework of Vasicek (1977) and then use this to price the bonds as a comparison to the pricing model of NSS. However, this comparison is done by Eom et al. (2004) who got scarce differences in results when comparing the two approaches for the M model. We therefore decided to only include a discrete yield curve estimation using the NSS method when pricing bonds through this M

[2] A visual presentation of the NSS yield curve modeled on Norwegian government bonds is displayed in Appendix A2.

model.[3] Yield curve modeling using the Vasicek (1977) approach will be utilized for the LS model in this paper.

3.1.3 Payout ratio

The M model introduced by Eom et al. (2004) contains a variable for the issuer's payout ratio. To gain a good approximation of the payout ratio, we need data on the firm's leverage level (measured as total liabilities over market capitalization), dividend yield, share repurchases, and debt coupon level. The payout ratio is calculated as a weighted average of the coupon and the share-repurchase adjusted dividend yield. The coupon amount is weighted by the leverage ratio, whereas the share-repurchase-adjusted dividend yield is weighted by one minus the leverage ratio. Finally, the whole equation is divided by the total asset value to get a ratio format for the variable. The equation is displayed as:

$$\delta = \frac{Coupon * Leverage + Share-Repurchase-djusted-Dividends * (1 - Leverage)}{Total\ Asset\ Value} \quad (3.7)$$

Coupon rates are sourced from the Stamdata database, where they are presented either in full or as a spread above particular IBOR rates. To ensure consistency, we collect the IBOR rates on the same day we retrieve the coupon data from the database. We gather data on dividends and shares outstanding for each issuer directly from Infront. Share repurchases are assessed based on the decrease in shares outstanding from one period to the subsequent one.

3.1.4 Recovery rate

The recovery rate is the amount paid to bondholders in a redemption or settlement, which bondholders are particularly interested in because it predicts how much of the bond's coupons and principal will be repaid in the event of default. Eom et al. (2004) use a constant recovery rate of 51.31% of face value across their bond sample, which is based on previous research. Nonetheless, this estimate is outdated and based on American firms, prompting us to delve into more recent research. Aarvik & Nordli (2016) find an average recovery rate of 38.6%, measured between May 2014 and September 2016, across their sample based on 78 high-yield bonds in the Nordics. This result is equivalent to what Jankowitsch et al. (2014)

[3] We discuss the possible inherent bias to using Norwegian government bonds modeled with NSS in chapter 6.4.2, compared to using a yield curve based on swap rates from the Norwegian interbank market, or rates from AAA-rated corporate bonds.

find in their similar analysis of the US market. Additionally, Altman & Kishore (1996) stated at the time of their article's publication that the average prices of bonds have been 40% of face value. This rate has been commonly used in both academia and the industry for many years. We decided to align our estimate with previous research and use a recovery rate of 40%.

3.1.5 The default barrier K

The barrier value, K, remains fixed from the issue date to the maturity date, as outlined in Chapter 3.1. This value is derived by taking each issuer's total liabilities as a percentage of total assets on the issue date. These calculated percentages help determine the K value relative to the face value, V. Adopting this method allows the asset value to dip below K at certain times during the bond's life, influenced by asset volatility. If, on the other hand, K is dynamically set as a proportion of the asset value, it would never drop below the asset value. Therefore, setting a consistent value for K is a crucial aspect of the model.

3.2 The Leland & Toft (1996) model

In the model of Leland & Toft (1996), coupons are paid continuously, where the total coupon amounts to C every year. For this thesis, we use a semi-annualized coupon. The model is presented as:

$$P^{LT}(0,T) = \frac{C}{r} + \left(P - \frac{C}{r}\right)\left(\frac{1-e^{-rT}}{rT} - I(T)\right) + \left(wV_B - \frac{C}{r}\right)J(T) \tag{3.8}$$

where:

$$I(T) = \frac{\left(G(T) - e^{-rT}F(T)\right)}{rT} \tag{3.9}$$

$$J(T) = \frac{1}{zs\sqrt{T}}\left[-e^{(z-a)b}N\big(q_-(T)\big)q_-(T) + e^{-(z+a)b}N\big(q_+(T)\big)q_+(T)\right] \tag{3.10}$$

$$G(T) = e^{(z-a)b}N\big(q_-(T)\big) + e^{-(z+a)b}N\big(q_+(T)\big) \tag{3.11}$$

$$F(T) = G(T)|_{z=a} \tag{3.12}$$

and:

$$a = \frac{r-\delta}{\sigma_V^2} - \frac{1}{2} \qquad (3.13) \qquad\qquad b = \ln\left(\frac{V_0}{V_B}\right) \qquad (3.14)$$

$$z = \left(a^2 + \frac{2r}{\sigma_V^2}\right)^{\frac{1}{2}} \qquad (3.15) \qquad\qquad q_{\mp}(t) = \frac{-b \mp z\sigma_V^2 t}{\sigma_V \sqrt{t}} \qquad (3.16)$$

In the price equation, r represents the fixed interest rate for each bond, P represents the principal payment upon the maturity date, T is the time to maturity, V_B is the default boundary, and α represents the loss rate in the event of default. The firm is solvent for the condition $V > V_B$, implying that the equity value is positive. $V = V_B$ is the bankruptcy point and investors need to provide additional cash flow to keep the firm solvent. $V < V_B$ will never happen as equity appreciation never will be larger than the contribution investors make to keep the firm solvent. The firm will go bankrupt and a share w of the assets will be paid out to shareholders.

Among the three models we examine, only the LT model integrates taxes by considering the tax-deductibility of interest payments, displayed in the default boundary equation presented in Chapter 3.2.2. The tax rate is set to 22%, as a standard for Norwegian corporate tax. Beyond the computation of V_B, this model reuses variables from the M approach, which makes it considerably easier to implement this model than the M model.[4]

3.2.1 Fixed interest rate

The LT model assumes fixed interest rates meaning that it values the continuous coupons with a fixed rate until maturity for every date. To assign an interest rate to each bond with identical maturity, we utilize the yield curves that were also employed in the M model. By doing this we follow the analytical framework of Eom et al. (2004) to ease comparability of results. In this approach, every bond is matched with an interest rate based on its corresponding maturity date on its initial data point in the sample. This rate then remains constant throughout the duration of each bond's existence.

[4] For the sake of brevity, we leave it to the reader to further investigate the equations of the LT model in Leland & Toft (1996).

3.2.2 The default boundary V_B

The LT model presents V_B as a variable that is being determined endogenously by the following equation:

$$V_B = \frac{\left(\frac{C}{r}\right)\left(\frac{A}{rT} - B\right) - \frac{AP}{rT} - \frac{\tau C x}{r}}{1 + \alpha x - (1-\alpha)B} \tag{3.17}$$

where:

$$A = 2ae^{-rT}N\left(a\sigma_V\sqrt{T}\right) - 2zN\left(z\sigma_V\sqrt{T}\right) - \frac{2}{\sigma_V\sqrt{T}}n\left(z\sigma_V\sqrt{T}\right) + \frac{2e^{-rT}}{\sigma_V\sqrt{T}}n\left(a\sigma_V\sqrt{T}\right) + (z-a) \tag{3.18}$$

$$B = -\left(2z + \frac{2}{z\sigma_V^2 T}\right)N\left(z\sigma_V\sqrt{T}\right) - \frac{2}{\sigma_V\sqrt{T}}n\left(z\sigma_V\sqrt{T}\right) + (z-a) + \frac{1}{z\sigma_V^2 T} \tag{3.19}$$

$$x = a + z \tag{3.20}$$

Here, the default boundary is determined by several known variables, such as C, r, T, P and α. The variables A, B and x are also determined by variables that are previously computed. $N(.)$ is the cumulative standard normal distribution, $n(.)$ denotes the standard normal density function, and τ is the tax rate. V_B is computed for each bond based on its initial data point in the dataset, typically its issuance date. After establishing the default threshold, it remains unchanged for the entirety of each bond's duration. Leland & Toft (1996) specifically highlight that the V_B derived from their model varies with the chosen debt maturity, given the fixed principal and coupon amount. This is contrary to the model presented by Kim et al. (1993), which establishes a default threshold in line with flow-based bankruptcy, emphasizing that V_B is unaffected by the maturity of the debt.

3.3 The Longstaff & Schwartz (1995) Model

The Longstaff & Schwartz (1995) model builds on the model of Black & Cox (1976), but further extends it to develop a continuous-time valuation framework for risky debt that allows for both default risk and interest rate risk. The model considers a defaultable bond that pays semi-annual coupons and is given by:

$$P^{LS}(0,T) = \frac{c}{2}\sum_{i=1}^{2T-1} D(0,T_i)[1 - wQ^{F_i}(0,T_i)] + \left(1 + \frac{c}{2}\right) D(0,T)[1 - wQ^{F_i}(0,T_i)] \quad (3.21)$$

In the model, $D(0,T_i)$ represent the time 0 value of a T_i-maturity default-free zero-coupon bond as determined by the Vasicek (1977) model (see Chapter 3.3.1), c is the coupon rate, w is the recovery rate, and $Q^{F_i}(0,T_i)$ is the time 0 probability of default over $(0,T_i]$ under the T_i-forward measure. The equation can also be displayed as:

$$P^{LS}(0,T) = \frac{c}{2}\sum_{i=1}^{2T-1}[D(0,T_i) - D(0,T_i)wQ^{F_i}(0,T_i)] + \left(\left(1 + \frac{c}{2}\right) D(0,T) - \left(1 + \frac{c}{2}\right) D(0,T)wQ^{F_i}(0,T_i)\right) \quad (3.22)$$

where the equation breaks down the bond price into two essential components. The initial component aggregates the disparities between every upcoming discount factor for each bond and its counterpart, which is modified to account for default risk. Meanwhile, the latter component contrasts the present value of the bond's concluding payment with its version that has been adjusted for default considerations. The approach is the same as the M model, valuing coupons and the final principle as a portfolio of zero-coupon bonds. The recovery rate in this model is the same as what we used in the M model.[5]

3.3.1 Risk-free rate using Vasicek (1977)

The LS model assumes that the dynamics of interest rates are drawn from the term structure model of Vasicek (1977), which uses short term spot interest rates to model the rest of the yield curve. This is computed with the following equation:

$$dr = \kappa(m - r)dt + \sigma dz \quad (3.23)$$

[5] For the sake of brevity, we leave it to the reader to further investigate $Q^{F_i}(0,T_i)$, and the rest of the equations, in Longstaff & Schwartz (1995).

where κ is the speed of mean reversion, m is the long-term mean interest rate, r is the current interest rate, dt represents the change in time, whereas the term σdz is the stochastic component of the equation. The stochastic component is supposed to capture the randomness of interest rate changes and consists of $dW(t)$ which is the Wiener process, and σ being the standard deviation of the stochastic component and ultimately determines the magnitude of the randomness.

The model describes how the short-term interest rate r, is evolving through time. The equation is fully determined by the stochastic term σdz when $m = r$. This creates an opportunity for the model to compute negative short-term rates at times. This feature of the Vasicek (1977) model has been widely criticized in academia. However, the probability of negative interest rates occurring is small when realistic parameter values are used. Additionally, given that the current value of the spot rate is positive, the dynamics of the model always imply positive expected values of interest rates. [6]

3.4 Spread Prediction

The models in Chapter 3 compute semi-annual prices for each bond in the sample. To make the results comparable to previous research, we find the spread for each computed price using the following equation:

$$S(0,T) = \frac{\ln\big(D(0,T)\big) - \ln\big(P(0,T)\big)}{T} \tag{3.24}$$

The equation calculates the yield difference between the price of a risk-free bond $D(0,T)$ and a risky bond $P(0,T)$, quantifying the spread between the two. This approach is also used in Ytterdal & Knappskog (2015), where they determine spreads using prices derived from an extended Merton model. Given that this method of estimating spreads through modeled prices and discount rates is established in prior research, we have chosen to adopt this methodology. However, alternative spread calculation methods exist, such as asset-swap-spread, maturity spread, and z-spread, which we will discuss further in Chapter 6.4.4 of our analysis.

[6] A visual presentation of the Vasicek yield curve is displayed in Appendix A3. Moreover, we leave it to the reader to further investigate the equations of the Vacisek model in Vasicek (1977).

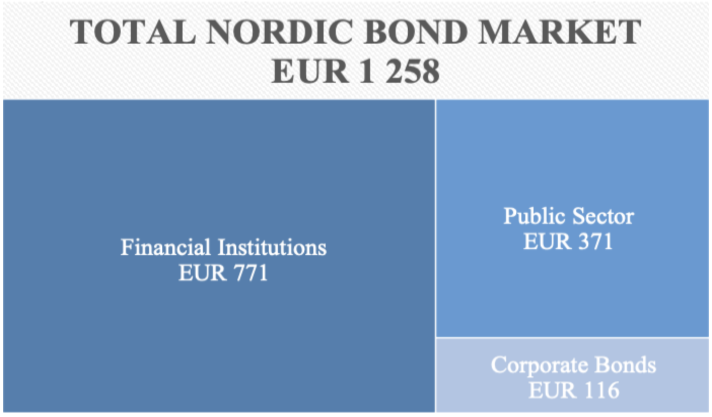
4. DATA

This chapter offers a detailed walkthrough of how we narrowed the sample from the raw data. Initially, we describe the Norwegian high-yield bond market, highlighting its size and significance within the broader Nordic bond market. This information aims to provide readers with a clear perspective on the scale of the market under study. Following this overview, we delve into the methodology used to transform raw data into our specifically defined sample, explaining the steps and criteria involved in this process.

4.1 The Presence of Norwegian High-Yield in the Nordics

Within the broader Nordic bond market, the corporate bond segment occupies a minor portion. Figure 5 offers a visualization of the Nordic bond market's structure. The majority, at 61%, is made up of financial institutions, followed by the public sector with 30%, and corporate bonds trailing at 9%.

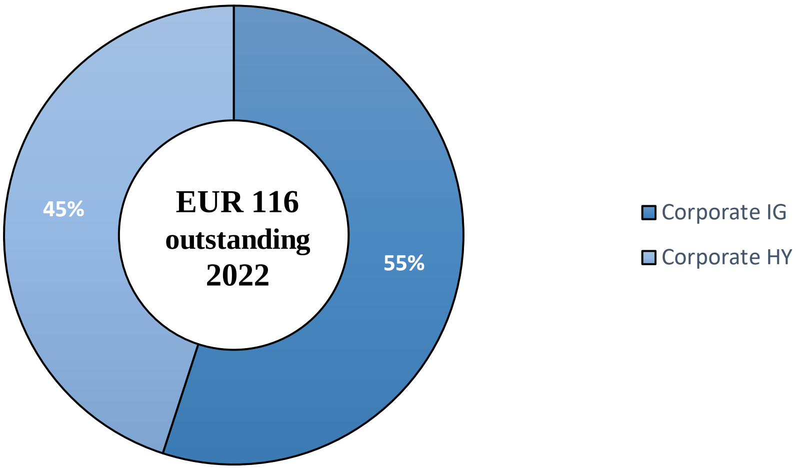
Figure 5: The Nordic bond market at year-end 2022. Measured in millions.



Source: Nordic Trustee.

Furthermore, the corporate bond market splits into two categories based on the inherent risk of debt: IG bonds and HY bonds. These categories represent 55% and 45% of the entire market respectively, as presented in Figure 6.

Figure 6: Distribution of the Nordic corporate bond market as of year-end 2022. Measured in millions.



Source: Nordic Trustee.

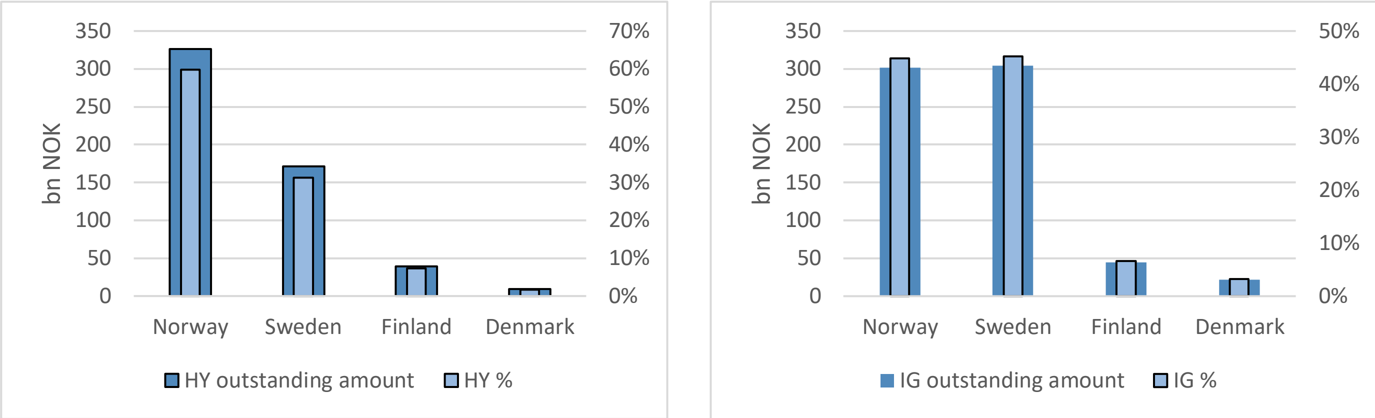
The upcoming figures highlight the individual share each Nordic country holds in the overall corporate bond market. They detail the outstanding sums in both IG and HY bonds for each respective country in NOK. Identified by their ISIN codes, the bonds referenced in Table 2 and Figure 7 are exclusive to NO, SE, FI, and DK ISINs. It is worth noting that while the ISINs designate the trading country, some bonds might have issuers outside the Nordic region.

Table 2: The distribution of the Nordic Corporate Bond Market across countries. Numbers in NOK as of year-end 2022.

Country	Norway	Sweden	Finland	Denmark	TOTAL
Total outstanding amount	628	476	84	31	1 220
Total %	52%	39%	7%	3%	100%
IG outstanding amount	301	304	45	22	672
IG %	45%	45%	7%	3%	100%
HY outstanding amount	327	171	40	9	547
HY %	60%	31%	7%	2%	100%

Source: Nordic Trustee.

Figure 7: Visualization of the corporate bond market distribution across countries. Numbers as of year-end 2022.

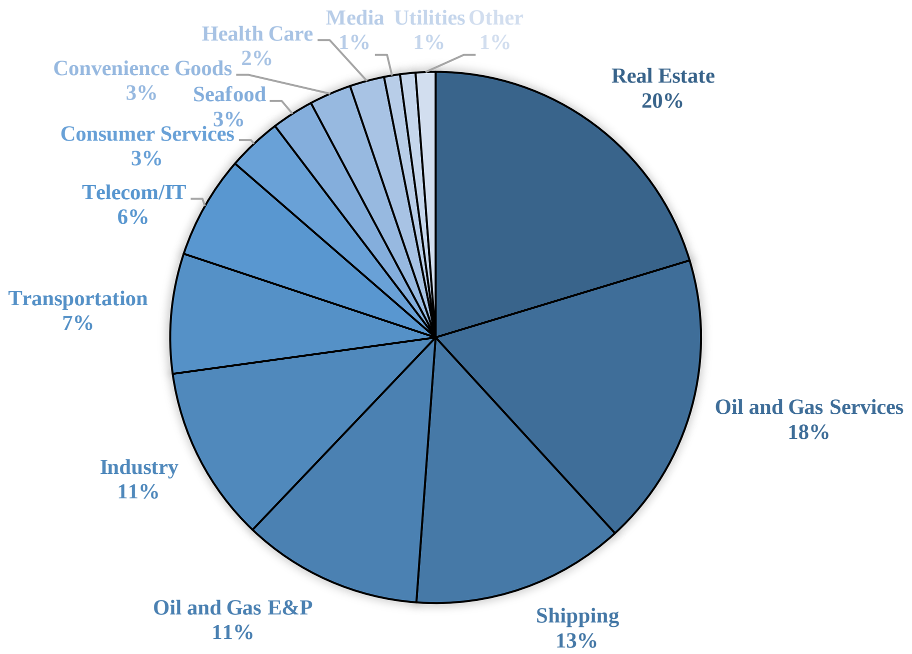


Source: Nordic Trustee.

Overall, both the IG and HY markets see a substantial presence from Norwegian and Swedish corporate bonds. The volume of NO and SE ISINs in the IG market is almost at the same level. However, in the Nordic HY market, Norwegian ISINs claim a dominant position with a 60% share, compared to 30% for Swedish ISINs.

As illustrated in Figure 8, the Norwegian HY market's landscape at the close of 2022 is diversified across several industries. The distribution in the figure is organized on the operational sector of each issuer, specifically omitting financial entities like banks, insurance firms, and other financial institutions. The real estate sector stands out, constituting 20% of the market based on outstanding amounts. The Oil and Gas industry, a large sector in the Norwegian economy, has a notable representation in the HY market. This is evident in Figure 8, which shows a significant presence of Oil and Gas Services (18%) and Oil and Gas E&P (11%). Furthermore, the shipping industry is not to be overlooked with an 18% market share.

Figure 8: Distribution of the Norwegian High-Yield market by sector, measured by outstanding amount at year-end of 2022. Financial companies are not included.



Source: Stamdata.

4.2 Nordic Trustee

Nordic Trustee is the leading and most experienced provider of bond trustee and loan agency services in the Nordic region. The main purpose of a bond trustee is to protect the interest of bondholders while being a contact point for issuers as the trustee represents investors on a collective basis. Nordic Trustee was established between 1993-1995 by large institutional investors and banks, and played an active role in developing an efficient bond infrastructure in the Nordics during the '90s, participating in actively building the Norwegian high-yield market through 2004-2008. Today, Nordic Trustee have over 3000 active assignments in the non-banking lending sector accounting for more than 850 different issuers from 30 different countries. Through its subsidiaries, Stamdata and Nordic Bond Pricing, Nordic Trustee is the main data provider for this paper.

4.2.1 Stamdata

Being a subsidiary of Nordic Trustee, Stamdata delivers insights on a vast number of debt securities within the Nordic markets, presenting bond-specifics exclusively upon issuance. Stamdata provides insights into various bond metrics, including outstanding amounts, bond ratings, issuer industry classifications, redemption type, debt risk profiles, return type (either floating or fixed rates), green label (either green or non-green), and both coupon rates and reference rates, among others. These metrics play a pivotal role in the analyses presented in Chapter 6 of this paper.

4.2.2 Nordic Bond Pricing

High-yield bonds are mostly traded over the counter. Additionally, they are often traded between parties without a financial intermediary logging the transactions to the public. For this reason, finding corporate bond prices is challenging, and most corporate bond prices are not available in financial terminals such as the Infront Professional Terminal and the Bloomberg Terminal.

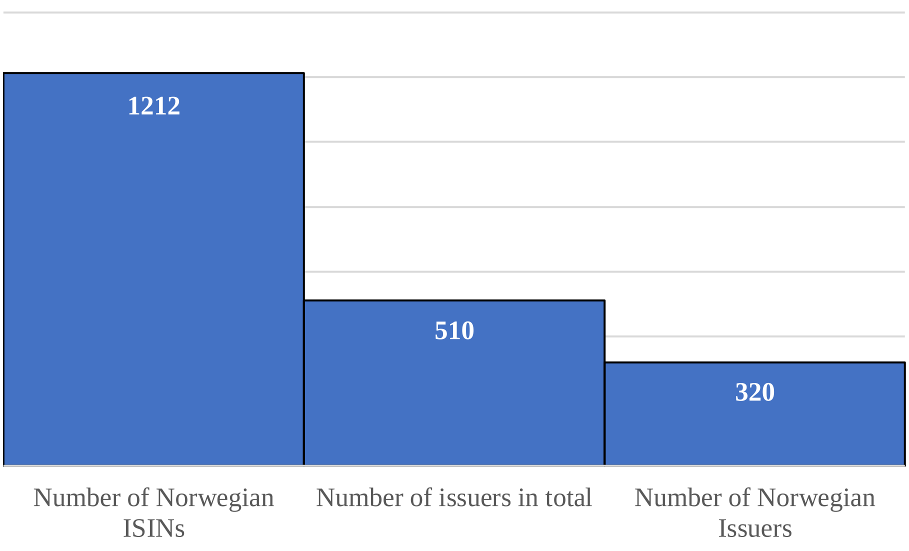
Nordic Bond Pricing (NBP) is a subsidiary of Nordic Trustee, and provides daily updates on bond prices, spreads, durations, and yields among other financial metrics. NBP has agreements with the main brokerage houses to gather the needed data for calculating bond prices. The bond prices are then processed through financial models to create bond prices and other metrics in line with best practice standards. An independent auditor confirms the adherence of all procedures and financial models to the company's official guidelines. In this thesis, we contrast the bond prices we model against those computed by NBP, considering NBP's prices as the benchmark. NBP's data is pivotal for the empirical research in this thesis.

An important note about the data is that it is not without its imperfections. Prices of some bonds are not always accurate as they may stay fixed over periods of time. This thesis implements the structural bond pricing models on semi-annual periods which decreases the severity of this issue as the data is not as affected by weeks of fixated prices. We discuss biases of potential NBP spreads further in Chapter 6.5.

4.3 The Raw Data

This paper collects necessary data for bond pricing through three channels. Specific bond characteristics for Norwegian HY bonds are gathered from the Stamdata database, actual daily bond spreads are collected from the NBP database, and market data is collected from the Infront Professional Terminal. The goal of this thesis is to measure the benchmark spreads calculated by NBP against our modeled spreads. NBP provided us with daily spreads for Norwegian ISINs from the middle of 2014 and until the present. However, we decide to start spread modeling from the beginning of 2015 based on the availability of yield curve data points provided by the Norwegian central bank (See Chapter 3.1.2). Raw data for Norwegian high-yield bond ISINs are extracted from the Stamdata database and overall statistics are exhibited in Figure 9.

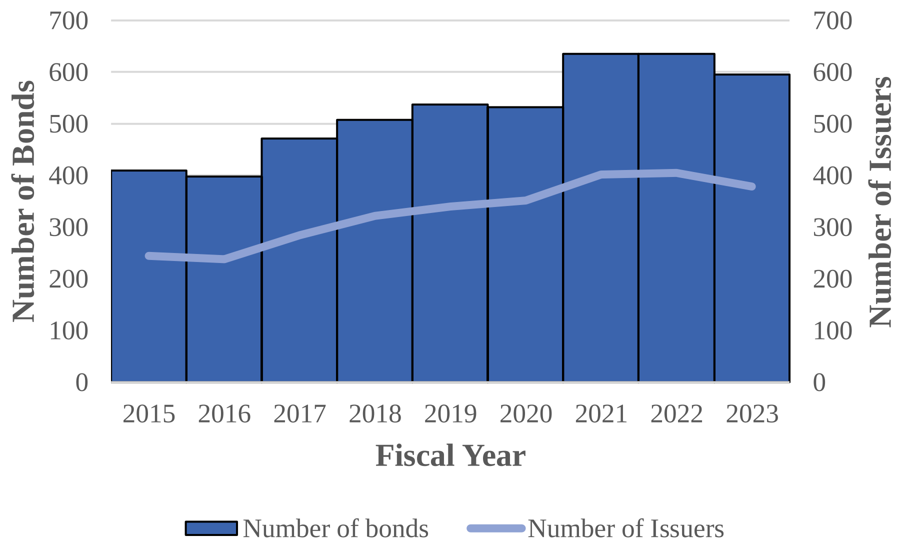
Figure 9: Overall statistics of the Norwegian high-yield bond market.



Source: Stamdata.

It is not a perfect match between the number of bonds and the number of issuers. This is because some issuers have several unique bonds outstanding. Both the number of bonds and the number of issuers in the market have increased over the sample period as exhibited by Figure 10. The numbers for both bonds and issuers are lower in 2023 because only bonds with an issue date before 08.01.2023 are included.

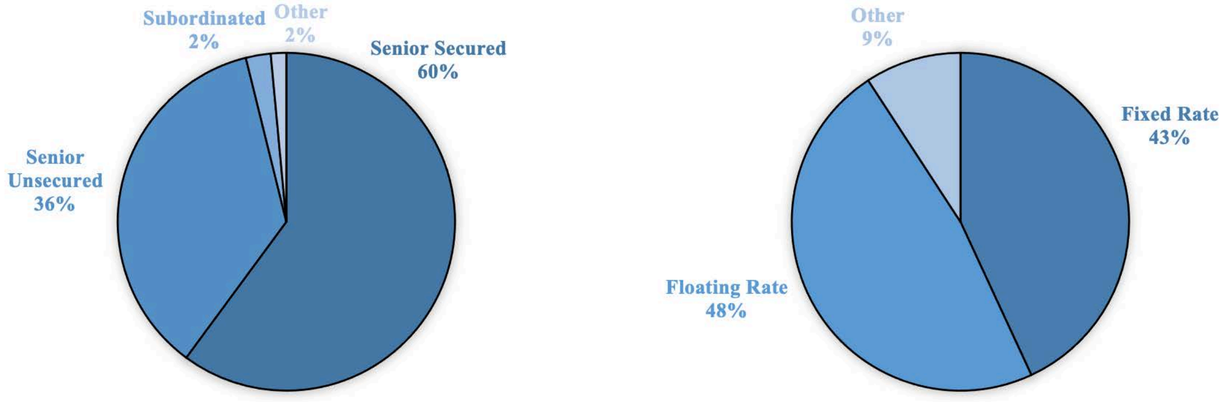
Figure 10: Total number of Norwegian high-yield bonds and issuers of Norwegian high-yield bonds in each year between 01.01.2015 and 08.01.2023.



Source: Stamdata.

Figure 11 displays that during the sample period, 60% of the total issued HY bonds have been senior secured bonds, which have asset backing to provide security in case of default. On the other hand, 36% were senior unsecured bonds, which do not possess asset backing if default occurs. A minimal portion, 2%, were categorized as subordinated risk bonds. The split between fixed-rate and floating-rate bonds has nearly been equal, with fixed-rate bonds making up 43% and floating-rate bonds accounting for 48%. Other types of return structures, such as step rate, zero-coupon rate, and adjustable rate, have been a smaller fraction, totaling only 9% of the overall issuance.

Figure 11: Distribution of risk types and return types across the Norwegian high-yield market.



Source: Stamdata.

4.4 Shaping the Sample

In shaping the dataset extracted from Stamdata, we exclusively select bonds issued by non-financial companies. The intent is to ensure that leverage ratios are comparable across companies, considering the fundamental differences in balance sheet structures between non-financial companies and financial companies. This specific limitation aligns with the approach taken by Eom et al. (2004). Accordingly, bonds issued by banks, insurance companies, and other financial institutions are therefore excluded.

The dataset is narrowly defined to include solely plain vanilla bonds. This involves the exclusion of bonds with additional features, such as convertible bonds, as well as callable and puttable options, and any bonds with attached warranties. Moreover, the decision to eliminate financial companies from the dataset also effectively filters out corporate debt classified as Tier 1-3.

The most significant narrowing of the dataset occurs during the filtering of bond issuers by nationality and the presence of exchange-traded equity. This paper focuses specifically on the pricing of corporate high-yield bonds through structural models, which necessitates the use of market data on stock prices and balance sheet information. For the sake of simplicity, we have chosen to concentrate on Norwegian HY bond issuers based on these companies being the majority of issuers in the dataset.

The bonds we are left with after narrowing down the initial dataset have been issued by a variety of Norwegian companies, all of which are either currently listed or have previously been listed on the Oslo Stock Exchange. It should be noted that the majority of these bond issues are characterized as floating-rate bonds, indicative of their interest payments being tied to a variable benchmark interest rate. This financial instrument is often preferred by companies seeking to align their debt costs with current market rates.

4.5 Infront Professional Terminal

The utilization of structural bond pricing models requires balance sheet information and market data for each issuer in the sample. We have gained access to the Infront Professional Terminal and are therefore able to extract the needed data. Infront is a provider of real-time market data and is used by thousands of wealth managers, traders and, other finance professionals.

4.5.1 Balance sheet data

We utilize Infront's Excel Add-In to streamline the collection of balance sheet data for our modeling. We retrieve balance sheet data every quarter to maximize data frequency. We aim to present balance sheet data monthly, which entails replicating the quarterly recorded data for the subsequent two months. This approach facilitates our computation of semi-annual bond prices.

For each bond issuer, we gather data on total liabilities, current and non-current liabilities, total outstanding shares, and a multiple for dividends per share. However, the extent of the archival data before 2015 on the terminal is limited. Consequently, our analysis of the bond market is anchored to the timeframe for which the Infront system provides comprehensive and reliable balance sheet data.

When analyzing publicly listed issuers, we encounter the limitation that structural data is only refreshed quarterly. This means that while market capitalization figures are updated monthly, metrics such as total liabilities, interest expenses, and outstanding shares remain static until the next quarter's publication. Consequently, since market capitalization is the only variable that changes monthly, it directly affects the total assets and, by extension, variables such as asset volatility, payout ratio, and leverage ratio. As the market capitalization is updated monthly, the measures are not static.

Langdalen & Johansen (2016) address the challenge of static debt figures by applying a linear approximation from the start-of-year to the end-of-year leverage for the firms in their sample. This method is effective for historical data analysis but falls short for real-time application, as periods between quarterly statements would require estimation of future leverage. This means speculative assumptions about whether and how much a firm's leverage would fluctuate, an approach that is not only time-consuming but also impractical for analyzing large datasets. Therefore, we do not modify this measure.

4.5.2 Market data

We utilize the Infront Excel Add-In to retrieve market data. Each bond issuer's monthly share prices are gathered to determine monthly market capitalizations. Furthermore, we gather market data for interest rates, enabling us to construct yield curves for discounting purposes. We decide to collect monthly NIBOR rates for computing the Vasicek (1977) model, which we present in Chapter 3.3.1. It is worth noting that some bonds from the Stamdata database only display a coupon margin tied to reference rates such as NIBOR, STIBOR, LIBOR, or EURIBOR. For simplicity, we extract the corresponding IBOR rates from Infront on the day of data retrieval, and then integrate them with the coupon margin to determine the full coupon values.

4.6 Final Sample

Having applied multiple constraints to the initial dataset, we are left with a final sample comprising 141 bonds across 43 issuers. This represents a substantial decrease from the initial collection of 1212 bonds and 510 issuers (320 Norwegian issuers), as presented in Figure 9. Additionally, Figure 12 presents the count of bonds and issuers active in each year throughout the sample period.

Figure 12: Sampled number of high-yield bonds and issuers for each year.

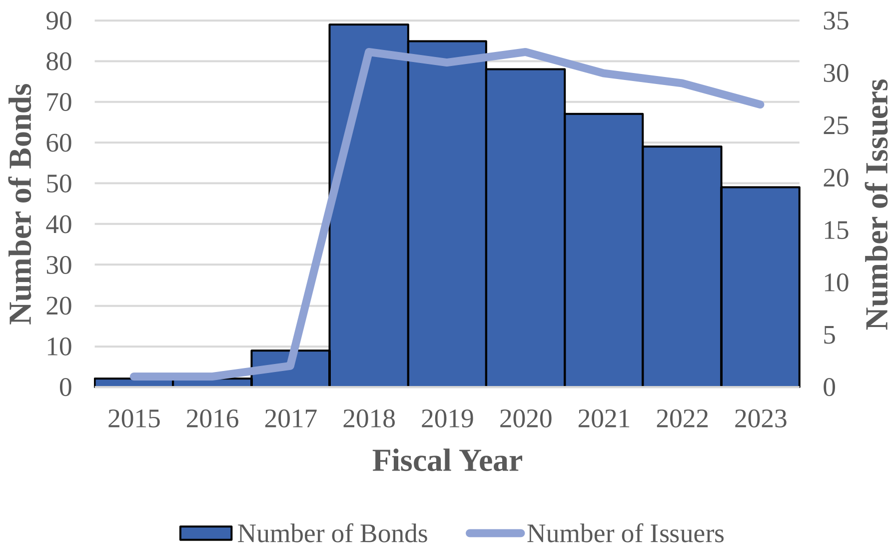


Figure 12 illustrates the bonds that were issued in or before each specified year, with maturities falling within or after that same year. The numbers of bonds and issuers displayed in Figure 12 are exclusively those that can be fully priced using the structural models discussed in Chapter 3. These are the bonds for which we can secure all the necessary data to calculate the output prices. The sample of high-yield bonds suffer from missing market data in the earliest years, which is the reason why the sample only has a few bonds before 2018. In attempting to extract data from Infront, we encountered difficulties in retrieving quarterly balance sheet information for the early years, which often resulted in receiving "N/A" values. Towards the end of the sample period, we observe a downward trend in the number of issuers and bonds. This decline can be attributed to several factors: some issuers initially included in the sample were either delisted from the Oslo Stock Exchange or faced defaults during the period. Furthermore, as bonds reach maturity and are redeemed, they are removed from the sample, along with their issuers if they have only one outstanding bond. Such a trend is not surprising, considering that our research focuses on higher-risk firms in the HY bond market, where such occurrences are more common.

The sample consists of bonds with issuers operating in several different industries. Table 3 displays the distribution of issuers across different industry groups as of August 2023. The most interesting part here is that the presence of the real estate industry is surprisingly small. This is in stark contrast to the initial dataset extracted from the Stamdata database, where real estate accounted for 20% of all Norwegian HY bond ISINs, as indicated by Figure 8. Within the sample, the shipping sector dominates with a 24% share, followed by the industry sector at 19%, and closely by Oil and Gas E&P and Services at 17% and 14%, respectively. The prominence of these four sectors in the sample is not unexpected, reflecting their substantial representation in the original dataset.

Table 3: Sample distribution of bonds across industry groups in our sample.

Industry Group	Number of Bonds	% of Total
Health Care	1	1%
Industry	27	19%
Oil and Gas E&P	24	17%
Oil and Gas Services	20	14%
Pulp, Paper and Forestry	3	2%
Real Estate	6	4%
Seafood	9	6%
Shipping	34	24%
Telecom/IT	9	6%
Transportation	6	4%
Utilities	2	1%
Total	141	100%

Table 4 offers a statistical overview of the variables relevant to the HY bond sample, which will be utilized for structural bond pricing analyses in Chapters 5 and 6. The bonds in the sample have an average maturity of approximately 6.46 years, with the range spanning from a minimum of around one year to a maximum of 16 years. The table reveals substantial variation within the sample, as evidenced by the large standard deviations and the span between the minimum and maximum values for asset volatility, payout ratio, and leverage ratio. This underscores the diversity among the issuers in the sample, reflecting substantial differences in company size and capital structure.

When comparing the NSS rate, which is incorporated into the M and LT models, with the Vasicek rate used in the LS model, we observe that the NSS rate exhibits less volatility. The adoption of a deterministic approach in the M and LT models contrasts with the stochastic application in the LS model, which is particularly evident in the Vasicek rate's large maximum value. This difference highlights the Vasicek rate's great sensitivity to the recent short-term interest rate fluctuations observed in the market.

Table 4: Summary statistics of bond characteristics.

Variable	Mean	Std	Min	25th Percentile	75th Percentile	Max
Maturity	6.46	3.17	1.00	4.00	9.00	16.00
Asset Volatility	29.48%	124.56%	0.00%	12.71%	40.96%	2361.04%
Payout Ratio	13.44%	83.61%	0.00%	0.19%	1.59%	600.43%
Leverage Ratio	120.56%	186.25%	1.35%	21.35%	136.60%	2118.57%
NSS rate	1.38%	0.93%	-0.04%	0.73%	1.60%	4.23%
Vasicek rate	2.01%	2.34%	0.00%	0.48%	2.08%	13.85%

5. RESULTS

5.1 Modeled Spreads Against Actual Spreads

This section delves into the visual presentation of the modeled spreads and the actual spreads of the three models. To effectively visualize the results, we compute the mean of the spreads for all the bonds at every date, as is done by Feldhütter & Schaefer (2018). We also include summary statistics on an annual basis for each model. Note that both the modeled and actual spreads are represented by two lines. The smooth line in these representations is a trend line for the average spreads, while the volatile line is the true average spreads.

The implementation of structural bond pricing models is not without its intricacies. As the complexity of these models escalates, they become increasingly sensitive to the specific conditions of input variables, such as interest rates, asset volatility, and leverage ratios. This sensitivity can result in corporate bond spreads becoming incalculable, particularly within the LT and LS models. The challenge lies in the multitude of variable calculations required. An extreme value in any one of these can disrupt the pricing mechanism and result in incalculability. We try to mitigate this problem by winsorizing the variables where the problem is most common, but we still lose many observations.[7]

Consequently, our analysis is restricted to the subset of bonds for which each model can reliably compute spreads, and this is why the average actual spreads between models are not the same. This limitation implies that our testing and visual comparisons do not necessarily compare the spreads of the same sample of bonds between models. Nonetheless, while the visual comparability of the bond pricing models with observed spreads may not be particularly informative between models, their average spreads speak to the yearly variations and trends in the samples.

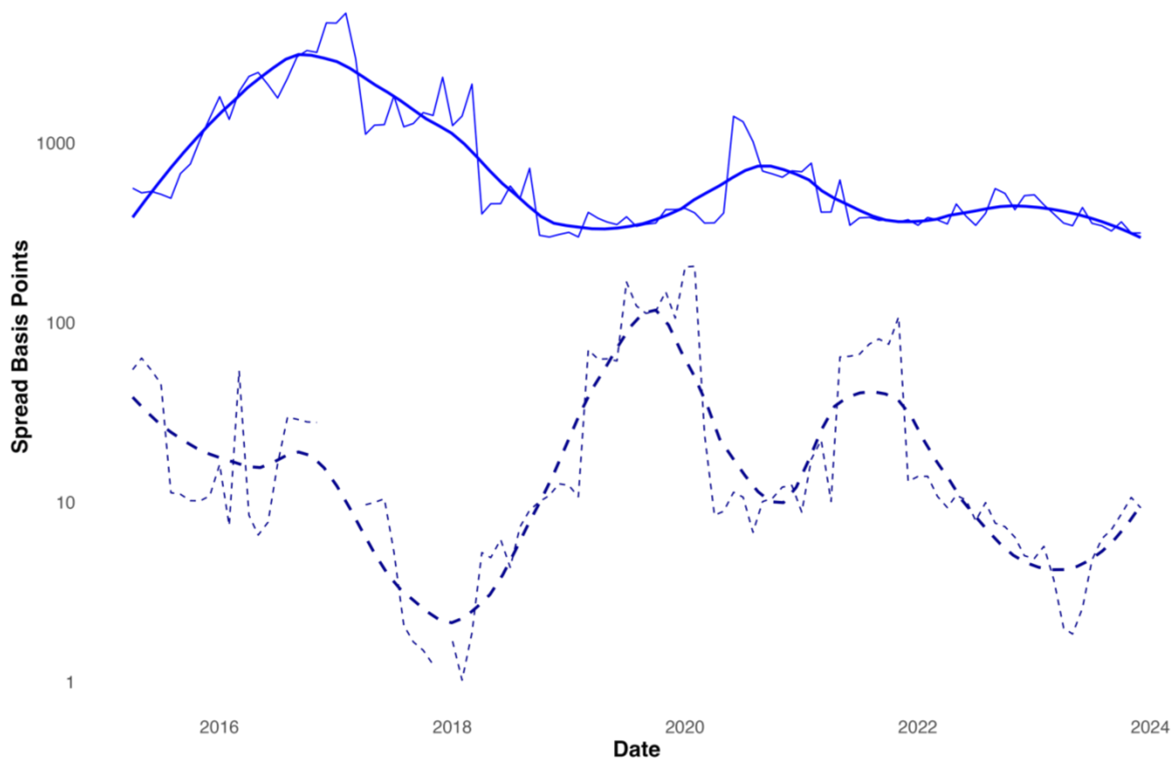
Figure 13 presents the modeled spreads from the M model against the benchmark spreads from NBP. In line with Eom et al. (2004), Ytterdal & Knappskog (2015), and Langdalen & Johansen (2016) the M model severely underpredicts the spreads on average for every date when compared to the spreads observed in the market. The summary statistics for the model

[7] We observe extreme values for asset volatility, payout ratio, and leverage ratio. Winsorizing of these variables is performed on both the 5% and 95% level.

display extreme max values for the model in 2019, and a large standard deviation. As is evident by the summary statistic, the model consistently models mean spreads in the single- and double digits. Notably, during most years, the M model has the least standard deviation of all the models.

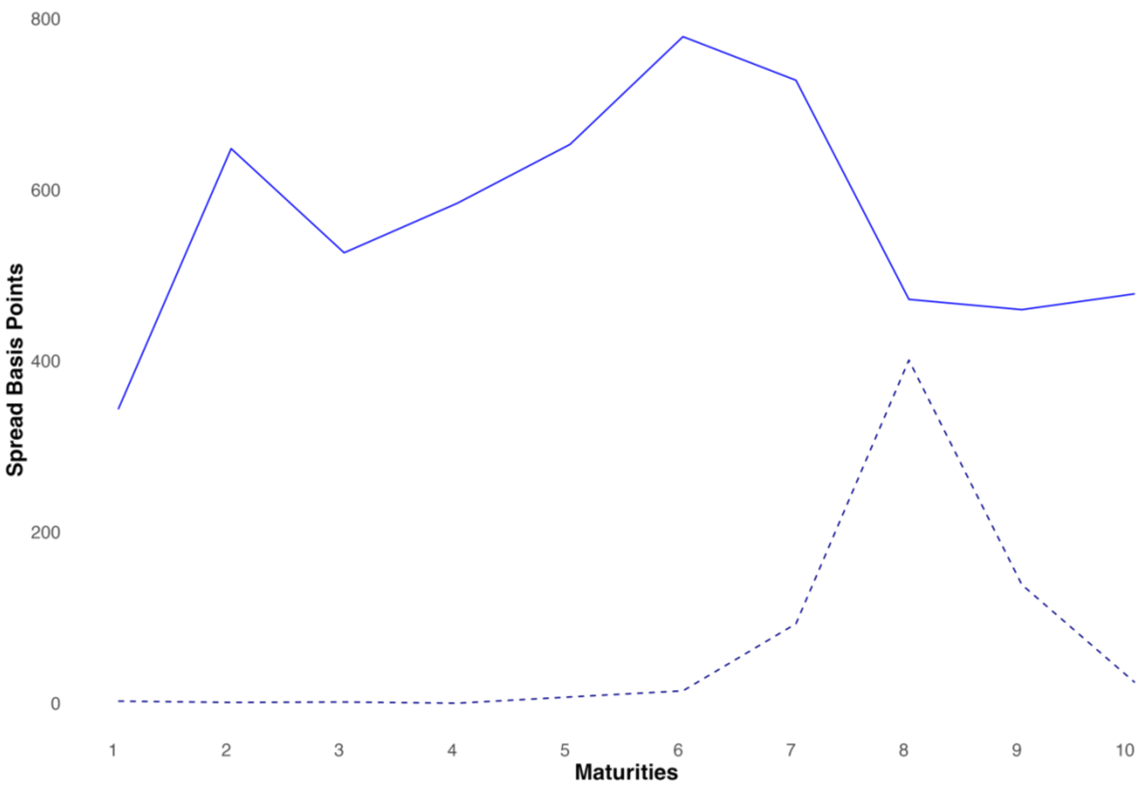
Figure 13: Actual spreads vs modeled spreads using the M model. Actual spreads are in solid line and modeled spreads are in dashed line. Numbers in basis points.

Year	Mean Actual Spread	Mean M Spread	Std	Min	25th percentile	75th percentile	Max
2015	905.05	32.08	30.98	-12.93	27.28	28.59	80.85
2016	2459.21	25.07	26.36	-14.40	-4.11	39.42	68.47
2017	1655.75	2.89	12.25	-10.75	-5.09	21.16	23.86
2018	509.42	7.91	17.92	-31.11	-4.62	23.76	52.85
2019	394.73	124.11	675.88	-40.94	-5.76	24.39	7702.33
2020	711.74	11.73	23.88	-40.76	-5.39	24.14	174.59
2021	398.36	50.53	311.25	-42.86	-5.73	26.08	3621.75
2022	429.74	8.08	18.27	-46.76	-5.96	24.75	67.37
2023	360.86	5.43	21.77	-40.57	-8.09	23.83	75.85



To benchmark the results with Eom et al. (2004) we also look at spreads in with the bond's maturity. For high-yield bonds, they find extremely large spreads for short-maturity bonds, and the spread estimation gets more accurate as the maturity increases. We find some of the same effects for bonds with maturities up until 8 years, and then it decreases substantially in the following maturities.

Figure 14: Spreads measured across different maturities using the M model. Actual spreads are in the solid line and modeled spreads are in the dashed line. Spreads in basis points. Maturities in years.

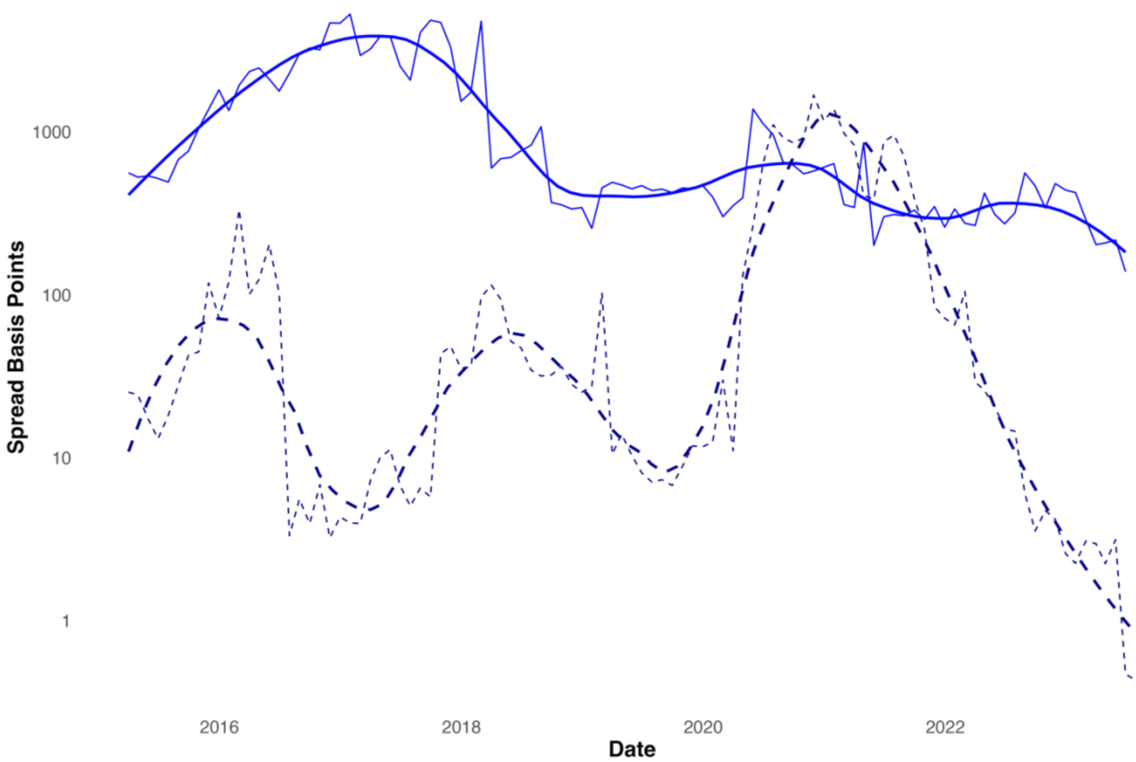


Looking at Figure 15 we observe that the LT model for most of the sample period underpredicts spreads when compared to the actual spreads. This is in sharp contrast to the previous literature of Eom et al. (2004), where they find that the LT model overpredicts spreads. From 2020-2022, we observe that the model sharply rises from underestimation to overestimation of spreads. From 2022 to the end of the estimation period, the spreads drop. This movement directly coincides with the inverse of the interest rate environment in the market, in the NSS yield curve presented in Appendix A2.

We find large outliers and standard deviations, in line with the research of Eom et al. (2004). The standard deviation of the mean spreads for LT is the largest of the models for almost every year, and this is evident by the large difference between minimum and maximum values, and even the 25th percentile and the 75th percentile.

Figure 15: Actual spreads vs modeled spreads using the LT model. Actual spreads are the solid line and modeled spreads are the dashed line. Numbers in basis points.

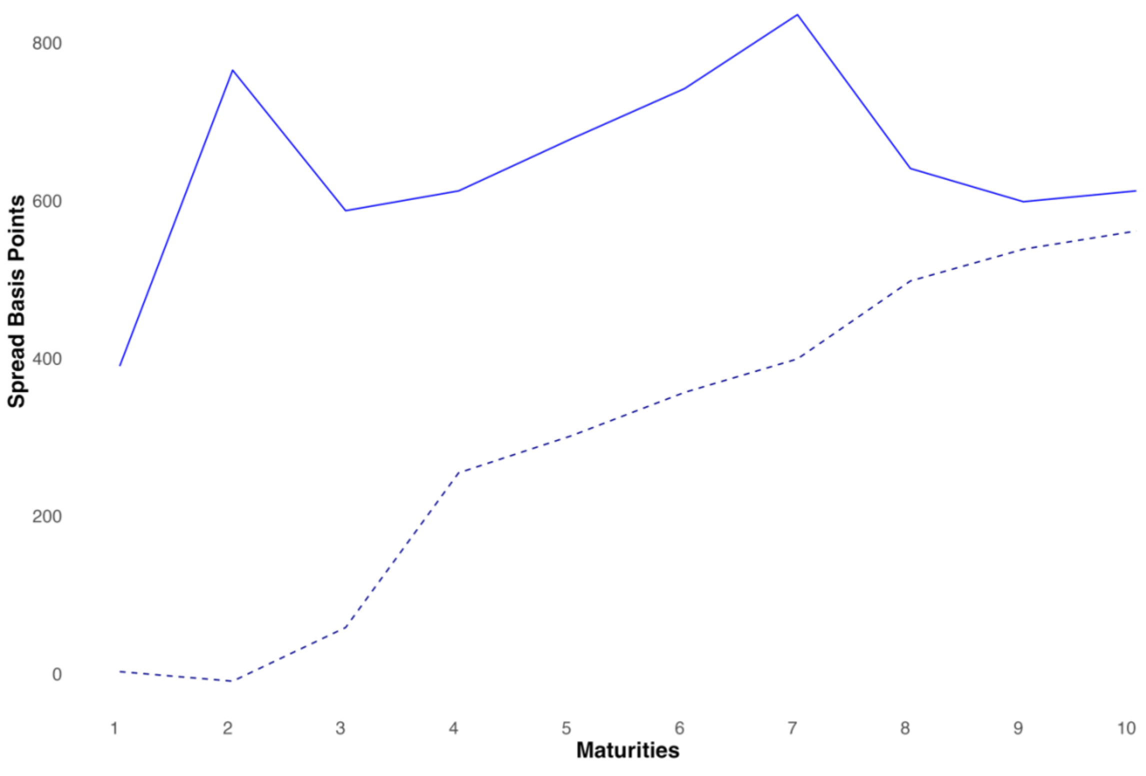
Year	Mean Actual Spread	Mean LT Spread	Std	Min	25th percentile	75th percentile	Max
2015	905.05	239.88	204.11	-86.96	95.76	443.38	504.43
2016	2459.21	152.17	215.19	-202.89	-96.98	322.56	342.04
2017	2848.55	246.36	175.27	14.15	160.52	316.53	741.86
2018	784.37	266.77	233.58	-183.67	102.06	476.55	792.02
2019	453.07	344.23	184.96	-248.19	185.69	498.98	632.60
2020	631.76	84.28	341.15	-1062.64	-109.70	405.31	991.19
2021	360.97	17.42	279.52	-646.03	-205.88	224.54	738.75
2022	340.85	50.40	200.57	-194.33	30.51	80.20	807.91
2023	225.70	10.30	98.20	-10.23	5.13	13.12	542.94



Looking at the relationships between maturities and spreads for the LT model, our results are in sharp contrast to what is found in Eom et al. (2004), and other previous literature. The LT model strongly underpredicts spreads for shorter maturities, however, the degree of underprediction decreases substantially for longer maturities.

Previous empirical literature, such as Eom et al. (2004), notes that the continuous coupon for high-coupon bonds is the reason for the consistent overprediction of spreads for the model on high-yield bonds. We investigate this effect and find that the model consistently underpredicts both high and low coupons, however, the underprediction decreases for higher coupons.

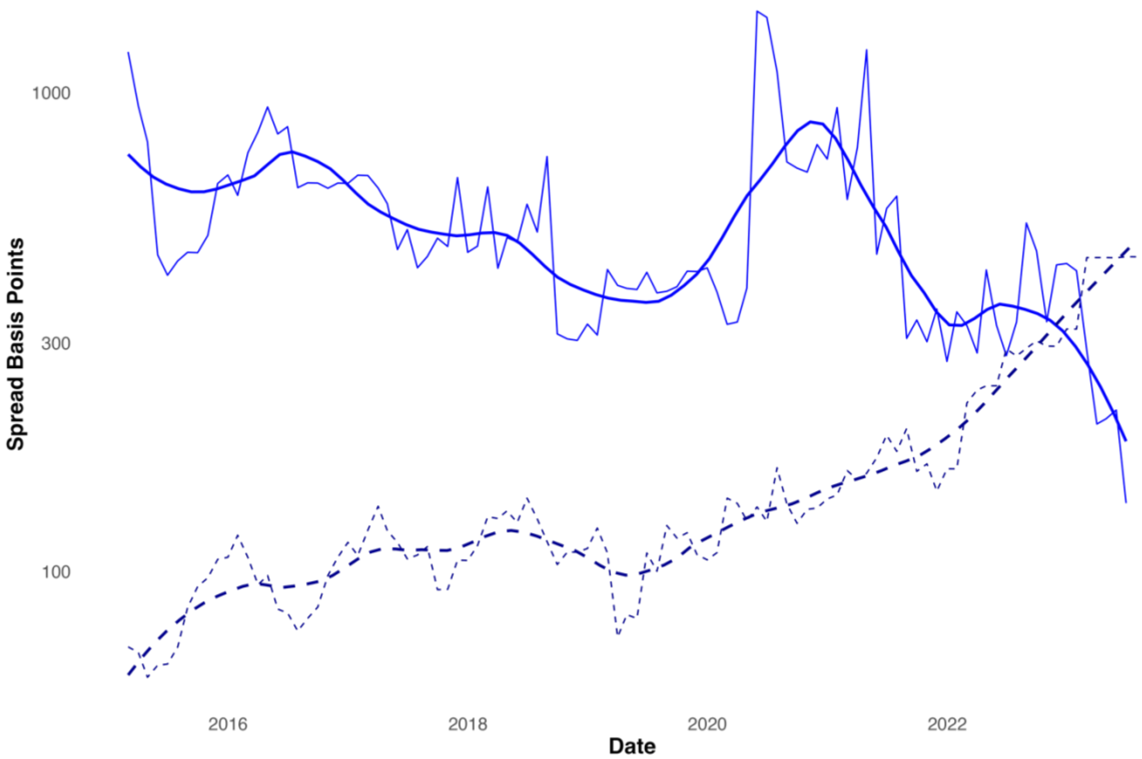
Figure 16: Spreads measured across different maturities using the LT model. Actual spreads are the solid line and modeled spreads are the dashed line. Spreads in basis points. Maturities in years.



We also find differing results from previous research when conducting the visual inspection of the LS modeled spreads against actual spreads in Figure 17. We observe the consistent underprediction of spreads by the LS model until the end of the sample period, where the corporate bond spreads are overpredicted. As is evident in the visual representations and summary statistics, there is a positive trend in the mean LS-modeled spreads. The max spreads of the model remain remarkably consistent throughout the sample period, the same is the case for the standard deviation.

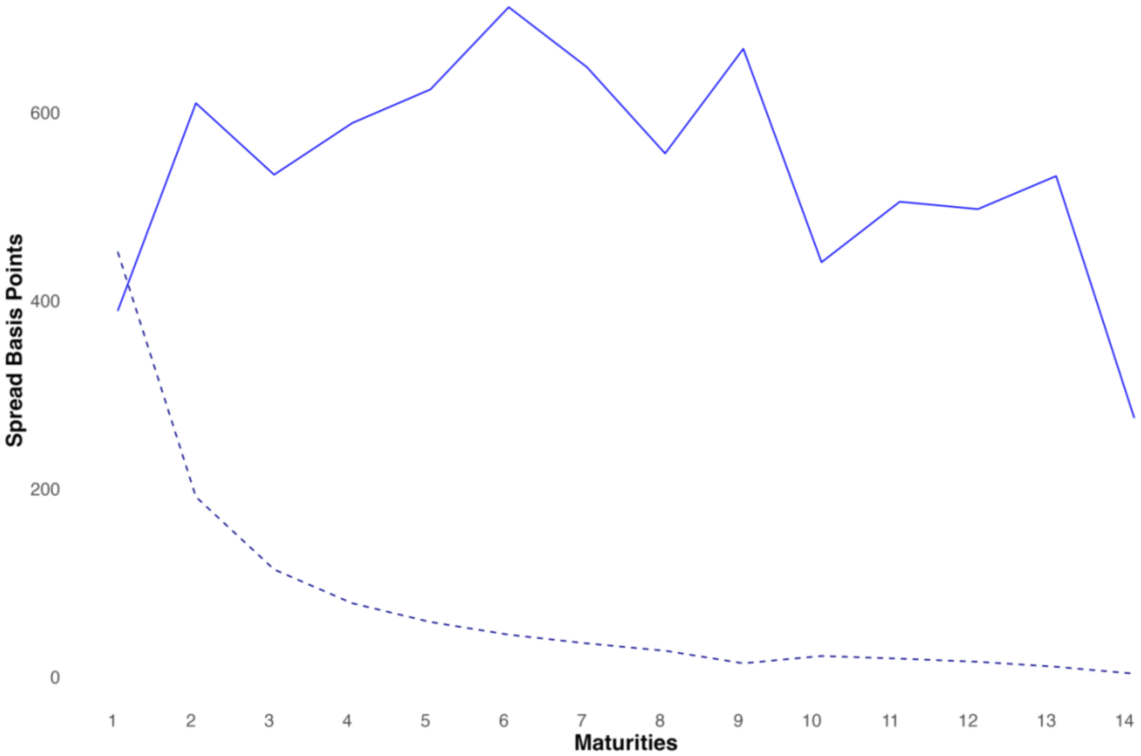
Figure 17: Actual spreads vs modeled spreads using the LS model. Actual spreads are the solid line and modeled spreads are the dashed line. Numbers in basis points.

Year	Mean Actual Spread	Mean LS Spread	Std	Min	25th percentile	75th percentile	Max
2015	574.05	86.18	108.08	-2.26	28.55	78.39	455.14
2016	737.57	94.44	112.91	14.94	30.87	113.20	455.56
2017	547.53	111.67	120.90	12.26	37.61	114.27	452.68
2018	505.86	124.04	129.94	16.95	38.02	190.17	454.29
2019	409.83	105.94	133.94	-65.24	30.15	113.82	454.93
2020	741.24	138.74	127.91	29.43	58.08	190.07	463.54
2021	542.63	167.23	132.72	44.98	78.49	191.95	455.09
2022	351.88	275.45	149.15	79.94	115.46	452.17	455.02
2023	225.70	452.64	2.00	449.45	449.81	454.26	454.63



For the shortest maturity bonds, we find that the LS model overpredicts spreads, however, as is evident by Figure 18 we find that the model consistently underpredicts spreads for all other maturities. This is also in contrast to the findings of Eom et al. (2004), in which they found that LS consistently overpredicts spreads for longer maturity high-yield bonds.

Figure 18: Spreads measured across different maturities using the LS model. Actual spreads are the solid line and modeled spreads are the dashed line. Spreads in basis points. Maturities in years.



5.2 Comparing The Models

The visual comparison in the previous section speaks to the general trend of underprediction of spreads by the models. However, mean average spreads for the sample of bonds are not particularly informative for comparative purposes. To compare the M model to the other two structural models, we use the comparative framework established by Eom et al. (2004), including a metric to measure the effect of large errors. In addition, this allows us to compare with the results found by Eom et al. (2004). To better follow this section, consider the value 0 as the best possible result for the models. The formulas for these measures can be found in Appendix B1.

Table 5: Error estimation between modeled spreads and actual spreads.

	M	LT	LS
Mean Percentage Pricing Error	-27.86%	222.11%	-88.34%
Mean Absolute Percentage Pricing Error	50.71%	348.18%	94.41%
Mean Spread Percentage Error	-93.86%	-53.49%	-38.16%
Mean Absolute Spread Percentage Error	99.41%	93.10%	97.65%
Root Mean Squared Spread Percentage Error	103.43%	134.16%	185.20%

Table 5 displays a thorough comparison of the models' pricing errors. Mean percentage pricing error tells us how accurate the model pricing is on average, and simple economic intuition tells us that the M model, prices the bonds 27.86% lower than real market prices on average. Looking at the mean absolute pricing error measure we observe that the measure almost doubles when we isolate and homogenize the errors. The mean spread percentage error is close to 93.86% less than the actual observed spreads, suggesting that the model severely underpredicts spreads. This relationship is interesting as the mean percentage pricing error suggests that the model underpredicts spreads. The mean absolute spread percentage error has the lowest variation of the three models, which suggests that there are not many large positive errors from the model that increase the perceived overall inaccuracy for the mean spread percentage error. The high absolute spread percentage error shows that the model consistently underpredicts spreads. Root mean squared percentage error (RMSE) is included in the model to investigate and highlight the effect of large errors and observe that the M model has the lowest overall RMSE.

Looking at the mean percentage pricing error, the LT model appears to be consistently overpricing corporate bonds in the sample. The mean absolute percentage pricing error captures the fact that there are also large negative errors, meaning that the LT model also finds prices for bonds that are substantially lower than real market prices. Note both its mean error, and absolute mean error, which is the largest of the three models for structural bond pricing. The mean spread percentage error of -53.49% is the second-best mean spread error in the sample, however by also inspecting the mean absolute spread error we can see that this model suffers from variation. The observant reader might notice something peculiar about these results. The table shows that the model overpredicts bond prices, while it underpredicts spreads when compared to NBP's spreads. The reason for this lies in the risk-free rates used to calculate the spreads and different measures of spread calculation. The average bond prices are therefore three times the actual prices, while the spreads still are lower because we calculate higher risk-free rates. As described by Eom et al. (2004), we find that there are some extreme value outputs by the LT model, and this is evident by looking at the RMSE measure. The RMSE of the model is larger than that of the M model, but interestingly it is smaller than what is found by the LS model.

The LS model finds corporate bond prices that on average are 88.34% lower than the actual traded prices. Additionally, the mean of the absolute percentage pricing error tells us that the mean is not as skewed by outliers as the others and is centered around the mean. This is evident by the summary statistics in Chapter 5.1, where we find that the model has consistent max values for all sample years of close to 450 basis points. The mean spread percentage error finds that the modeled spreads are -38.16% lower than observed market spreads, which is the best estimation of all the models. The mean absolute spread percentage error is close to the other models, which speaks to the fact that there are large positive errors that influence the spreads, as was the case for the LT model. This is also evident by the RMSE, which is the largest in the sample, confirming the existence of outliers that are affecting the modeled spreads. The reader might initially be puzzled as to why the M model is the most accurate at pricing, and the LS model is better at spread prediction. However, the reason for this is natural, and at the core center is the role of the risk-free rate in the calculation of spreads. As is described in the methodology section, the spreads are calculated with different risk-free rates, M with the NSS rate, and LS with the Vasicek rate. This means that the bond prices are not subtracted by the same value of risk-free bonds, which in turn creates a difference in comparison. This was also the case in Eom et al. (2004).

Turning to the comparison of the results from the models and those of Eom et al. (2004) we find substantial differences. Whereas the models in Eom et al. (2004) find pricing errors in the range of -1.97% to 1.69%, our findings are in the range of -88.34% to 222.11%. This speaks to more volatility for our pricing models than theirs, in addition, their models are substantially more accurate in pricing the bonds in their sample. The same result is found in their mean absolute percentage pricing error, their worst-performing model is still more accurate at pricing their bonds than our best-performing model by a difference of 38.07 percentage points. However, when it comes to spreads there is a lot more comparability between our findings. Eom et al. (2004) find mean spread prediction errors of -50.42% for the M model, while we find -93.86%, which is still worse than what they find by a relatively large margin. The LT and LS models respectively outperform their counterparts in Eom et al. (2004) by 62.2, and 4.77 percentage points. Interestingly, as was discovered in Chapter 5.1, they find positive spreads indicating overprediction, while we find negative spreads indicating underprediction. The mean average spread prediction error finds that there are larger errors in our modeled spreads than in their research.

6. ANALYSIS

Following our comparison of the M, LT, and LS models' predictive strengths in Chapter 5, we now narrow our focus to the M model exclusively for an in-depth analysis. The next chapter is dedicated to uncovering the reasons behind the credit spread puzzle, which is the gap between actual spreads and those projected by the M model. Reflecting on recent studies that analyze the credit spread puzzle using an extended version of the Merton model, we apply a similar methodology, allowing us to benchmark our results against the findings of these studies.

We introduce a new variable named "Mispricing," defined as the discrepancy between the actual and modeled spreads.[8] Subsequently, we conduct regression analyses on this mispricing variable to identify potential causes of the spread differentials. Our research, and the significance of its results, are dependent on it being unbiased and that we avoid econometric pitfalls. We are dealing with time series data, and as such, it is crucial to address issues such as autocorrelation and heteroskedasticity to ensure the robustness and validity of our findings.[9]

The sample consists of both time-variant and time-invariant variables, and all these variables can be used as covariates to the dependent mispricing variable. Time-variant variables are variables that vary across time for each entity, such as liquidity and market factors. Time-invariant variables are variables that do not vary across time for entities but still is different between each entity, such as industry type or return type. In this analysis, we run separate regressions for time-variant and time-invariant variables. By keeping these variables separate, we can run time-variant variables in a panel data regression, making us able to effectively secure the coefficients for potential unobservable biases that vary across time and ISINs. The time-invariant variables are run in a regular multivariate OLS regression. In this regression, we do not group the variables by their respective ISIN, letting us effectively catch the variance of spreads and time-invariant variables in the dataset. Separating the variables enables us to get unbiased coefficients and at the same time compare our findings

[8] This approach is the inverse of Eom et al. (2004), and Langdalen & Johansen (2016). When their coefficients are negative, and ours are positive, we find that coefficients influence mispricing similarly.
[9] As seen in Chapters 3 and 4, we calculate semi-annual bond prices and spreads for all 141 bonds.

to the research of Ytterdal & Knappskog (2015), who study Nordic high-yield bonds, and Langdalen & Johansen (2016), who study corporate bonds in the Norwegian market.

6.1 Two-Way Fixed Effects Regression

The regression presented in Table 6 employs a two-way fixed effects regression model within a panel data framework, accounting for both entity-specific (each ISIN is an entity) and time-specific unobserved heterogeneity. This approach is particularly appropriate for the dataset, where several explanatory variables, such as the market factors, are consistent across bonds issued by the same entity but may vary over time. By incorporating both entity and time-fixed effects, the model effectively controls for unobservable characteristics unique to each issuer that could influence the observed mispricing. This methodology ensures a more accurate estimation of the independent variables' effects on bond mispricing. The specific independent variables utilized in the analysis are detailed in the subsequent sections before the two-way fixed effects regression is presented in Table 6.

6.1.1 Variable specifications

Upon recognizing the existence of the credit spread puzzle within the M model in Chapter 5, we have expanded our analytical scope by introducing several variables in Table 6. These variables are organized into groups: bond-specific determinants aimed at forecasting spreads as per the M model, liquidity variables, and variables for market risk. The inclusion of liquidity and market variables is in accordance with Andreassen & Semmen (2023), who suggest that credit spreads are reflective of a liquidity premium as well as a market risk premium.

Both time to maturity, asset volatility, leverage ratio, and payout ratio are used to calculate the modeled spreads as explained in the methodology chapter. We include these bond-specific variables in the regression to see if they have an impact on the mispricing variable. The general idea is that if these variables prove to be statistically significant as explanatory variables for the mispricing variable, then the M model does not fully account for this.

Amihud & Mendelson (1986) were among the first to suggest that investors demand compensation for illiquidity. This has garnered a lot of attention, and subsequently research into direct and indirect measures that attempt to capture the effect of liquidity. Sæbø (2015) and Ytterdal & Knappskog (2015) find in their analyses that liquidity explains a large part of the credit spread puzzle on Norwegian corporate bonds. Additionally, Feldhütter & Schaefer (2018), who examine corporate bonds issued in USD, find a strong relation between average pricing errors and bond illiquidity in speculative HY bonds specifically, suggesting that the model underprediction for HY bonds is subject to liquidity premium. To investigate the impact of liquidity on bond mispricing, we introduce two variables to further investigate this claim.

Longstaff et al. (2005) identify the default and non-default components in corporate spreads. They introduce the bid-ask spread of individual bonds as a proxy for the liquidity of these bonds. The bid-ask spread covariate in the dataset is calculated by subtracting the bid prices from the corresponding ask prices. These prices are collected from the benchmark prices provided by NBP. As we treat the NBP prices as the actual high-yield bond prices, the bid-ask spread serves as a proxy for the actual liquidity in the market. This way we can examine if liquidity risk is contributing to mispricing.

Bao et al., (2011) argue that while the bid-ask spread in certain conditions can be an appropriate measure of bond liquidity, it fails to capture the bond market trends. They introduce a different proxy, one which reflects the impact of bond market illiquidity on bond prices. We include this proxy and estimate it as:

$$Illiquidity = -Cov(P_t - P_{t-1}, P_{t+1} - P_t) \tag{6.1}$$

Bao et al., (2011) find that the variable can explain an important portion of the yield spreads of corporate bonds, and that it exceeds the explanatory power of the bid-ask spread. The difference in approach is that it captures the covariance in month-to-month prices, and uses this as a proxy for illiquidity

A field with extensive research in finance is risk premium for asset pricing. Fama & French (1993) proposed three factors to describe systematic risk: size, value, and a market factor. The three-factor model for bond spreads was further expanded upon by Carhart (1997) which introduced a momentum factor. To investigate the effects of systematic risk on corporate bond spreads, we look to Elton et al. (2001). They find that a large part of the spreads on corporate bonds can be explained by the three-factor model. In Table 6 we include factors for size, value, and momentum. However, the coefficients should be interpreted differently from the coefficients of the bond-specific variables. The market factors are not endogenously determined by the M model and will, by showing significance, be interpreted as explanatory variables for a potential credit spread puzzle.

To incorporate size as an explanatory variable we create a dummy where companies with a market capitalization greater than the median of each year, are 1, and where companies below the median are 0. We use the median to account for market capitalization outliers in the sample. To capture the growth factor of the firms, we estimate the price/book (P/B) ratio, a widely used ratio for valuation purposes. The P/B ratio is computed by dividing each company's market capitalization by their respective book value of equity at the corresponding time. To calculate the momentum variable, we create a dummy variable by assigning 1 to the top 25 percentile of returns on market capitalization for a 6-month, and 0 to the remaining companies not included in the 25 percentiles. Note that while most research on momentum looks at the momentum of both winners and losers, the dummy only captures whether firms are winners or not.

Incorporating multiple covariates in regression analysis inherently raises the risk of correlation and multicollinearity among the independent variables. Our analysis reveals a substantial degree of correlation, with the most pronounced being between the payout ratio and the price/book ratio, which exhibit a correlation coefficient of 0.935. This high correlation indicates the presence of imperfect multicollinearity, which, while not introducing bias into the coefficient estimates, inflates their standard errors. This inflation may adversely affect the t-statistics associated with these coefficients, making the coefficients difficult to interpret. [10]

To assess the need for the potential removal of these variables, we use the Volatility Inflation Factor (VIF) to gauge the severity of multicollinearity among the variables. The VIF values for the payout ratio and the price/book ratio are 8.02 and 8.00 respectively. This is in stark contrast to all other variables, which have VIF values below 2. Rohrer (2020) suggests a VIF

[10] See Appendix D1 for the full correlation matrix.

threshold of 10 as a rule of thumb for the exclusion of variables from regression analysis due to excessive multicollinearity. Given that the VIF values fall below this threshold, we decided to keep all variables in the regression, as detailed in Table 6. This decision is predicated on the understanding that while multicollinearity does increase the standard error, it does not make the coefficient estimates biased, as long as the VIF remains within an acceptable range.[11]

6.1.2 Bond-specific variables

Table 6 displays that the time to maturity variable is not statistically significant in any of the five models. The interpretation is that this variable is not a driver of the mispricing variable, and this suggests that we have successfully incorporated this variable in the pricing of credit risk through the model. Asset volatility is not statistically significant in any of the models apart from model 3, where it becomes significant at the 10% level after controlling for the illiquidity variable. The potential explanation for this is described in Chapter 6.1.3. This observation implies that the M model sufficiently considers asset volatility when comparing actual market spreads to modeled spreads. However, the insignificance of asset volatility in our findings presents a contrast to the results of Eom et al. (2004), who identified consistent significance for this variable using a similar computational approach. It is important to note that Eom et al. (2004) price bonds annually and thereby have fewer data points in their analysis, which could potentially lead to more pronounced variations for their asset volatility variable.

Both the leverage ratio and payout ratio are statistically significant on the 1% level. However, while the coefficient for the leverage ratio is positive throughout all models, the coefficient for the payout ratio is negative in all models. The interpretation of the significance of these independent variables might be that the actual observed spreads are more affected by these variables than what the theoretical M model accounts for. A one-unit increase in leverage ratio (a 100% increase) raises the mispricing variable in the range of approximately 54 – 66 basis points across all five regression models. On the other side, a one unit increase in payout ratio (a 100% increase) decreases the mispricing variable in the range of roughly 498 – 576 basis points across all regression models in Table 6. The effect of a one-unit increase in payout ratio is substantially larger than the equivalent increase in

[11] See Appendix D2 for the full VIF estimation table.

leverage ratio, measured in absolute terms. However, the significance of this coefficient could be distorted due to the coefficient being subject to imperfect multicollinearity, as discussed in Chapter 6.1.1.

Comparing our findings with those of Langdalen & Johansen (2016) and Eom et al. (2004), both studies identify the payout ratio as statistically significant. Furthermore, our analysis diverges regarding the leverage ratio, where we observe a consistent positive influence that is not present in their study. Moreover, the magnitudes of their coefficients are generally more pronounced than those in our analysis, suggesting that their bond pricing model may not account for these factors as effectively as ours does. This difference could suggest that the model may offer a closer approximation of the real-world impact of these variables on bond mispricing.

6.1.3 Bond liquidity premium

The bid-ask spread variable is included in Models 2 and Model 6 of the regression analysis where we also control for bond-specific risk and all other covariates respectively. As exhibited by Table 6, this independent variable is not statistically significant in either of the models, in contrast to the results of Ytterdal & Knappskog (2015) and Langdalen & Johansen (2016) where they find liquidity to be a significant factor. Interestingly, the insignificance of the bid-ask spread in explaining the mispricing variable is the same as suggested by Bao et al., (2011) in previous research on the topic.

Because the bid-ask spread was insignificant, we introduce the illiquidity variable introduced by Bao et al. (2011) to investigate liquidity further. The variable is estimated on 5-months benchmark bond price movements, as provided by NBP. Model 3 exhibits that a one percent increase for the illiquidity covariate increases the dependent mispricing variable by only 0.0001 basis points, however, it is also not statistically significant. Although not displayed in Table 6, we incorporated an approach for the illiquidity variable using a 10-months measure instead of 5 months. This change resulted in the variable remaining statistically insignificant when regressed on the mispricing variable when also controlling for bond-specific variables. Interestingly, our results differ from those of Bao et al. (2011). This may be because their research was on day-to-day prices, while our results are obtained from 5–10-month periods, meaning that daily variations in prices better reflect illiquidity in the bond market. However, what is the most likely cause of the difference in results is that their modeling is based on American IG bonds. IG and HY bond markets have different credit risks, and our analysis points to the fact that the illiquidity factor is not a driver of the mispricing variable.

Our study offers a distinct perspective from previous research on the credit spread puzzle in Norwegian high-yield bonds, particularly with liquidity measures like bid-ask spread and illiquidity not being statistically significant. Yet, an intriguing aspect emerges in Model 3: when accounting for illiquidity, asset volatility becomes statistically significant on the 10% level. This could suggest that the illiquidity variable captures some of the inexplainable variation attributed to asset volatility, implying that controlling for illiquidity uncovers asset volatility's true impact on mispricing. This outcome is intriguing, especially since incorporating bid-ask spread in Model 2 does not yield a similar effect, despite it also being a liquidity indicator based on the same actual spreads. A plausible reason might be the longer time frame of the illiquidity calculation, spanning a 5-month period, which potentially aligns more closely with the monthly measurement of asset volatility compared to the daily measured bid-ask spread. However, note that the value of the correlation coefficient between the two variables is 0.001 which suggests that there is a minimal direct linear relationship between the illiquidity variable and asset volatility. The revelation of a statistically significant asset volatility coefficient could indicate that the illiquidity factor is correlated with an omitted variable, but since we control for this with two-way fixed effects regression this is probably not the case. This does not disprove that illiquidity uncovers the effect of asset volatility on the mispricing variable, however, what may be the reason for asset volatility becoming statistically significant is the difference of observations between models. We observe a substantial drop in observations, close to 15% of the total observations, between Model 2 and Model 3, and an increase in observations from Model 3 to Model 4. Due to the estimation period of the illiquidity variable, there are fewer observations when including it in the regression. In summary, although asset volatility appears to be statistically significant, this could be attributed to a smaller number of observations. In addition, it is important to note that the statistical significance is only moderate, indicated by a 10% confidence level.

Notably, both Models 2 and 3, which include liquidity measures, exhibit higher adjusted R^2 than Model 1, with Model 3 showing the highest adjusted R^2 of all. This indicates that including liquidity factors in the analysis could be crucial, despite these factors not showing statistical significance on their own.

6.1.4 Market risk premium

The results from Models 5 and 6 regarding the market factors reveal a compelling narrative: issuers with strong recent stock performance, captured by the momentum variable, are associated with higher mispricing, widening the spread difference by approximately 75-77 basis points. This finding offers a contrast to the earlier studies by Ytterdal & Knappskog (2015) and Langdalen & Johansen (2016), where the momentum factor was absent, and size and value were identified as principal influences on bond spreads. The significant momentum coefficients, in addition to the correlation matrix, suggest that the momentum factor captures, and is potentially the main reason for the effect size has had in previous research. This could imply that momentum, likely influenced by market sentiment or speculative trends, applies a more immediate and pronounced effect on bond mispricing than the fundamental and slower-evolving variable size. Our study, therefore, extends the existing literature by highlighting the importance of considering momentum, which may act as the actual driver behind the mispricing out of the market factors, especially when size does not fully explain the variations in bond spreads.

As for the value factor P/B in Model 4 and Model 5, we find that it decreases the mispricing of the model respectively with a value of -0.233 and -0.220 basis points. Whereas previous research has found P/B to be statistically significant, interestingly our analysis finds that the coefficient is statistically insignificant. Looking at the correlation matrix we can see that there is a strong direct relationship between the P/B and the payout ratio variables, with a value of 0.935. The relationship suggests that when the P/B of the firm increases, the payout ratio is likely to increase. This tendency posits that firms with higher P/B are more inclined to distribute earnings in the form of dividends or engage in share repurchase activities. This is discussed more in detail in Chapter 6.3.1.

Table 6: Two-ways fixed effects regression.

This table shows the result of the panel data regression for the mispricing variable on bond specific variables, liquidity, interest rates, and various market factors. The regression includes time fixed effects and entity fixed effects. Standard errors are robust and clustered on ISIN level to avoid problems with heteroskedastic and serial correlated error terms.

	Mispricing				
	Model 1: *Bond Specific*	*Model 2:* *Bond Specific + Bid-Ask Spread*	*Model 3:* *Bond Specific + Illiquidity*	*Model 4:* *Bond Specific + Market Factors*	*Model 5:* *Bond Specific + Market Factors + Bid-Ask Spread*
Time To Maturity	4.129	9.056	20.668	2.440	7.292
	(25.427)	(23.902)	(22.736)	(26.240)	(24.724)
Asset Volatility	-0.930	-0.891	-1.107*	-1.001	-0.974
	(0.606)	(0.583)	(0.457)	(0.540)	(0.514)
Leverage Ratio	54.505***	66.148***	62.242***	53.950***	65.545***
	(14.547)	(18.126)	(14.362)	(15.152)	(18.688)
Payout Ratio	-549.245***	-545.527***	-576.108***	-497.956***	-498.254***
	(72.272)	(72.822)	(84.140)	(40.622)	(40.546)
Bid-Ask Spread		-34.141			-35.621
		(23.218)			(23.317)
Illiquidity			0.0001		
			(0.0002)		
Price/Book				-0.233	-0.220
				(0.342)	(0.345)
Size				-6.855	-43.732
				(57.300)	(55.007)
Momentum				75.130*	77.189*
				(34.634)	(33.983)
Fixed Effects	Two-ways	Two-ways	Two-ways	Two-ways	Two-ways
Observations	2,575	2,575	2,173	2,531	2,531
R^2	0.293	0.314	0.384	0.294	0.317
Adjusted R^2	0.235	0.257	0.328	0.234	0.258
F Statistic	246.474*** (df = 4; 2378)	217.772*** (df = 5; 2377)	248.467*** (df = 5; 1992)	138.878*** (df = 7; 2331)	135.146*** (df = 8; 2330)

6.2 Multivariate OLS Regression

Table 7 displays a multivariate OLS regression that includes time-invariant factors to improve our understanding of their effect on bond mispricing. This method clarifies the role of features that stay the same for the duration of the bond's life. Fixed effects are not used in this regression. This is because variables that do not change over time for each bond would overlap completely with fixed effects, which means we would not be able to measure their impact on mispricing.

The OLS regression utilizes bond characteristic variables from the Stamdata database, which are transformed into dummies for analytical purposes. We include a constant in the regression to avoid falling into the dummy variable trap. Yet, potential correlations among these covariates could introduce bias. When assessing multicollinearity, the VIF estimator is unsuitable due to the categorical nature and multiple number of dummies for each covariate, like sectors. Instead, we employ the Generalized Variance Inflation Factor (GVIF) which adjusts for degrees of freedom in $GVIF^{1/(2*DF)}$. Buteikis (n.d.) suggests that this estimator must be squared to make it comparable to the VIF estimator, and thereby adhere to a similar rule of thumb of $(GVIF^{1/(2*DF)})^2 < 10$ as a threshold to not exclude covariates due to extensive multicollinearity. In Appendix D3 we find all variables in Table 7 to fall below this limit, affirming no significant multicollinearity concerns within our regression.

Turning to Table 7, the coefficients for the bond-specific variables are somewhat different from their corresponding coefficients in Table 6. Although asset volatility is not significant across these models, time to maturity has some statistical significance in Models 6 and 7, whereas it turns insignificant when including all control variables in Model 8. Models 6-8 still display a strong statistical significance for leverage ratio and payout ratio after controlling for time-invariant covariates. This substantiates Table 6 by concluding that these variables are not being sufficiently accounted for in the M model, as opposed to their importance in actual spreads. This is in line with the research of Eom et al. (2004), Langdalen & Johansen (2016), and Ytterdal & Knappskog (2015), all implementing versions of the Merton model, who find no changes to statistical significance when introducing several of the same control variables in the regression.

6.2.1 Bond type risk

The coefficient for green bonds, which is a dummy variable with a value of 1 when the bond is green and 0 otherwise, is negative and statistically significant, implying that green bonds reduce mispricing. This is in line with economic intuition, stating that green bonds should have lower spreads because the perceived risk of holding these is also lower. Schoenmaker & Schramade (2019) find that 82% of all outstanding green bonds are issued in the IG market and that the energy sector is the main industry involved, based on green bonds issued between 2007 and 2017. However, the industry sector is the main green bond issuer in the sample, where none of the high-yield green bonds in this analysis is subject to Oil and Gas E&P, Oil and Gas Services, or the Shipping sector.[12] This is intriguing since when controlling for these industry groups in Model 8, both the coefficient for green and the coefficient for the industry sector becomes insignificant at the same time. Firstly, we see that the industry sector itself does not contribute additional explanatory power to the analysis regarding mispricing in Model 8. Secondly, the explanation as to why the green coefficient is not significant might be because it predominantly is subject to a sector that does not significantly affect mispricing. Our findings of insignificance for Norwegian green-labeled bonds align with the findings of Wensaas & Wist (2019) who find a green-premium in the Nordic market only for bonds issued in Swedish Krona and for IG bonds from 2013 to 2016. It is also important to note that our analysis contains a clear limitation of having a green label only attached to bonds in the industry sector and not across the three other sectors.

In Model 6, senior unsecured high-yield bonds are shown to reduce mispricing by about 104 basis points, and this effect decreases to 233 basis points in Model 8, which includes industry group controls. This indicates that the model might not fully account for the perceived lower risk associated with these bonds, seen in comparison to bonds that either have senior secured or subordinated risk type. The primary reason for including a risk type covariate in our regression is based upon the urge from Langdalen & Johansen (2016) for future research in doing so. Despite the economic expectation that senior unsecured bonds should command higher spreads due to the lack of collateral, the market seems to price them lower. This might be because issuers capable of offering bonds without collateral are often seen as financially stable, and likely to fulfill coupon and principal payments. This aligns

[12] In fact, only 10 of 141 bonds in the sample are labeled as “Green”, which is displayed in Appendix C2.

with findings from Table 6, where issuers with high payout ratios, a potential indicator of stable income, are associated with lower observed spreads.

Bonds characterized as floating rate notes are associated with a substantial reduction in mispricing of 111 basis points in Model 6 and 74 basis points in Model 8. This reflects the market's tendency to adjust for the interest rate risks inherent in floating rate notes. However, it is important to highlight that in the dataset, these bonds were treated as having fixed rates following Eom et al. (2004). This treatment contrasts with their actual behavior in the market where interest payments vary with the underlying rates. Consequently, the significant negative coefficients for the floating rate variable in the regression suggest that the M model underestimates the risk adjustments made by the market. This difference in spreads indicates that while market spreads for floating rate notes incorporate their lower interest rate risk, the M model does not. Our assumption of treating all bonds as having fixed rates makes the floating rate note coefficient less suitable for comparison with previous research.

6.2.2 Sector-specific risk

The work on credit spreads in the Norwegian bond market of Sæbø (2015) found that sector aversion could explain parts of mispricing from the extended Merton model. Following his research, we include sectors as dummy variables, like Ytterdal & Knappskog (2015), and Langdalen & Johansen (2016) also do in their papers.

Table 7 shows that all industries are consistently significant throughout both Model 7 and 8, apart from the industry sector that becomes insignificant in Model 8. These sectors are included in the regression because they represent the majority of issuers in the sample. The remaining bonds within the not-included sectors are accounted for by the constant. Hence, we see that the M model fails to fully account for the inherent risk. High-yield bonds issued by Oil and Gas Services companies create the largest mispricing variable with roughly 553 basis points in Model 7 and 663 basis points in Model 8. The effect of bonds issued by Oil and Gas E&P issuers on mispricing is 253 basis points in Model 7 and 237 basis points in Model 8. The effect is far less for the shipping sector where mispricing increases by 34 basis points in Model 7 and 137 basis points in Model 8.

Table 7: Multivariate OLS Regression

This table shows the result of the standard OLS regression for the mispricing variable on bond specific variables, bond characteristics, and industry types. The regression models report heteroskedasticity-robust standard errors to account for non-constant variance of errors across observations.

	Mispricing		
	Model 6: Bond Specific + Bond Type	*Model 7: Bond Specific + Industries*	*Model 8: All Controls*
Time To Maturity	10.192**	7.625*	6.686
	(4.653)	(3.966)	(4.707)
Asset Volatility	-0.562	-0.127	-0.405
	(1.095)	(1.118)	(1.168)
Leverage Ratio	51.018***	40.645***	27.640***
	(6.524)	(9.195)	(9.582)
Payout Ratio	-541.709***	-525.139***	-572.387***
	(68.845)	(69.323)	(68.613)
Green	-123.522***		35.474
	(38.583)		(38.927)
Senior Unsecured	-104.199***		-233.014***
	(23.625)		(23.410)
Floating Rate Note	-110.810***		-73.964**
	(26.663)		(32.094)
Industry		-58.524**	24.914
		(24.963)	(23.867)
Oil and Gas E&P		253.358***	237.481***
		(33.780)	(46.949)
Oil and Gas Services		552.534***	663.392***
		(107.028)	(107.163)
Shipping		34.276**	136.635***
		(15.507)	(16.996)
Constant	546.560***	303.836***	518.517***
	(41.497)	(29.380)	(44.977)
Observations	2,481	2,481	2,481
R^2	0.191	0.245	0.257
Adjusted R^2	0.188	0.243	0.254
F Statistic	83.246*** (df = 7; 2473)	100.248*** (df = 8; 2472)	77.627*** (df = 11; 2469)

6.3 Premiums Not Fully Reflected in the M Model

Model 1 in Table 6 displays that the explanatory power, measured by R^2 of the mispricing variable is 23.5% when controlling for time to maturity, asset volatility, leverage ratio, and payout ratio. The explanatory power increases to 32.8% in model 3, where we also control for the illiquidity variable. When controlling for market factors and the bid-ask spread simultaneously, we reach an explanatory power of 25.8%. Regarding these measures of R^2, we see that there is a minimum of 67.2% of the actual spreads that we are not able to explain, and hence there is clear evidence of a credit spread puzzle in the M model when controlling for both entity and time fixed effects. The same evidence can be gathered from Table 7, where we find a maximum explanatory power of 25.4% in Model 8 when incorporating all control variables in an OLS regression.

Combining the breakdown of the credit premium outlined in Figure 2 with the credit spread equation from Equation 2.1, we find ourselves focusing on the risk premium and liquidity premium. After accounting for the loss given default, probability of default, and time to maturity in the structural pricing, these two premiums emerge as potential key drivers of the credit premium. This chapter will therefore concentrate on exploring these premiums as possible explanations for the credit spread puzzle.

6.3.1 There is a significant risk premium

Market factors, like the ones presented by Fama & French (1993) and the momentum factor we include in Table 6, are generally perceived to have a risk premium in the stock market. However, many of the principles of these factors can also be seen to affect the bond market. Elton et al. (2001) argue that as much as 85% of the is spread not explained by default and taxes relate to market factors like size and growth. In line with economic theory, market factors are assumed to create a risk premium for bonds.

A low P/B might suggest that a firm exhibits low growth or has assets that are not being effectively leveraged, potentially indicating higher credit risk and thus a higher risk premium for its bonds. However, a low P/B can also be a proxy for financial distress, which is especially relevant for HY bonds. Bonds from companies with low P/B ratios might be seen as riskier because the market perceives these companies as closer to default, thereby requiring a higher yield. Depending on the size of an issuer, bonds might also have a risk premium that is embedded into the credit spread. Small companies typically have less diversified operations and weaker bargaining power, which can impact their creditworthiness, hence inflicting a higher risk premium on their bonds. Additionally, larger firms often have better access to capital markets and more financing options, which can lower their cost of capital and thereby reduce the risk premium required by bond investors.

However, the momentum factor emerges as a significant explanatory variable for the credit spread puzzle, whereas the size variable is insignificant. The potential explanation for this might be attributed to the comprehensive nature of the momentum factor. Stock price momentum efficiently captures real-time shifts in a company's size. As companies expand or contract, these changes are swiftly reflected in their stock prices, indicative of broader market perceptions and expectations. Therefore, the momentum factor may be effectively integrating key aspects of the size of the issuing companies, offering a more holistic understanding of the observed mispricing. However, it should be noted that including the market factors does not increase the adjusted R^2 substantially.

Both Table 6 and the correlation matrix in Appendix D1 display that the payout ratio might be acting as a proxy for the value effect within the models, overshadowing the P/B ratio's expected contribution to mispricing. This explains why the P/B ratio emerges as statistically insignificant in our analysis, as its effects are possibly being accounted for by the large negative statistically significant movements in the payout ratio.

Our findings contrast with Langdalen & Johansen (2016) which finds that both the P/B and payout ratio are statistically significant, indicating that both the payout ratio and the P/B variable have a significant effect on mispricing on their own. However, the most likely reason for the deviation is determined by the difference in sample types between our study and theirs. While their sample consists of mostly investment grade bonds, we only examine the high-yield market. The investment grade market usually holds large and robust firms that are unlikely to exhibit large shifts in P/B, while they usually pay dividends and repurchase shares in the market regardless of their earnings due to expectations. The high-yield market usually houses higher-risk issuers with P/B’s that are likely to be more volatile, and such issuers are not able to distribute as much capital. Higher-risk firms that post good earnings results and thereby increase their dividend payments or share repurchases are likely to exhibit growth in P/B as their share prices increase. This indicates that higher-risk firms that can distribute capital, are likely to be firms with high P/B ratios and consequently the P/B is captured by the payout ratio.

The regression analysis in Table 7 displays that investors demand higher premiums for investing in certain industries, which is not accounted for in the M model. The significant observed sector premiums, particularly in Oil and Gas E&P, Oil and Gas Services, and Shipping, may be attributed to the cyclical nature of these industries, which are closely tied to commodity cycles. This is reflective of the broader Norwegian economy, which is also influenced heavily by commodity market fluctuations. Langdalen & Johansen (2016) propose that the profitability of these sectors follows cyclical patterns, potentially leading to an underestimation of credit risk during prosperous periods. Investors, being aware of these cycles, might demand a higher risk premium as protection against the potential downturn of an upward cycle. This cyclical risk could be a plausible explanation for the observed risk premiums in these sectors.

However, quantifying this risk is challenging, as it is not easily measurable in practical terms. Furthermore, if this cyclical risk were the primary driver of the risk premium, we would expect to see a similar pattern in the Industry sector, which is also subject to cyclical forces. Yet, our regression analysis does not show an embedded risk premium for the Industry sector, suggesting that the fear of an ending upward cycle may not be the predominant factor explaining the risk premiums in our study. Therefore, it seems unlikely that the risk associated with the end of an upward cycle is the key factor behind the risk premium component of the credit spread puzzle, indicating the need to explore other underlying drivers of these sector-specific premiums.

Langdalen & Johansen (2016) suggest that leverage ratio might explain sector premiums, noticing a trend where certain industries consistently show high leverage ratios. However, our data presents a different picture. In the sample, the Industry and Oil and Gas Services sectors exhibit median leverage ratios above 1, markedly higher than other sectors. Despite this, the industry sector does not show significant effects in the regression analysis, and Oil and Gas Services stand out with a notably positive coefficient among sector variables. This undermines the argument that the leverage ratio alone drives sector premiums. Instead, it implies that the differences we observe in sector premiums relative to the M model are likely influenced by other factors beyond just leverage ratios. This points to the complexity of sector dynamics and the need to consider a broader range of influences when analyzing sector-specific risk premiums in high-yield bonds.

The sector premiums in our analysis might be influenced by the use of a constant recovery rate across different industries. The M model might not fully capture the varying levels of risk associated with different recovery rates in specific sectors. For instance, Ytterdal & Knappskog (2015) found that the Shipping sector tends to have significantly lower recovery rates compared to other industries. However, our regression results in Table 7 do not show a distinct pattern for the Shipping sector. Its coefficients are similar to those of the Oil and Gas E&P and Oil and Gas Services sectors, except that Shipping has the lowest coefficient among them. This similarity suggests that the sector premiums we observe are probably not solely due to the absence of sector-specific recovery rates in our model. The lack of a clear difference in the coefficients indicates that other factors, beyond recovery rates, may be driving the sector-specific risk premiums in Norwegian high-yield bonds.

Generally, there are numerous potential explanations for the origins of risk premiums. Thus, it continues to be a complex task to explain the credit spread puzzle by identifying possible risk premiums.

6.3.2 The liquidity premium is linked to the risk-free rate

In our research, paralleling the findings of Eom et al. (2004), we did not observe a significant liquidity premium within credit spreads. Their study encompassed both IG and HY bonds, incorporating the use of "Old Bonds" as a proxy for lower liquidity. [13] Similarly, Langdalen & Johansen (2016) failed to find significant evidence of a liquidity premium, despite their incorporation of both issuer-specific and market-wide liquidity measures. In contrast, Feldhütter & Schaefer (2018) and Ytterdal & Knappskog (2015) identified a liquidity premium, utilizing bond age and equity market liquidity indicators, respectively. These studies attempted to construct liquidity proxies based on actual bond transaction prices, but they were not able to do that because of the lack of access to daily traded bonds. Our study benefits from being able to leverage actual transaction data for this purpose through the use of liquidity measures based on daily quoted prices from NBP.

[13] They define "Old Bonds" as bonds that are issued five, or more, years ago. They argue that this serves as a good proxy for illiquidity since such bonds have less trading volume and are therefore less liquid.

This paper underscores the importance of a liquidity premium as a fundamental component of credit spreads. Holding less liquid bonds typically entails an opportunity cost, as it may prevent investors from engaging in more profitable opportunities. The process of buying or selling these bonds can be slow and subject to significant price volatility, impacting the time value of money and introducing additional uncertainty in terms of sale price realization. This is especially relevant in the HY bond market, characterized by a smaller pool of market participants, increased default risks, higher volatility, and greater information asymmetry. Furthermore, given Norway's heavy reliance on commodities, the liquidity of its HY bond market is likely to be correspondingly affected during economic downturns.

Consequently, one would anticipate that there is a liquidity premium within the credit spreads of the Norwegian high-yield bond market. However, as indicated by the regression results in Table 6, the M model appears to have already incorporated this liquidity premium. Rakkestad & Hein (2004) advocate for the use of swap rates over government bonds as a more appropriate proxy for risk-free rates, citing the small and less liquid nature of the Norwegian government bond market. Therefore, our reliance on Norwegian government bond rates to model term structures may have inherently included a liquidity premium within our modeled spreads.

The illiquidity present in the Norwegian government bond market suggests a parallel in the Norwegian corporate bond market. Thus, the liquidity premium embedded in our modeled spreads might reflect the liquidity premium inherent in the actual spreads. NBP calculates its spreads using 3-month NIBOR rates, which, while generally more liquid than Norwegian government bonds, may also experience liquidity constraints during market downturns. This overlap complicates the interpretation of the regression coefficients from Table 6, as isolating a liquidity premium independent of the risk-free rates becomes a complex endeavor. The lack of a clear liquidity effect in our results does not mean that liquidity is not important. Rather, it implies that liquidity is embedded within the array of factors affecting credit spreads, thereby not emerging as an isolated factor in our analysis.

6.4 Potential Biases in the M Model

When using a structural model to forecast credit spreads, there is inherent uncertainty regarding the model's effectiveness in accurately predicting these spreads. Consequently, it becomes challenging to observe if a credit spread puzzle truly exists or if the model is subject to biases that prohibit optimal spread calculation. In the subsequent paragraphs, we examine potential biases in the model by contrasting our methodology with those utilized by other researchers.

6.4.1 The leverage ratio possesses a dynamic nature

In structural models, the impact of pronounced variations in leverage ratio is a critical factor. The M model operates with static leverage ratios, assigned uniquely to each bond ISIN, which may not accurately reflect their inherently dynamic nature in actual market conditions. In reality, leverage ratios are subject to change due to strategic choices, the cyclical nature of sectors, and other market dynamics. This discrepancy is highlighted by Eom et al. (2004) and Langdalen & Johansen (2016), who emphasize the crucial role of leverage ratios in the credit spread puzzle. To address the consistent underestimation of spreads by the M model, various approaches have been suggested. For instance, Collin-Dufresne & Goldstein (2001) propose using mean-reverting leverage ratios. However, as Eom et al. (2004) point out, this method still tends to underestimate the spreads of corporate bonds. While our study adopts the methodology of Eom et al. (2004), it can be argued that the M model's reliance on static leverage ratios introduces an inherent bias. Transitioning to a dynamic approach for measuring leverage ratios could more accurately represent real-world dynamics, potentially resolving the associated risk premium and eliminating its influence as a primary factor in the credit spread puzzle.

6.4.2 The yield curve matters

The interest rate term structure represents a critical input factor for the M model. In aligning with the methodologies used by Eom et al. (2004), our approach to modeling the term structure is based on the NSS model, consistent with the original methodologies employed by the creators of the extended Merton model we use in this paper. In addition to the bias of having a liquidity premium naturally attached to it, as discussed in Chapter 6.3.2, the yield curve as modeled by the NSS model may also exhibit a bias in its fit across various maturities. Lakhany et al. (2021) point out that the NSS model's mathematical function is non-linear and possesses multiple local optima. This complexity means that even minor adjustments to a parameter can lead to substantial and unpredictable changes in how well the model aligns with actual data across different maturities.

Alternative options for determining risk-free rates include utilizing yields from the swap market or yields from AAA-rated corporate bonds. Rakkestad & Hein (2004) suggest that swap yields might be more suitable for representing long-term risk-free rates due to their relative stability. Although we followed the methodology of Eom et al. (2004) closely, using swap yields obtainable through Infront as the risk-free rate might have offered a different perspective on spreads and could potentially reduce bias in the model. AAA-rated corporate bonds could potentially have the same effect as the swap rates, however, yields from these bonds are available only to a limited extent in the Stamdata database. Most corporate bonds in this database have «shadow ratings» rather than official ratings from agencies like Moody’s or S&P, making it difficult to separate the AAA-rated bonds from the rest.

6.4.3 The payout ratio might be distorted

In our research, we have identified a distinctive relationship between firms' capital distribution policies and bond pricing. Notably, the payout ratio exhibits a consistently negative and significant association with bond mispricing. This suggests that as companies increase their capital distributions through dividends, share repurchases, or coupon payments to creditors, the observed market spreads align more closely with those predicted by the model. This alignment may reflect the market's confidence in the fiscal prudence and solvency of these firms, challenging the view that higher payouts inherently signal greater risk. The addition of the payout ratio to the M model by Eom et al. (2004) to capture this capital redistribution fails to account for most of the effect it has on investor sentiment and bond pricing.

Following the methodology outlined by Eom et al. (2004), we needed to acquire data on dividend payments, share repurchases, and bond coupon payments to create a variable payout ratio. Unfortunately, we encounter a barrier in collecting share repurchase data through Infront, which does not offer such data. Eom et al. (2004) use Compustat in their research, however, this service does not provide data on Norwegian firms. To avoid the labor-intensive process of manually checking public announcements for share repurchases for every issuer in the dataset, we apply a more streamlined approach. We decided to use the decrease in number of shares from one quarter to the next as a proxy for share repurchases, which is a measure we can obtain from Infront. However, this approach has its limitations. A decrease in outstanding shares could also result from other corporate actions, like reverse stock splits, retiring shares after mergers or acquisitions, or during restructuring. Considering that the payout ratio is a critical factor for explaining the credit spread puzzle, as exhibited by both Tables 6 and 7, an inaccurate measurement of this ratio could potentially lead to biases in the spreads predicted by the M model. Any inaccuracy could bias the M model's spread predictions. Including correctly collected share repurchase data could refine the payout ratio calculation, potentially resolving its influence on the credit spread puzzle.

6.4.4 The spread measure

There are alternative spread measurement methods not utilized in Chapter 3.4 of our study, as detailed by Chaplin (2010). The asset swap spread is derived by equating the value of fixed cash flows from a bond to the floating payments in an interest rate swap, essentially involving the purchase of a risky bond and entering into a swap agreement that yields an initial cash flow. The maturity spread contrasts the bond's yield with that of the corresponding point on the swap curve, based on the bond's maturity. However, the z-spread is the more prevalent measure. It represents the parallel upward shift required in the zero-coupon swap curve to equate the bond's market price with the present value of its future cash flows discounted at the adjusted curve. These methods are commonly applied in bond pricing practice and inherently relate to the swap market. Our analysis, which employs government bonds within the M model, does not incorporate these spread measures due to the chosen term-structure approach, which does not align with swap market-based methodologies. The bias of applying our approach is linked to the use of Norwegian government bonds as discussed in Chapter 6.4.2, and the possible bias that the illiquidity of such risk-free rates might inflict on spreads, as discussed in Chapter 6.3.2.

6.4.5 The floating note problem

Eom et al. (2004) omitted bonds with floating coupon payments from their analysis. Most bonds in the sample feature floating payment structures. Rather than discarding these bonds, which would substantially reduce the sample size, we choose to treat all bonds as if they had fixed payment structures. Yet, as shown in Table 7, there is a notable difference between the spreads predicted by the M model and the actual observed spreads, and a part of this mispricing is due to the inherent lower risk associated with floating payment structures. This observation suggests that the lower risk associated with floating rate notes is not adequately reflected in the model.

6.4.6 A constant recovery rate causes problems

In Chapter 3, we identified the recovery rate as a key determinant for modeling bond spreads, and we adopted a constant recovery rate following the methodology of Eom et al. (2004). The composition of the recovery rate, however, is complex, involving both the decrease in firm value and the deadweight loss from financial distress. Unlike Eom et al. (2004), who analyzed both IG and HY bonds, our study focuses solely on HY bonds. Given the higher risk profile of the HY bond sample, the recovery rate should logically be lower, something we have accounted for when implementing a measure of 40% (compared to 51.35% in Eom et al. (2004)). Moreover, recovery rates differ across bond ISINs, which is a variation the model does not incorporate. Therefore, the persisting credit spread puzzle may be influenced by our assumption of a uniform recovery rate.

Calculating individual recovery rates for each bond ISIN is something we considered doing. However, Ytterdal & Knappskog (2015) model recovery rates in their research but observe very low explanatory power while doing so. They argue that the recovery rate is specifically difficult, or even impossible, to model since the outcome of a default event often is a bargaining process between debt holders and debt issuers.

6.5 Potential Biases in the NBP Spreads

In our study, we have applied credit spread data from NBP as a benchmark for actual market spreads. However, it is essential to recognize that an underlying assumption in our analysis is the unbiased nature of NBP's spread calculations. This assumption is pivotal in accurately isolating the true dynamics behind the credit spread puzzle.

Our understanding of NBP's specific methodologies for determining market spreads is limited. Nevertheless, we acknowledge the strength of their data collection process. NBP's method of obtaining market data directly from brokerage houses, which serve as intermediaries in bond transactions, lays a strong foundation for generating precise estimates for market spreads.

However, it is important to note that our research does not delve into the specific impacts of potential biases in NBP's spreads on our results. This is primarily due to our limited insights into their method of modeling. Consequently, this chapter aims to offer a broader reflection on the possibility that biases in the spreads reported by NBP could distort the actual nature of the credit spread puzzle, thereby affecting our interpretation of the explanatory variables.

7. CONCLUSION

This article directly predicts spreads using a model that expands upon the Merton (1974) model, and compares it to two additional structural bond pricing models, using a sample of noncallable Norwegian high-yield bonds belonging to non-financial companies in the period from 2015 to mid-2023. In particular, we implemented an extended Merton model, as presented by Eom et al. (2004), and compared it to the Leland & Toft (1996) model, and the Longstaff & Schwartz (1995) model. With a unique dataset provided by Nordic Bond Pricing, we measure the models' accuracy in pricing and spread prediction.

Our analysis reveals substantial differences in price predictions across the three models compared to the benchmark study by Eom et al. (2004). They reported pricing errors ranging from -1.97% to 1.69%, whereas our results vary more widely, from -88.86% to 222.11%. Nonetheless, when focusing on spreads, our findings are more closely aligned. Eom et al. (2004) report an average spread prediction error of -50.42% for the extended Merton model, where our figure stands at -93.86%, representing a significantly larger deviation. In our study, the Leland & Toft and Longstaff & Schwartz models surpass their corresponding models from Eom et al. (2004) by margins of 62.2 and 4.77 percentage points, respectively. Notably, the Longstaff & Schwartz model demonstrates superior performance within our sample, yielding the smallest spread prediction error when benchmarked against the spreads provided by Nordic Bond Pricing.

Our results affirm the presence of a credit spread puzzle across all models. To delve deeper into this phenomenon, we specifically scrutinize the difference between actual and modeled spreads within the extended Merton model, aiming to align our research with previous studies into the puzzle. In contrast to prior studies, we uncover no evidence that a liquidity premium affects this puzzle, implying that such a premium may already be integrated into the spreads generated by the extended Merton model. Moreover, we find that stock price momentum plays a significant role in explaining the puzzle, while traditional market factors like size and value, which are commonly highlighted in Norwegian bond research, do not exhibit significance. Our research indicates that the momentum factor might be capturing essential elements of size, while value, due to the dynamics of firms in the high-yield sector, is captured by the payout ratio of the firms. Additionally, our observations reveal that investors seek extra risk premiums for bonds in sectors such as Oil and Gas E&P, Oil and

Gas Services, and Shipping. This suggests that the increased risk of being invested in any of these sectors is not fully captured by the extended Merton model, which is in line with findings in previous studies.

Predicting bond spreads is inherently challenging, and variables in structural bond pricing models are often affected by assumptions that limit perfect and unbiased predictions. Our analysis points to the significance of leverage and payout ratios, indicating that the extended Merton model may not adequately consider these factors. Further, the use of Norwegian government bonds as proxies for risk-free rates and the assumption of constant recovery rates could introduce biases in this model's spread calculations. Without clear insight into the methodology behind Nordic Bond Pricing's benchmark spreads it is challenging to discover the absence of inherent biases, complicating the task of isolating the credit spread puzzle.

While our research uncovers multiple explanations for the credit spread puzzle, we acknowledge the inherent difficulty in calculating a mispricing variable that accurately reflects the puzzle's true nature. The complex challenge of predicting spreads and explaining the puzzle remains a recurrent theme in the literature on the subject.

8. CRITICISMS

The scope and implications of this thesis are intrinsically linked to the quality and range of the data. Our analysis specifically focuses on the subset of bonds issued by publicly listed firms. This focus was primarily chosen to streamline the data collection process but consequently limits our insights to this segment of the market. It is important to recognize that the dynamics governing the bonds of non-publicly listed firms could diverge substantially from our findings due to their distinct financial environments.

The dynamics of bonds issued by non-public issuers may be quite different, and to illustrate this we look to the momentum factor. We find that this factor is statistically significant for publicly listed firms, however, it cannot be applied directly to non-public firms as it is modeled on semi-annual stock prices. Non-public firms can keep their performance to themselves, and such information is therefore not available to potential bond investors. As such, the conclusions from this thesis may not be indicative of the structural bond pricing models on the greater high-yield bond market.

In this thesis, we have been working with large datasets of financial data and have applied these to the models. All the models are relatively complex, however the Leland & Toft (1996) model and Longstaff & Schartz (1995) model are especially so due to their many sub-calculations to arrive at the semi-annual prices. These models were created by researchers and proven on standardized data that was typical for the firms at the time of the paper's publication, however, they are not as practical for real-life data. Some input variables in these models are fixed, and some change from period to period, and the models do not adapt well to changes in firm characteristics such as risk profile. As was brought up in Chapter 5, this led to some prices becoming incalculable, and the data was therefore omitted.

An issue that arises with this is that the comparability of individual bond spreads on a model-to-model basis is inadequate and that we were not able to compare the models on the same subset of bonds. Note that structural bond pricing models are famously inaccurate at predicting spreads of individual bonds, and therefore we did not expect this to be possible either way. However, this may be the reason why we find that the LS model is best at predicting spreads. As the sample of bonds is not the same between models, the LS model may be better at predicting spreads for firms with a certain type of data that survives being passed through the model.

This leads to a form of survivorship bias where we only look at the results where the data was calculable and discard the data that did not pass cleanly through the models. Later research on structural bond pricing models for high-yield bonds should take note of this and either make use of models that are not as rigid and allow for more extreme data conditions or find a solution on how to treat the data to mitigate the problem.

In Chapter 6, we recognize a gap in our analysis stemming from the absence of an exploration of the explanatory variables in the LT and LS models. Our decision to concentrate on the M model was natural, given its widespread discussion in existing literature which facilitates comparative analysis. Our analysis finds that the credit spread puzzle exists, yet it prompts further investigation. Despite the M model's historical precedence in academic studies, our findings suggest that later models like the LT and LS models may offer greater predictive accuracy and could potentially reshape understanding in this domain.

Future studies should broaden the scope of research to include a variety of other models, particularly those that can be applied to Norwegian benchmark spreads. There lies an opportunity to delve into whether the nature of the credit spread puzzle manifests across these alternative models as well. Such research could not only validate our current observations but also contribute to a more nuanced comprehension of credit spread behaviors and the factors that influence them.

Furthermore, incorporating a comprehensive analysis of explanatory variables in the LT and LS models could yield new insights. This would not only fill the gap identified in this thesis but also enhance the understanding of structural bond pricing models. By doing so, subsequent research could offer a transformative perspective on the mechanisms underlying the credit spread, particularly within the context of the Norwegian high-yield market.

9. APPENDIX

Appendix A1 - The Extended Merton Model.

The M model is presented in the main text as:

$$p^M(0,T) = \sum_{i=1}^{2T-1} D(0,T_i) E^Q \left[\left(\frac{c}{2}\right) I_{\{V_{T_i} \geq K\}} + min\left(\frac{wc}{2}, V_{T_i}\right) I_{\{V_{T_i} < K\}} \right] +$$

$$D(0,T) E^Q \left[\left(1 + \frac{c}{2}\right) I_{\{V_T \geq K\}} + min\left(w\left(1 + \frac{c}{2}\right), V_T\right) I_{\{V_T < K\}} \right] \quad \text{(A1.1)}$$

The equation can be separated into two parts for intuition purposes:

1. Discounted value of the expected payoff before redemption date (coupons):

$$\sum_{i=1}^{2T-1} D(0,T_i) E^Q \left[\left(\frac{c}{2}\right) I_{\{V_{T_i} \geq K\}} + min\left(\frac{wc}{2}, V_{T_i}\right) I_{\{V_{T_i} < K\}} \right] \quad \text{(A1.2)}$$

2. Discounted value of the expected payoff from the last payment (principal and last coupon):

$$D(0,T) E^Q \left[\left(1 + \frac{c}{2}\right) I_{\{V_T \geq K\}} + min\left(w\left(1 + \frac{c}{2}\right), V_T\right) I_{\{V_T < K\}} \right] \quad \text{(A1.3)}$$

The M model equation is simply the present value of the expected payoff from all future coupons and the principal. The two parts above look really alike, however, the latter include a «1» to account for the principal payment. Both parts incorporate both the probability of default and loss given default:

1.

$$\underbrace{\left(\frac{c}{2}\right)}_{\text{Expected payoff if not default}} \overbrace{E^Q I_{\{V_{T_i} \geq K\}}}^{\text{Probability of not default}} + E^Q \left[\overbrace{min\left(\frac{wc}{2}, V_{T_i}\right)}^{\text{Expected payoff if default}} \underbrace{I_{\{V_{T_i} < K\}}}_{\text{Probability of default}} \right] \quad \text{(A1.4)}$$

2.

$$\underbrace{\left(1 + \frac{c}{2}\right)}_{\text{Expected payoff if not default}} \overbrace{E^Q I_{\{V_T \geq K\}}}^{\text{Probability of not default}} + E^Q \left[\overbrace{min\left(\left(w + \frac{wc}{2}\right), V_T\right)}^{\text{Expected payoff if default}} \underbrace{I_{\{V_T < K\}}}_{\text{Probability of default}} \right] \quad \text{(A1.5)}$$

In the equations A1.4 and A1.5 we see that the model incorporates the recovery rate w, which is the same as one minus loss given default. These equations also calculate the expected payoff based on the probability of the firm not defaulting, and the probability of the firm defaulting. The probability of the firm not defaulting is determined by either $E^{Q}I_{\left\{V_{T_i} \geq K\right\}}$ or $E^{Q}I_{\{V_T \geq K\}}$, meaning that the asset value at time T for the ith coupon date is at or above the default barrier. $E^{Q}I_{\left\{V_{T_i} < K\right\}}$ and $E^{Q}I_{\{V_T < K\}}$ on the other side represent the probability of default and is a bit more of a complex measure. However, its dynamic is described by the following equation:

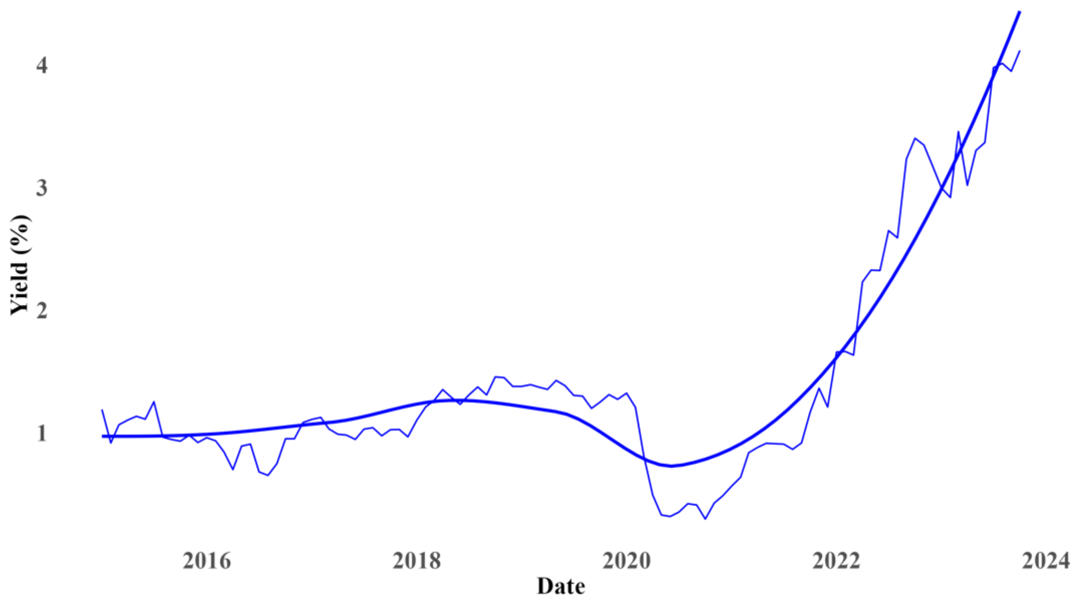
$$E^{Q}\left[I_{\{V_t < K\}} min(\text{Y}, V_t)\right] = \overbrace{V_0 D(0,t)^{-1} e^{-\delta t} N\big(-d_1(\text{Y},t)\big)}^{\text{Expected value if } V \text{ is below } Y} + \underbrace{\text{Y}\left[N\big(d_2(\text{Y},t)\big) - N\big(d_2(K,t)\big)\right]}_{\text{Probability of } Y < V < K \text{ at maturity.}} \tag{A1.6}$$

Y in isolation is the expected payoff if $Y < V < K$, and is the expected recovery rate value of the last coupon payment and the principal. The intuition is that promised payments are paid in full when the asset value V is larger than the default barrier K. Furthermore, when $V < K$ but larger than Y, the payoff is restricted to Y. Moreover, if V is less than Y, then the payoff is equal to V.

Appendix A2 – The risk-free rates used in the extended Merton model.

Figure A2.1 is a representation of a forward-looking market sentiment, and clearly exhibits a steep increase in the zero-coupon yield of Norwegian government bonds around the start of 2021.

Figure A2.1: Visual presentation of the NSS yield curve modeled on Norwegian government bonds. Interest rates are measured as an average of all maturities modeled at each date.

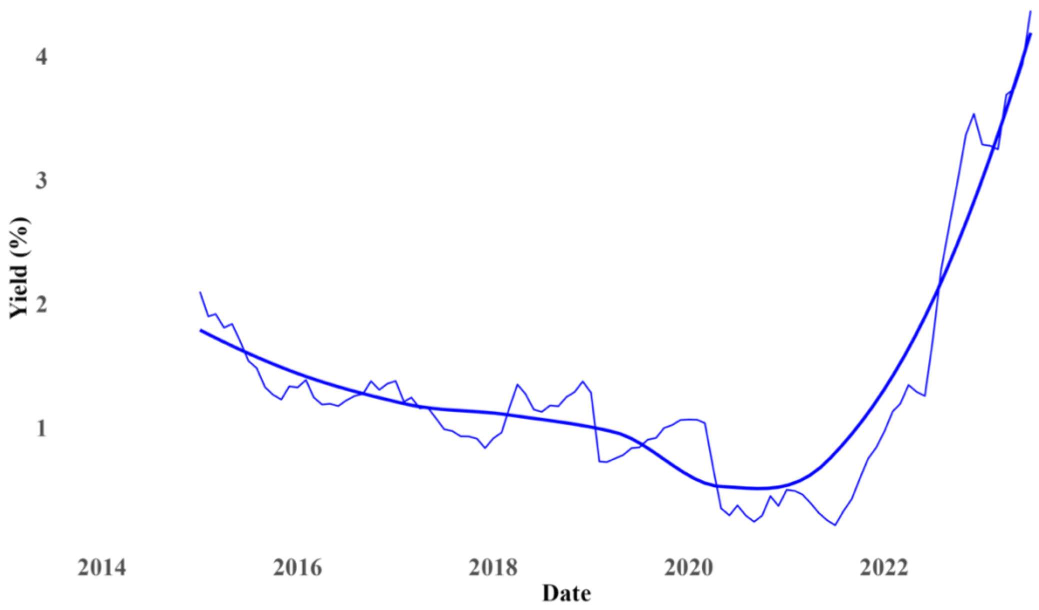


Source: Nullkupongrenter (n.d.)

Appendix A3 – The risk-free rates used in the Longstaff & Schwartz (1995) model.

Figure A3.1 presents the Vasicek yield curve, which we utilize in the LS model to discount high-yield bonds. This curve is modeled using 3-month NIBOR rates from 2005 to 2023. We estimate the rates 10 years back and make a rolling implementation of the model estimates every 6 months. This means that the rates for January of 2023 were estimated on daily 3-month NIBOR rates going back to January 2013. The Vasicek model's sensitivity to fluctuations in short-term rates is evident in Figure A3.1, especially with the sharp increase in rates observed at the beginning of 2021, mirroring the volatility in short-term rate changes.

Figure A3.1: Visual presentation of the Vasicek yield curve. Interest rates are measured as an average of all maturities modeled at each date.



Appendix B1 – Formulas for model comparison

$$Mean\ Pricing\ Percentage\ Error = \frac{\sum_{i=1}^{n}\left(\frac{Modeled\ Pricing_i - Actual\ Pricing_i}{Actual\ Pricing_i}\right)}{n} \quad \text{(B1.1)}$$

$$Mean\ Absolute\ Percentage\ Pricing\ Error = \frac{\sum_{i=1}^{n}\left|\frac{Modeled\ Pricing_i - Actual\ Pricing_i}{Actual\ Pricing_i}\right|}{n} \quad \text{(B1.2)}$$

$$Mean\ Spread\ Percentage\ Error = \frac{\sum_{i=1}^{n}\left(\frac{Modeled\ Spread_i - Actual\ Spread_i}{Actual\ Spread_i}\right)}{n} \quad \text{(B1.3)}$$

$$Mean\ Absolute\ Spread\ Percentage\ Error = \frac{\sum_{i=1}^{n}\left|\frac{Modeled\ Spread_i - Actual\ Spread_i}{Actual\ Spread_i}\right|}{n} \quad \text{(B1.4)}$$

$$Root\ Mean\ Squared\ Spread\ Percentage\ Error = \sqrt{\frac{\sum_{i=1}^{n}\left(\frac{Modeled\ Spread_i - Actual\ Spread_i}{Actual\ Spread_i}\right)^2}{n}} \quad \text{(B1.5)}$$

Appendix C1 – Issuers and their respective number of high-yield bonds.

Ticker	Number of Bonds	Ticker	Number of Bonds
AFK	2	KCC	3
AKER	8	KMCP	3
AKH	1	MGN	3
AKRBP	2	MOWI	3
AKSO	4	NAS	6
AMSC	2	NRC	1
AQUIL	2	NSKOG	3
AUSS	3	OCY	8
BFISH	1	ODF	8
BNOR	5	OKEA	6
BONHR	9	PEN	3
BOR	2	PEXIP	1
CODE	1	PGS	1
CRAYN	5	PRS	5
CSAM	1	RECSI	3
DDRIL	2	SBO	1
DNO	6	SCATC	4
EIOF	1	SOFF	1
ENDUR	1	WAWI	12
GSF	2	WEST	1
HEX	1	ZAL	2
IOX	2	**TOTAL**	**141**

Appendix C2 – Characteristics for the high-yield bond sample

Risk Type	Unique Bonds
Senior Secured	38
Senior Unsecured	102
Subordinated	1
Total	**141**

Rating	Unique Bonds
BBB	12
B	1
Not Rated	128
Total	**141**

Green Label	Unique Bonds
Green	10
Not Green	131
Total	**141**

APPENDIX D1 – Correlation matrix

	Mispricing	Maturity	Asset Volatility	Leverage Ratio	Payout Ratio	Bid-Ask Spread	Illiquidity	Price/Book	Size	Momentum
Mispricing	1	-0.003	-0.009	0.449	-0.323	0.628	0.019	-0.299	-0.188	-0.09
Maturity	-0.003	1	-0.052	-0.023	0.092	0.046	-0.04	0.083	-0.009	0.024
Asset Volatility	-0.009	-0.052	1	-0.008	0.004	-0.009	0.001	0.006	0.012	0.015
Leverage Ratio	0.449	-0.023	-0.008	1	-0.014	0.672	-0.104	-0.039	-0.168	-0.101
Payout Ratio	-0.323	0.092	0.004	-0.014	1	0.009	-0.004	0.935	-0.097	-0.058
Bid-Ask Spread	0.628	0.046	-0.009	0.672	0.009	1	-0.009	-0.0005	-0.149	-0.081
Illiquidity	0.019	-0.04	0.001	-0.104	-0.004	-0.009	1	-0.004	-0.024	-0.008
Price/Book	-0.299	0.083	0.006	-0.039	0.935	-0.0005	-0.004	1	-0.076	-0.052
Size	-0.188	-0.009	0.012	-0.168	-0.097	-0.149	-0.024	-0.076	1	0.536
Momentum	-0.09	0.024	0.015	-0.101	-0.058	-0.081	-0.008	-0.052	0.536	1

APPENDIX D2 – VIF-table

	Maturity	Asset Volatility	Leverage Ratio	Payout Ratio	Bid-Ask Spread	Illiquidity	Price/Book	Size	Momentum
VIF	1.02	1.00	1.90	8.02	1.86	1.02	8.00	1.45	1.41

APPENDIX D3 – Generalized VIF-table

	Maturity	Asset Volatility	Leverage Ratio	Payout Ratio	Green	Risk	Return Type	Industry Group
Generalized VIF	1.33	1.00	1.95	1.18	1.52	2.68	52.66	116.92
Degrees of Freedom	1	1	1	1	1	1	3	9
GVIF^(1/(2*Df))	1.15	1.00	1.40	1.09	1.23	1.64	1.94	1.30
GVIF^(1/(2*Df))^2	1.32	1.00	1.96	1.19	1.51	2.69	3.76	1.69

10. REFERENCES

Aarvik, G. A., & Nordli, S. M. (2016). The determinants of recovery rates in the Nordic high yield bond market: An empirical study of the price development and market-based recovery of defaulted high yield bonds [Master thesis]. https://openaccess.nhh.no/nhh-xmlui/handle/11250/2432162

Altman, E. I., & Kishore, V. M. (1996). Almost Everything You Wanted to Know about Recoveries on Defaulted Bonds. Financial Analysts Journal, 52(6), 57–64.

Amihud, Y., & Mendelson, H. (1986). Liquidity and Stock Returns. Financial Analysts Journal, 42(3), 43–48.

Andreassen, H. M., & Semmen, K. (2023). Samling 1—Grunnleggende Renteanalyse [Lecture Notes]. Norwegian School of Economics.

Bai, H. (2021). Unemployment and credit risk. Journal of Financial Economics, 142(1), 127–145. https://doi.org/10.1016/j.jfineco.2021.05.046

Bao, J., Pan, J., & Wang, J. (2011). The Illiquidity of Corporate Bonds. The Journal of Finance, 66(3), 911–946.

Berk, J., & DeMarzo, P. (2019). Corporate Finance, Global Edition (5th edition). Pearson.

Bhamra, H. S., Kuehn, L.-A., & Strebulaev, I. A. (2009). The Levered Equity Risk Premium and Credit Spreads: A Unified Framework (SSRN Scholarly Paper 1016891). https://doi.org/10.2139/ssrn.1016891

Black, F., & Cox, J. C. (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. The Journal of Finance, 31(2), 351–367. https://doi.org/10.2307/2326607

Black, F., Derman, E., & Toy, W. (1990). A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options. Financial Analysts Journal, 46(1), 33–39.

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81(3), 637–654.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. The Journal of Finance, 52(1), 57–82. https://doi.org/10.2307/2329556

Chen, H., Cui, R., He, Z., & Milbradt, K. (2018). Quantifying Liquidity and Default Risks of Corporate Bonds over the Business Cycle. The Review of Financial Studies, 31(3), 852–897. https://doi.org/10.1093/rfs/hhx107

Chen, L., Collin-Dufresne, P., & Goldstein, R. S. (2009). On the Relation between the Credit Spread Puzzle and the Equity Premium Puzzle (SSRN Scholarly Paper 1468193). https://papers.ssrn.com/abstract=1468193

Collin-Dufresne, P., & Goldstein, R. S. (2001). Do Credit Spreads Reflect Stationary Leverage Ratios? The Journal of Finance, 56(5), 1929–1957.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. Econometrica, 53(2), 385–407. https://doi.org/10.2307/1911242

Cremers, K. J. M., Driessen, J., & Maenhout, P. (2008). Explaining the Level of Credit Spreads: Option-Implied Jump Risk Premia in a Firm Value Model. The Review of Financial Studies, 21(5), 2209–2242. https://doi.org/10.1093/rfs/hhn071

Elton, E. J., Gruber, M. J., Agrawal, D., & Mann, C. (2001). Explaining the Rate Spread on Corporate Bonds. The Journal of Finance, 56(1), 247–277.

Eom, Y. H., Helwege, J., & Huang, J.-Z. (2004). Structural Models of Corporate Bond Pricing: An Empirical Analysis. The Review of Financial Studies, 17(2), 499–544.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Feldhütter, P., & Schaefer, S. M. (2018). The Myth of the Credit Spread Puzzle. The Review of Financial Studies, 31(8), 2897–2942.

Geske, R. (1977). The Valuation of Corporate Liabilities as Compound Options. The Journal of Financial and Quantitative Analysis, 12(4), 541–552. https://doi.org/10.2307/2330330

Huang, J.-Z., & Huang, M. (2012). How Much of Corporate-Treasury Yield Spread is Due to Credit Risk? (SSRN Scholarly Paper 307360). https://doi.org/10.2139/ssrn.307360

Jankowitsch, R., Nagler, F., & Subrahmanyam, M. G. (2014). The determinants of recovery rates in the US corporate bond market. Journal of Financial Economics, 114(1), 155–177. https://doi.org/10.1016/j.jfineco.2014.06.001

Kim, I. J., Ramaswamy, K., & Sundaresan, S. (1993). Does Default Risk in Coupons Affect the Valuation of Corporate Bonds?: A Contingent Claims Model. Financial Management, 22(3), 117–131. https://doi.org/10.2307/3665932

Langdalen, H., & Johansen, V. H. (2016). On the extent and origins of the Merton model`s credit spread puzzle: A study of the credit risk pricing of Norwegian corporate bonds 2003-2014 [Master thesis]. https://openaccess.nhh.no/nhh-xmlui/handle/11250/2403486

Leland, H. E., & Toft, K. B. (1996). Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads. The Journal of Finance, 51(3), 987–1019. https://doi.org/10.2307/2329229

Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. The Journal of Finance, 60(5), 2213–2253.

Longstaff, F. A., & Schwartz, E. S. (1995). A Simple Approach to Valuing Risky Fixed and Floating Rate Debt. The Journal of Finance, 50(3), 789–819. https://doi.org/10.2307/2329288

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. The Journal of Finance, 29(2), 449–470. https://doi.org/10.2307/2978814

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. The Journal of Business, 60(4), 473–489.

Nullkupongrenter. (n.d.). Retrieved December 12, 2023, from https://www.norges-bank.no/tema/Statistikk/statsrenter/statsrenter/

Sæbø, J. K. (2015). The Credit Spread Puzzle Does Exist—But Is It Really a Puzzle? The Journal of Fixed Income, 25(1), 75–83. https://doi.org/10.3905/jfi.2015.25.1.075

Schoenmaker, D., & Schramade, W. (2019). Principles of Sustainable Finance (Illustrated edition). Oxford University Press.

Sundaresan, S. (2009). Fixed Income Markets and Their Derivatives (3rd edition). Academic Press.

Svensson, L. E. O. (1994). Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994 (Working Paper 4871). National Bureau of Economic Research. https://doi.org/10.3386/w4871

Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5(2), 177–188. https://doi.org/10.1016/0304-405X(77)90016-2

Wensaas, E., & Wist, J. B. (2019). Greenium or myth: Do green labels affect bond yields in the Nordic markets? [Master thesis]. https://openaccess.nhh.no/nhh-xmlui/handle/11250/2644900

Ytterdal, A. G., & Knappskog, B. H. (2015). Predicting spreads in the Nordic high yield bond market: A study of credit pricing in the years 2000-2012 [Master thesis]. https://openaccess.nhh.no/nhh-xmlui/handle/11250/302160

Zhang, B. Y., Zhou, H., & Zhu, H. (2009). Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms. The Review of Financial Studies, 22(12), 5099–5131.