CREDIT WHERE CREDIT'S DUE

AN EMPIRICAL STUDY OF DEFAULTS IN THE SWEDISH CORPORATE BOND MARKET BETWEEN 2004 AND 2023

KRISTOFFER ÖSTLIN
LUDVIG SVIBERG

Master Thesis in Finance

Stockholm School of Economics

2023

Corporate Bond Defaults: An Empirical Study of Defaults in the Swedish Corporate Bond Market Between 2004 and 2023⋆

Abstract:

This paper studies the relationship between bond defaults and yield spread, and other bond and company variables observable at issuance using a comprehensive dataset of matured and defaulted bonds from non-financial Swedish firms, covering the period 2004-2023. We find that higher yield spreads are correlated with higher default probabilities, particularly in the high-yield (HY) bond segment. More specifically, we find that a one percentage point increase in yield spread increases the probability of default with 0.8 and 2.2 percent for the full sample and the HY sample, respectively. Moreover, the effect on default probabilities increases non-linearly with higher yield spreads. Other notable variables we find as strong predictors of bond default are fixed rate bonds, time to maturity, if the bond is issued in NOK or USD, dividend payout ratio, and company age. Surprisingly, however, we do not find accounting ratios to be strong predictors of bond default.

Keywords:

Corporate Bonds, Yield Spread, Default, Swedish Bond Market, Logit Regression Model

Authors:

Kristoffer Östlin (24040)
Ludvig Sviberg (24099)

Tutor:

Bo Becker, Professor, Cevian Capital Professor of Finance, Department of Finance, Stockholm School of Economics

Master Thesis
Master Program in Finance
Stockholm School of Economics

⋆ We want to thank our tutor, Bo Becker, for his guidance throughout the thesis process. We would also want to thank Jakob Eliasson at Nordic Credit Partners for his engagement and thoughtful insights regarding this topic and the Swedish corporate bond market. Finally, we want to thank Stamdata and Nordic Trustee for generously providing us with access to their database, which was essential to accomplish the study.

Contents

1. Introduction

This paper studies the predictive ability of yield spread, and other observable bond and company variables at the time of issuance and default events in the Swedish corporate bond market. This is done using an extensive data set of matured and defaulted bonds issued by Swedish non-financial firms extracted from the Stamdata database and hand-collected data extracted from bond term sheets covering the period 2004-2023[1]. The bond data is tabulated with financial statement data from the Serrano database provided by the Swedish House of Finance (SHoF).

Capital markets are an essential part of firm's strategy for optimal financial management through a combination of bank lending, private placements, and public issues of equity and debt. Historically on the Swedish market, debt funding by publicly issued corporate bonds has been subordinate to traditional bank lending (Gunnarsdottir and Lindh, 2011). However, in the period since the global financial crisis, stricter regulatory reforms in the financing landscape have increased the cost of debt through traditional bank loans. This has led Swedish corporations to search for alternative debt financing options, such as corporate bonds (Bonthron, 2014). An attractive characteristic of corporate bonds is that bonds are generally issued more frequently than equity, making corporate bonds essential for firms' continuous funding and a vital element in firms' capital structure. Hence, the Swedish corporate bond market has gained a substantial increase in interest from both issuers and investors in search for capital and returns.

A fundamental truth expected to hold true in financial markets is that returns should follow risk. The expected loss on corporate debt is an important input in several corporate financing decisions, and not least for the issuance of and investment in corporate bonds. Thus, correct pricing of corporate bonds is vital for both issuers and investors. The risk adjusted price paid by the issuer and the return gained by the investors can be said to be the yield spread on the bond, defined as the difference between the yield on the bond and a treasury note or government bond[2] of the same maturity (Elton et al., 2001).

In theory, all things equal, the yield spread should reflect the additional risk investors are exposed to from investing in corporate bonds rather than risk-free

[1] Year-to-date until October 2023.
[2] Hereafter, treasury notes and government bonds will be jointly referred to as "government bonds" only.

government bonds, implying that the yield spread should compensate for the credit risk of firms defaulting on their debt. However, researchers have found that yield spreads on corporate bonds are many times higher than what would be expected from credit risk alone, only representing a fraction of the total yield spread, with other risk factors such as liquidity risk and interest risk also explaining the spread. This phenomenon is called the "credit spread puzzle" and has been widely studied in the U.S. corporate bond market and primarily from a secondary market perspective (e.g., Amato and Remolona, 2003, Huang and Huang, 2012). However, the research on the Nordic corporate bond market in general and the Swedish corporate bond market in particular is limited, especially from a primary market perspective, i.e., at the time of issuance of the bonds.

There are several peculiarities in the Swedish corporate bond market that make it especially relevant to study the determinants of default, such as the low frequency of credit ratings, shorter durations, and the large degree of floating rate bonds (Bonthron, 2014). With low documentation requirements, the absence of an official credit rating requirement, and the domination of over-the-counter (OTC) trades, both issuers and investors have suboptimal information of what the cost of debt is compensating for or what idiosyncratic risks a bond contains.

With this background in mind, we investigate the Swedish corporate bond market's efficiency in the primary market. More specifically, we investigate the primary market's ability to predict default by applying a binary logit regression model on matured and defaulted Swedish corporate bonds between the years 2004-2023 to answer the research question:

Is the yield spread, and other observable bond and company variables at the time of issuance related to the probability of default of Swedish corporate bonds?

We break down the analysis of this question to study if the effects vary between the full sample of bonds and a High-Yield[3] (HY) subsample given the bonds varying risk characteristics in each segment. To the best of our knowledge, this paper represents the

[3] HY consists of credit ratings corresponding to BB+ (S&P and Fitch) and Ba1 (Moody's) or lower.

first of its kind to apply a binary logit model on a dataset of defaults on Swedish corporate bonds in particular with the purpose to identify default predictors at the time of issuance.

Our analysis starts with descriptive statistics including mean and standard deviation of our full sample of variables, with the data split into a default and non-default segment. Welch's t-statistics are calculated for each variable to characterise statistical differences between the two groups. As expected, the differences in mean for most variables are statistically significant between the two groups.

The main analysis is then conducted using four different models of a multiple logistic regression. The dependent variable investigated in all models is default in the form of a credit event, as defined by Stamdata, by applying one model of regressors including only bond characteristics and one model using both bond variables and company variables for a sample of all bonds and a subset of only HY bonds, respectively. Our main variable of interest, the yield spread at issuance, is found to be statistically significant in a positive direction in all four models thus answering our primary research question. Several other bond variables and company variables are also found to have statistically significant relationships with default probability, but in various directions and of different degrees between the models. For bond variables, we find especially fixed rate bonds and longer maturities to have a positive relationship with the probability of default. Also, bonds issued in other currencies than SEK, specifically NOK and USD, are associated with a higher probability of default. Compared to previous literature using a similar method of predicting firm failure and bond default using accounting ratios, we find a weaker predictive power in our sample for these variables. However, we do find that company age and dividend payout ratio to have statistically significant negative relationship with default probability.

To further investigate the effect of yield spread at issuance on probability of default, we calculate the average marginal effect (AME). For the full bond sample, on average, a one percentage point increase in yield spread increases the probability of default by 0.8 percentage points and the equivalent effect for the HY sample is 2.2 percentage points, ceteris paribus. However, due to the non-linearity of the logit regressions, we find that the marginal effect changes non-linearly for different levels of yield spread. Additionally, we estimate the optimal yield spread to default probability ratio in our sample. For this ratio, we find that for the full sample the optimal yield spread is 3.2 percent and for the

HY sample the optimal yield spread is 4.2 percent. For yield spreads above these thresholds, the increase in risk compensation is diminishing. In addition, we conduct three tests to evaluate or models, namely Pseudo-R^2, the Hosmer-Lemeshow goodness of fit test, and the Area under the Receiver Operating Characteristic (ROC) curve. All three test results indicate a god model fit and predictive power for all models.

In terms of research contribution, this paper adds to the limited existing literature of the Swedish corporate bond market by examining primary market spreads rather than secondary market spreads. This has two main advantages.

First, yields on new issues reflect actual transaction prices rather than brokers' "indicative prices" in the form of estimates derived from pricing matrices or dealers' quotes. Secondary yields could be especially inaccurate in the Swedish corporate bond market, which is characterised by being a relatively small market with many small participants, limited liquidity, and unreliable pricing because of a lack of transparency in pricing and trading (Wollert, 2020). Many investors also hold their bonds either to maturity or until default occurs (Grøstad, 2013), meaning liquidity risk is of less importance (Fabozzi, 2012) which, on the contrary, empirically has been found to make up a large share of the secondary yield spread. As such, yield spreads at issuance provide a more accurate measure of the actual risk premium for default risk demanded by investors.

Second, primary market spreads represent a better measure of the actual cost of debt faced by bond issuers (Gabbi and Sironi, 2002). Credit ratings of bonds have been found to be the main source of determining the yield spread at issuance in Europe as well as in Sweden when applicable (Gabbi and Sironi, 2002; Duan and Rimsaite, 2016). However, given the absence of credit ratings for a large share of bonds issued in Sweden, it is plausible that the cost of debt could be influenced by other factors than credit risk.

Furthermore, our paper also highlights potential variables to be used by industry professionals to incorporate in their investment decisions. This topic is of great interest to industry professionals in Sweden as it could be an additional information input in their investment analysis with the potential to improve their risk adjusted returns given the level of default risk. Such input is especially valuable for investors in the Swedish corporate bond market due to the limited use of credit ratings and limited availability of information about the bond and the issuer provided to the investors during the issuance and placement process (OECD, 2022). In addition, this would be of interest for companies issuing bonds as they could be more accurate when managing the cost of debt given their own risk profile and be able to evaluate their offering more critically from the view of an investor knowing their own performance on variables found to predict default.

The rapid development of the bond market in Sweden is also beneficial for our study as the growth supports increasing data availability and the ability to draw statistical conclusions compared to previous studies.

1.1. Development of the Swedish Corporate Bond Market

For a long time, bank lending dominated corporate credit markets in Sweden (Barr, 2011), with bond financing only being available for a few large companies. This statement still held true during the early 2000s up until the global financial crisis of 2008 (OECD, 2022). With banks retracting from the financing markets with increasingly stringent regulations, such as Basel III implementing higher capital requirements, non-bank lending and corporate financing via the bond market has been steadily on the rise both internationally and in Sweden for several years (Bonthron, 2014). This is further seen in the balance sheets of Swedish firms as the share of debt securities to total financial debt among non-financial firms has increased, although from a low initial share, with 45 percent between 2005 and 2020, which is a steeper increase than in both the Euro Area (42 percent) and the U.S. (15 percent) (OECD, 2022).

During the same time, the bond market has also benefitted from a low interest rate environment, which has made it easier for companies to obtain debt financing and has created a high demand for these securities among investors (Wollert, 2020). In 2021, the Swedish corporate bond market reached record both in terms of volume and new issues (Nordic Trustee, 2023). The proliferation of corporate bonds and the creation of new trading markets, such as First North Bond Market and RänteTorget, have also made the corporate bond market more accessible for smaller companies. According to OECD data (2022), the average number of unique issuers between 2016 and 2021 almost doubled compared to the period between 2010 and 2015, while the median amount issued declined by more than 60 percent, cementing the increased popularity of the Swedish corporate bond market. Consequently, the value of outstanding Swedish corporate bonds increased threefold between 2010 and 2019, making up 30 percent of the external debt funding mix of Swedish non-financial firms up from 17 percent in 2010 (Statistics Sweden, 2020).

The Swedish corporate bond market's sector composition has become increasingly diversified with time as well, although real estate makes up a significant share of market (Nordic Trustee, 2023). Despite that the issued amounts in the Swedish corporate bond market have increased significantly over the years, the traded volume of bonds has not expanded to the same extent resulting in poorer liquidity (Bonthron et al., 2016).

Furthermore, the quality of the issued bonds in Sweden has deteriorated during the 21st century. The low interest environment in combination with the Euro- and Covid-19 crises have driven investors towards higher yields in search for higher returns. Consequently, the value-weighted credit rating for bond issued in 2021 was marginally above BBB-/Baa3 which is the lowest IG rating (OECD, 2022). This deterioration of quality in issued bonds is further indicated by the prevalence of BBB/Baa2 bond issues. In 2021, almost three out of four issued bonds in Sweden were BBB/Baa2-rated compared to an average of less than 11 percent during the time before the global financial crisis (OECD, 2022).

Another fundamental change in the market characteristics of the Swedish corporate bond market during the last two decades is the currency composition. During 2000 to 2010, the Swedish corporate bond market was dominated by issues in EUR with only four percent of the annual issues being in SEK. The comparable share of bonds denominated in SEK between 2010 and 2021 was 42 percent, and in 2021 more than half of the bond issues were made in SEK. However, in terms of amounts issued, EUR is still dominating the market with an average issue amount that is five times larger than the average amount in SEK (OECD, 2022). This development further proves that the Swedish corporate bond market has been more accessible for smaller domestic firms in recent years.

Moreover, the Swedish corporate bond market differentiates itself in many ways from its European and U.S. counterparts. Apart from macro factors such as market size, tax-policies, investor preferences, correlation with global financial markets, and risk premium spreads in interbank offered rate, it also differs in its microstructure characteristics such as trading, regulations, high transaction costs, low liquidity, and few market actors (Wollert, 2020). Additionally, a significant difference is also the large share of issuers without an official credit rating, both among smaller companies and well-established firms (Wollert, 2020). This is partly due to the large costs associated with receiving a rating from the official rating agencies and partly due to the lack of demand from investor for such information. For a long time, Swedish banks provided unofficial ratings, so called “shadow ratings”, for Swedish bond issuers. However, these shadow ratings were discontinued as it breached European law. In the period between 2000 and 2021, 46 percent of issued bonds were unrated (OECD, 2022).

2. Literature Review

2.1. Default Prediction

Closest to our area of study is the topic of default prediction for corporate bonds, as well as firm failures, which has been extensively studied in the literature (Bellovary et al., 2007). However, to the best of our knowledge, there has been no studies applying this type of default prediction methodology on the Swedish corporate bond market specifically. Prominent papers on default prediction includes Beaver (1966), Altman (1968), Blum (1974), Ohlson (1980), Zavgren (1985), Skogsvik (1990), Rosengren (1993), Huffman and Ward (1996), Cotter and Peck (2001), Bernhardsen (2001), Shumway (2001), Marchesini et al. (2004), and many others.

Beaver (1966) pioneered the field of ratio analysis for predicting business failures. The study aimed to empirically verify the usefulness of financial ratios in predicting business failure, defining failure as a firm's inability to meet its financial obligations, with events such as bankruptcy, bond default, an overdrawn bank account, or non-payment of a preferred stock dividend marking a firm as failed. Beaver found a significant relationship between accounting data and firm failure, with the ratios Cash Flow to Total Debt, Net Income to Total Assets, Total Debt to Total Assets, Working Capital to Total Assets, and Current Ratio found to be the most useful for this purpose. Although the study suggested that these ratios are useful predictors of firm failure, Beaver acknowledged certain limitations and biases in the study, especially pertaining to the imperfect pairing of failed and non-failed firms in terms of asset size and the selected population.

Altman (1968) assessed the analytical quality of financial ratio analysis, including leverage ratios such as Working Capital to Total Assets, Retained Earnings to Total Assets, EBIT to Total Assets, Market Value of Equity to Book Value of Total Debt, and Sales to Total Assets. Based on data from publicly held manufacturing companies in the U.S., Altman combined these financial ratios into a discriminant-ratio model, known as the Altman Z-score, which could predict corporate bankruptcy with up to 95 percent accuracy and up to two years prior to the default.

Although many previous studies had analysed models to predict bankruptcy for firms, Huffman and Ward (1996) contributed to the literature on corporate bond defaults by estimating four models which predict bond default based upon public information at the time of issuance of HY bonds. From applying a logistic regression analysis, their results indicated that defaulted HY bond issues were characterised by having higher asset growth rates, lower operating profit margins, larger levels of collateralizable assets, and larger changes in net working capital.

Marchesini et al. (2004) investigated the ability the models developed by Altman (1968), Ohlson (1980), Gentry et al. (1985), and Aziz et al. (1988) to forecast the default of high-yield bonds. Surprisingly, the models performed poorly, with Gentry et al.'s model achieving the best score with a 61.5 percent accuracy rate one year prior to default. The authors' recommendation was to incorporate different variables, such as the logarithm of Total Assets and the Interest Coverage Ratio, to improve the predictive power.

Moreover, Duffie et al. (2007) presented a method to predict conditional probabilities of corporate default based on the Merton (1974) model. According to the authors, future corporate default probabilities depend on the company's distance to default (a volatility-adjusted measure of leverage), the firm's trailing stock return, as well as macroeconomic factors such as trailing S&P500 returns, and interest rates.

First to apply ratio analysis in a Swedish setting was Skogsvik (1990) who investigated the ability of accounting ratios to predict business failure on a sample of Swedish industrial firms using a probit model. With a sample consisting of 51 failure and 328 non-failure Swedish firms, Skogsvik found that current cost accounting ratios perform as well as historic cost accounting ratios in predicting corporate default with the variables Interest Cost to Liabilities, Return on Assets, and Equity to Total Assets being especially significant.

Related to bond default prediction in the Nordics, Grøstad (2013) investigated determinants of defaults experienced in the Norwegian HY bond market from 2006 to 2013. Although the study used a limited data set of 176 observations, some interesting findings were uncovered in his paper. Univariate analysis found that Norwegian firms issuing HY bonds experiencing default in the sample on average had lower levels of profitability already at issuance, were smaller in size, had higher yield spreads on their bonds at the time of issuance, and the issued amount was higher relative to the issuer's total assets. He also conducted a logistic regression analysis that corroborated the findings in the univariate analysis regarding profits relative to debt level, company size, and that younger firms were more prone to default on their bond issues.

More recently, Khan and O'Rawe (2016) built on the work of Grøstad (2013) by studying determining factors of default observable at the time of issuance for Nordic corporate bonds. This was done by analysing a sample of 627 original issue bonds in the period 2006 to 2014 using a binary logit model. Their results suggested that the yield spread and a combination of financial ratios, certain bond characteristics, a stock returns-based industry variable, the size of the issuer, and the firm's distance to default were useful estimates for predicting default. Our study differs in focusing only on the issuance of Swedish corporate bonds, in the choice of explanatory variables and by using a sample covering significantly longer period.

2.2. Determinants of Yield Spread

Theoretically, all things equal, the yield spread is a measure of default risk that applies to a risky bond compared to a risk-free government bond of similar maturity. Historically, researchers have tried to explain this relationship using structural diffusion models (e.g., Merton, 1974; Black and Cox, 1975), but the performance in explaining the yield spread level and variance has been uneven. Many studies, such as Jones et al. (1984), Delianedis and Geske (2001), Driessen (2005), Eom et al. (2004), and Huang and Huang (2012), have found a surprisingly small amount of the yield spread, to a varying degree, can be explained by default probability alone. For example, Delianedis and Geske (2001) found that for AAA/Aaa rated bonds, five percent of the yield spread can be attributed to default risk, and for BBB/Baa1-3 rated bonds the equivalent number was 22 percent. Furthermore, both Collin-Dufresne et al. (2001) and Huang and Huang (2012) indicated that the structural form models only explain 25 percent of the levels and change in yield spreads. In addition, Elton et al. (2001) found support for the existence of risk premiums on corporate bonds and Lin et al. (2011) suggested that as much as 53 percent of the bond risk is explained by non-default risks. Multiple of these studies also suggested other factors explaining the level of yield spread. For instance, Elton et al. (2001) decomposed spot rates on corporate bonds into expected loss from default, taxes, and a residual.

The yield spread has also been found to be related to a firm's ability to manage their capital structure efficiently. Longstaff et al. (2005) found that most of the yield spread consisted of a default component, with over 50 percent of the spread being explained by the default risk even for the highest-rated Investment-Grade[4] (IG) companies. In addition, they claimed that the non-default component of the yield spread was related to bond-specific illiquidity, which could explain why firms tend to use a suboptimal level of debt in their capital structure as the illiquidity effect affects the cost of debt equilibrium. This was further elaborated on by Collin-Dufresne et al. (2001) who concluded that accounting for a company's ability to manage its level of outstanding debt had a significant impact on yield spread predictions, implying a much lower sensitivity of changes in firm values on yield spreads than previously thought. Further, Chen et al. (2007) echoed the previous literature of Collin-Dufresne et al. (2001) and Huang and Huang (2012) by stating that credit spreads cannot be fully explained by default risk. In line with Longstaff et al. (2005), the authors found that liquidity improvements of a bond significantly reduced the yield spread even when controlling for bond-specific, firm-specific, and macroeconomic variables.

Furthermore, the yield spread has been studied from a primary market perspective. Fridson and Garman (1998) explored the determinants of the yield spread for new HY bonds in the U.S. market and found that up to 56 percent of the risk premium at issuance could be explained by factors such as credit rating, time to maturity, seniority in capital structure, call option, zero coupon, first-time issues, and secondary-market spreads. Gabbi and Sironi (2002) extended the literature by studying the determinants of yield spreads in the primary corporate bond market by analysing over 3,000 Eurobond yield spreads at issuance between 1991 and 2001 and found that the bonds' credit ratings were the most important determinant of the yield spread. However, in contrast to other studies on the yield spread, primary market efficiency and secondary market characteristics, such as liquidity, proved to have poor explanatory value at the time of bond issuance.

Some further insights regarding yield spreads in the Swedish market have been uncovered in previous theses. Duan and Rimsaite (2016) compared the pricing dynamics in the primary HY bond market in Sweden and Norway and found that up to 82 percent of the variation in yield spreads among Swedish HY bonds could be explained by various issue-specific, company-specific, and market-specific factors. Specifically, the time to maturity, credit rating, and equity volatility were found to significantly affect the yield

[4] IG consists of credit ratings corresponding to BBB- (S&P and Fitch) and Baa3 (Moody's) or higher.

spread, implying that Swedish investors require higher yields for low-rated bonds and that Swedish corporations face higher costs of debt. Moreover, Frisén and Lövstedt (2015) investigated underpricing of corporate bonds in the Swedish market. They found in line with previous literature, that corporate bond issues are on average underpriced by around 0.21 percent. In addition, they found that the observed underpricing is related to information problems and more specifically related to the issuer status, with underpricing in first-time issues being more than twice as high as in seasoned issues.

Additionally, Ekman and Tibell (2019) found a positive relationship between liquidity risk, interest risk, and credit risk with the secondary yield spread in the Swedish corporate bond market, in line with other international studies (e.g., Elton et al., 2001; Chen et al., 2007). Göransson and Olofsgård (2020) focused on uncovering the underlying determinant variables for yield spreads in the Swedish market. They found that international indices, estimated earnings-per-share (EPS), government borrowing rate, and changes in fund flows explained about 87 percent of the variance in yield spreads, which would facilitate the presence of a credit spread puzzle on the Swedish market as well. Further, Andersson (2021) extended Chen et al.'s (2007) study on liquidity and yield spreads by investigating whether the total amount issued of a company influences corporate bonds' yield spreads in the Swedish market. Consistent with Longstaff et al.'s (2005) result of the importance of the outstanding principal amount in explaining bond liquidity, Andersson (2021) found a convex relationship across the total amount of issued bonds by a company and the yield spread, indicating that the issued amount has an impact in the yield spread a company needs to pay.

3. Data and Methodology

3.1. Bond Selection and Data

There is no recognised consensus around the definition of the Swedish corporate bond market. Hence, we choose to define the market as bonds issued by companies domiciled and incorporated in Sweden. To derive our bond sample, we apply several constraints to our defined bond market. First, due to the development and maturity of the market in recent decades, we constrain our data extraction to bonds issued from January 2000. Second, the bonds must have matured, been redeemed, or defaulted as of October 2023 to be able to distinguish between defaulted and non-defaulted bonds, i.e., the bonds cannot still be outstanding. Third, similar to previous studies, we exclude all bonds with special features such us perpetual bonds, credit linked notes, commercial papers, and warrants, since these would be priced significantly different from the bonds included in the sample. Fourth, we exclude banks, financial firms, government-extensions, and government-owned firms, as well as municipality-owned firms. This also includes the exclusion of bonds that are government- or municipality guaranteed. These bonds are deemed to have a different leverage and risk profile compared to bonds issued by ordinary firms, which would impair the comparability of default probability.

The availability of credible data on the Nordic bond market is scarce. Stamdata, operated by Nordic Trustee, is considered the most comprehensive database on Nordic bonds, and is used as the main source of bond information in this paper. Stamdata provides detailed bond reference data including information from loan agreements, term sheets, and bond notices. Most of the bond information of interest for our study can be automatically extracted from the database; however, some of the information, such as coupon at issue, callability, and first-time issues, must be collected manually for each bond from the database and term sheets. In addition, the Stamdata database includes the most comprehensive information regarding bond defaults in the Nordic market that is available, which is a crucial component for our study. When identifying defaulted bonds, we rely on Stamdata's definition of default, which can be separated into three categories, namely bankruptcy, non-payment, and distressed exchange. These categories are further divided into subcategories, according to Stamdata. Bankruptcy includes both bankruptcy proceedings and voluntary liquidation. Non-payment involves a standstill and deferral of principal, instalments, and coupon payments. The final category, distressed exchanges, includes distressed companies who restructure or renegotiate the debt with its creditors, including changes in bond indentures or updated covenants and securities. All these scenarios are considered as a reason for default and will hence be used as the definition of default in this paper. Furthermore, for data concerning bond variables that are not available in Stamdata, official and public sources of information are used. For exchange rates, the Riksbank[5] provides historical daily exchange rates related to the SEK. Similarly, daily historical data on government bond yields is extracted from the relevant Central Banks' official sources. One exemption from this is for government bond yields related to bonds issued in EUR, for which average yields of all EUR-based countries is provided by Eurostat.

Given the restrictions and bond criteria above, a list of 2,145 matured, regular non-financial corporate bonds issued after January 2000 is extracted from the Stamdata database, including the exclusion of 344 government- and municipality-related bonds. Further, we manually investigate each bond to add complementary data and to validate the extracted data by going through all the available term sheets, loan agreements, and additional information provided on Stamdata. If a bond still has missing or inconclusive values after this process, such as missing coupon at issuance, bond amount or inconclusive guarantor, we eliminate the bond from the sample, which retracts another 75 bonds and results in a bond sample of 2,070 bonds. Finally, we exclude 129 bonds due to unavailable necessary accounting data. This provides us with a final data sample of 1,941 bonds whereof 96 have defaulted.

[5] Sweden's Central Bank.

Table I – Data Filtering Process

	Bonds	Defaults
Regular non-financial corporate bonds issued from January 2000	2,489	134
Government- and municipality-related bonds	- 344	0
Unavailable or inconclusive bond-specific data	- 75	- 3
Unavailable accounting data	- 129	- 35
Final data sample	**1,941**	**96**

Note: The table displays an overview of the filtering process of our bond data sample. Unavailable data includes both missing and inconclusive data from Stamdata, Serrano or other relevant information sources.

The final sample consists of 335 unique issuers, indicating an average of approximately six bonds per issuer. The first bond in the sample was issued in Q3 2004 and the most recent bond was issued in Q4 2022. The median bond was issued in 2016 which is consistent with the recent development and increased popularity of the Swedish corporate bond market. Further, the shortest maturity of a bond at issuance is six months and the longest maturity is 63 years, although most bonds in the sample are skewed towards short- and medium-termed maturities between two and five years. There are 14 zero coupon bonds, 69 convertible bonds, and 540 callable bonds in the sample. The highest yield spread in our sample is 18.0 percent and the lowest yield spread is -1.8 percent, with an average yield spread of 2.4 percent and a median yield spread of 1.1 percent. Additionally, 346 bonds were first-time issues representing approximately 18 percent of the sample. To determine if a bond was a first-time issue, we match the bonds in our sample with all issued bonds in Sweden, in accordance with our definition of the Swedish market, from the Stamdata database and check if it was the first time the relevant issuer had issued a bond. We also validate the information with the loan agreements and term sheets of the bonds. If a first-time issuer issued more than one bond on the same date, each bond on that date is considered a first-time issue. Moreover, the most denominated currency is SEK (84 percent) followed by EUR (11 percent) and NOK (three percent). The most common industry is Real Estate (45 percent) followed by Industry (17 percent) and Auto (14 percent). Overall, our data sample displays representative characteristics in line with the Swedish market in full. Finally, out of the 96 defaults in our sample, 12 is due to bankruptcy, 49 due to non-payment, and 35 due to distressed exchange. More details on our bond sample's characteristics are presented in Appendix AI.

3.2. Accounting Information

To complement the bond-specific data, we collect accounting information for all issuers related to the bonds. When collecting accounting information, it is crucial to determine for which year and for which entity the data should be gathered. First, since we are studying the bonds at the time of issuance, we choose to focus on the information available at the time. Thus, the base year for the accounting information is the year prior to the bond issuance. Second, some bonds are issued by parent companies or subsidiaries and using the accounting information for those entities solely could be misleading. To determine the appropriate issuing entity, we go through all available loan agreements and term sheets in Stamdata to identify potential guarantors of the bond. If a guarantor is explicitly stated that entity is considered as the real borrower. In cases where there are no or several guarantors stated, we access the Retriever Business database to manually assess group structures and annual reports to identify the real borrower. Furthermore, some companies issue bonds through holding companies or subsidiaries with the primary purpose of acting as a financial intermediary. To correctly reflect the real business and operational risk, we choose to collect consolidated statements on group level as representative accounting information for such cases. In cases where the guarantor or appropriate group entity is a foreign company, these bonds are excluded from our data sample since they are not part of our definition of the Swedish corporate bond market.

The accounting information is primarily accessed through the Serrano database provided by the Swedish House of Finance (SHoF). Serrano consists of historical accounting data for both public and private companies on company level from 1997 and is based on financial statement data from the Swedish Companies Registration Office[6]. In addition to the accounting data, general company data is available from Statistics Sweden (SCB) and group data from Bisnode's group register, which allows us to find and match group data using the identified issuer's organisation number. For companies issuing a bond during their first year as active entities, accounting data from the year prior

[6] In Swedish referred to as "Bolagsverket."

to the bond issue is unavailable and these bonds are removed from the sample due to this data limitation. Further, for some bonds we identify that there are missing values from the Serrano database. In these cases, we manually access the annual reports through Retriever Business to append the values. Similarly, we cross-validate the Serrano data with the annual reports for a several bonds to ensure the accuracy of the extracted data.

3.3. Explanatory Variables

To study predictions of default among corporate bonds at issuance, we identify several potential explanatory variables based on previous literature (e.g., Altman, 1968; Ohlson, 1980; Cheung et al., 1992; Marchesini et al., 2004; Grøstad, 2013; Khan and O'Rawe, 2016). The variables can be categorised into bond variables, company variables, and control variables. Bond variables consists of characteristics unique for each bond, while company variables are related to financial ratios for the company applicable for the year prior to issuance. For control variables, we aim to control for other unobservable characteristics that can influence the probability of default. An overview of all variables is presented in Table II.

Table II – Overview of Explanatory Variables

Variable Name	Exp. Sign	Definition	Var. Type
Bond Variables			
YIELD	+	Yield Spread at Issuance	Integer
FIXED	+	Fixed Coupon Bond	Indicator
CALL	+/-	Callable Bond	Indicator
CONV	+/-	Convertible Bond	Indicator
ZEROC	+/-	Zero Coupon Bond	Indicator
FIRST	+	First-time Bond Issue	Indicator
RISK	+/-	Risk Classification of Bond	Categorical
MATURITY	+/-	Time to Maturity in Years	Integer
AMOUNT	-	Natural Log of Raised Amount	Integer
CURRENCY	+/-	Currency of the Bond Issue	Categorical
Company Variables			
EBITMARG	-	EBIT / Revenue	Integer
ROE	-	Net Income / Total (Book) Equity	Integer
ROA	-	EBIT / Total Assets	Integer
RETAIN	-	Retained Earnings / Total Assets	Integer
ICR	-	EBIT / Interest Expenses	Integer
DPR	-	Dividend / Net Income	Integer
ISSIZE	+	Amount Raised / Total Assets	Integer
WCTA	-	Working Capital / Total Assets	Integer
CACL	-	Current Assets / Current Liabilities	Integer
CASH	-	Cash / Total Assets	Integer
REVTA	-	Revenue / Total Assets	Integer
SOLVENCY	-	Total Equity / Total Assets	Integer
COMPSIZE	-	Natural Log of Total Assets	Integer
COMPAGE	-	Age of Company	Integer
Control Variables			
INDUSTRY	+/-	Industry Classification	Categorical
QUART	+/-	Year and Quarter of Issue	Categorical

Note: The table displays an overview of all variables used in the statistical analysis of bond default predictions. The variable names are used in the presentation of our results except for categorical variables that will take the name of the respective categories.

3.3.1. Bond Variables

We have identified 10 different characteristics which we believe could be indicative and decisive for a bond to default.

Yield spread (YIELD). We assume that all bonds are issued at par which allows us to estimate the yield-to-maturity as the coupon at issuance for each bond[7]. To fairly compare different yields issued at different dates, we subtract the risk-free rate which isolates the risk associated with each bond and neutralises other market risks prevailing at the time of issuance. The risk-free rate is estimated as the yield of a government bond with the closest duration matching the bond in the same currency as the bond issue. If the yield of the relevant government bond was negative, we assumed the risk-free rate to be floored at zero percent. By studying the yield spread, we also capture the required return by investors and the real cost of debt for the issuer. Naturally, our hypothesis is that a higher yield spread is positively correlated with the probability of default as riskier companies would be required to compensate investors more than less risky companies.

Fixed coupon bond (FIXED). Coupon payments can either be determined as a fixed rate or at a floating rate. Floating rate bonds are usually determined by a benchmark rate, e.g., the three-month STIBOR, and an additional margin. This implies that bonds with a floating coupon rate do not carry interest risk to the same extent as fixed rate bonds. Hence, for floating rate bonds, the risk is more concentrated to credit risk while fixed rate bonds carry both interest risk and credit risk. As an indicator variable that takes the value of one if it is a fixed rate bond and the value of zero otherwise, we believe that fixed coupon bonds are positively correlated with the probability of default.

Callable bond (CALL). A call option for the issuer enables the company to redeem the outstanding bond earlier than the stated maturity date. However, the implication of the call option on probability of default is indecisive. On the one hand, riskier companies could embed this option at a relatively high redemption cost to attract investors and secure funding that otherwise would not have been possible. On the contrary, less risky firms could be in a better position to negotiate the loan agreements and include the call option. Yet, less risky firms could also negotiate better coupon rates and would be less in need of a call option than more risky firms. Hence, we cannot determine the expected relationship between a call option and the probability of default. The variable takes the value of one if it there is a call option and the value of zero otherwise.

[7] For zero coupon bonds, the stated or implied yield was used based on the issue price and/or the yield-to-maturity explicitly stated in the term sheet and loan agreement of the bond.

Convertible bond (CONV). A convertible bond option implies that bondholders could choose to convert the debt into equity at a given conversion rate. Companies in a good financial state would refrain from including a convertible option as it would be relatively costly to give up equity instead of paying interest. Similarly, a convertible option could be used by riskier companies as potential compensation to attract investors and secure financing. However, in contrast to this theory, Rosengren (1993) found that convertible bonds have lower default rates than non-convertible bonds. Thus, we are indecisive in the expected relationship between convertible bonds and the probability of default. The variable takes the value of one if there is a convertible option and the value of zero otherwise.

Zero coupon bond (ZEROC). As indicated, zero coupon bonds do not pay coupons over time. Instead, companies usually issue the bonds either at a discounted price or at a premium redemption price. Like previous variables, it is difficult to determine the expected relationship with probability of default. For instance, strong companies can negotiate the possibility to avoid regular interest payments; however, these companies would also refrain from issuing debt at a discounted price. Correspondingly, risky companies would have the reversed dilemma. The variable takes the value of one if there it is a zero coupon bond and the value of zero otherwise.

First-time bond issue (FIRST). This variable relates to companies issuing a bond for the first time. As first-time issuers do not have a track record, the credibility of these bonds is deemed to be lower than for multiple-time issuers. Hence, these bonds are expected to be riskier and a positive relationship with the probability of default is expected. The variable takes the value of one if the bond is a first-time issue and the value of zero otherwise.

Risk classification (RISK). With risk classification we refer to the risk and security level of the bond, namely senior secured, senior unsecured, and subordinated. Bonds with higher level of seniority contains a higher level of security for creditors and would have higher priority in case of default. Consequently, bonds with lower seniority could be deprioritised in cases where the company enters times of financial distress and would hence have an increased probability of default. On the contrary, it is plausible that lenders require to be further up in seniority to be willing to provide funding for riskier firms, and hence a higher share of potential default firms would have more secured bonds. Thus, the expected correlation is inconclusive. This variable is a categorical variable with senior secured as the base case (0), followed by senior unsecured (1) and subordinated (2) in descending order of seniority.

Time to maturity (MATURITY). At the time of issuance, the time to maturity is the expected lifetime of the bond. Cotter and Peck (2001) stated that debt with longer time to maturity experience lower likelihood to default. In support of this, shorter maturities could be insufficient time for companies to generate operational returns on the investments meaning that not enough funds have been generated to repay the debt. However, longer maturity also implies a greater risk for unforeseen operational events and market shocks, which would result in a greater risk for default. An aging effect has also been recorded in the literature (e.g., Asquith et al., 1989; McDonald and Van De Gucht, 1996; McDonald and Van De Gucht, 1999; Moeller and Molina, 2003), where the probability of default increases with the lifetime of the bond. The expected relationship between time to maturity and probability of default is therefore difficult to determine.

Size of issued amount (AMOUNT). The size of the issued amount is harmonised into SEK by using daily exchange rates for issues in other currencies as well as being logarithmic. Our hypothesis is that firms that can raise higher amounts have attracted investors who believe in a positive outlook for the firm, which indicate that the probability of default would be lower for these bonds. In addition, higher amounts are usually related to the size of the firms, and, in theory, larger firms tend to default less than smaller firms. Hence, the expected relationship is negative between the size of the issued amount and the probability of default.

Currency of issue (CURRENCY). Our final bond variable is related to the currency of the bond issue. Larger companies with multinational operations, which in theory tend to be less likely to default, could issue bonds in different currencies to match regional needs reflecting their international operations. However, smaller and riskier firms could turn to bond markets with higher liquidity and more experience of HY offerings to attract funding from investors with more flexible investment criteria. Thus, the expected relationship between currency and probability of default is indecisive. The variable is a categorial variable with SEK as the base case (0), followed by NOK (1), EUR (2), USD (3), GBP (4), and JPY (5).

3.3.2. Company Variables

For company variables, we have identified several important financial ratios to study related to profitability, leverage, liquidity, and other company-specific variables.

EBIT margin (EBITMARG). EBIT margin is a profitability measure indicating a company's ability to generate operating profit. A company with higher EBIT margin will have more funds available to meet interest payments which should generate a lower probability of default. We expect a negative relationship.

Return on equity (ROE). The return on equity measures a company's ability to generate profit from the shareholders' capital investments. A company with higher return on equity is perceived to be in a better financial state and the expected relationship with probability of default is negative.

Return on assets (ROA). A company's return on assets measures how efficiently the company uses its resources to generate profit. Consistent with other profitability-related variables, higher return on assets should be negatively correlated with probability of default.

Retained earnings to total assets (RETAIN). This financial ratio measures the firms' cumulative profitability over time and is a component in Altman's (1968) Z-score. Younger firms that have not been able to aggregate profits could potentially have a lower ratio than older firms, partially covering the effect of a company's age. Companies with a higher ratio also in general have a buffer of retained earnings that can be used in times of financial distress and would therefore be less likely to default. A negative relationship is expected.

Interest coverage ratio (ICR). The interest coverage ratio is part of Altman et al.'s (1977) ZETA™ model, an extension of the Z-score, for identifying bankruptcy risk of companies. The ratio signals a company's ability to pay interest on its outstanding debt. A low interest coverage ratio indicates that the current outstanding debt level could be concerning for the company and hence we expect a negative relationship with the probability of default.

Dividend payout ratio (DPR). The dividend payout ratio measures the amount of dividend paid out to shareholders relative to the generated profit. In general, a company that pays out dividend is assumed to be in a good financial position with solid liquidity.

A firm with high dividend payout ratio should in theory be less likely to default on their debt obligations and therefore we expect a negative relationship.

Issued amount relative the total assets (ISSIZE). This variable is defined as the amount issued in the bond offering relative to total assets of the company. Grammenos et al. (2008) stated that it should be alarming for investors when a firm wants to raise a substantial amount of funds compared to the value of its assets, and that the ratio is a good predictor of defaults. We hypothesise, in line with Grammenos et al. (2008), that a large value of this ratio is positively related to the probability of default.

Working capital ratio (WCTA). The working capital ratio is a liquidity measure and a component in Altman's (1968) Z-score. Working capital is defined as the difference between current assets and current liabilities. The ratio is used as an indication of a company's financial strength and its net liquid assets compared to its total assets. Higher values of the ratio are associated with better financial outlook for the company, and we expect a negative relationship with the probability of default.

Current ratio (CACL). The current ratio is a liquidity measure that indicates a company's ability to cover its short-term debt obligations. Firms with a lower current ratio are closer to financial distress which subsequently could lead to default. We expect a negative relationship.

Cash ratio (CASH). We define the cash ratio as cash and cash equivalents in relation to total assets. To have sufficient liquid funds available to cover interest and other debt obligations is essential for operational survival, although, too elevated levels of cash could be indicative of inefficient management or few investment opportunities. However, the risks of having too much liquid funds available is assumed to be neglectable and we hypothesise that a higher ratio is related to lower probabilities of default. Hence, a negative relationship is expected.

Asset turnover (REVTA). An asset turnover ratio measures a firm's capital intensity and how efficiently the firm generates sales from its resources. For creditors, this measure is of interest as it indicates how well a company can convert the potential funds into operational returns. We predict that higher asset turnover is related to lower probability of default and a negative relationship is expected.

Equity ratio (SOLVENCY). The equity ratio measures how much of a company's assets that are owned by the shareholders and are not leveraged. A high equity ratio implies that the company has a relatively lower debt burden, which, generally, is related to lower risk of default. We expect a negative relationship.

Company size (COMPSIZE). In general, larger companies are more financially flexible than smaller companies, implying that the risk for default should be lower. Hence, we include a variable for the total company size in terms of balance sheet total, with a predicted negative relationship. In accordance with previous literature on company defaults (e.g., Ohlson, 1980), we harmonise the variable value by taking the natural logarithm of the firm's total assets.

Company age (COMPAGE). A company's age is estimated as the years between the founding year of the company and the year of the bond issue. It is reasonable to expect that a company that has been established longer, is typically larger, has more cumulative liquidity generating more financial flexibility, is more recognised, has more experienced management, has a proven performance history, and well-established investor connections. All these characteristics are exemplary for firms with lower probability of default. Thus, the age of the company is expected to have a negative relationship with default probability.

3.3.3. Control Variables

We include two control variables to control for unobservable characteristics that could influence the probability of default.

Industry classification (INDUSTRY). Industry classification is based on Stamdata's variable "Industry Main Grouping". Industry categories available are Agriculture, Auto, Consumer, Convenience Goods, Finance, Health Care, Industry, Insurance, Media, Oil and Gas, Pharmaceuticals, Pulp, Paper, and Forestry, Real Estate, Shipping, Telecom/IT, Transportation, and Utilities. Industry classification is worthwhile to consider as different operating environments, financial characteristics, and industry idiosyncratic risks could impact the probability of firms to default. Overall, we do not expect either a positive or negative relationship with default probability as there could be large variations among the categorial variables.

Issue quarter (QUART). Issue quarter is defined as the year and quarter when a bond was issued. We elect to include this as a control variable because the timing of a bond issue could have a relationship with the likelihood of default as the credit quality of corporate bond issuers varies over the cycle (Greenwood and Hanson, 2011), as well as different prevailing market conditions at different times. Overall, we do not predict either a positive or negative sign for issue quarter as there could be large variations among the categorial variables.

3.4. Statistical Analysis

3.4.1. Welch's t-test of Sample Means

When analysing two independent samples, a Welch's t-test for sample means compare the means of a given variable between the two groups, in our study defaulted and non-defaulted bonds. The test allows us to investigate if there is a statistically significant difference in the mean for the two groups by calculating a t-statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} - \frac{s_2^2}{n_2}}} \tag{1}$$

where $\bar{x}_i$ is the sample mean, s_i^2 is the sample variance, and n_i is the sample size for the respective group.

The t-test tests if the difference in sample means between defaulted and non-defaulted groups are significantly different from zero. We use this test as part of our descriptive statistics overview to better understand the characteristics of our bond sample given our specified bond and company variables.

3.4.2. The Logit Regression Model

To investigate the probability of default we use a logit regression model, inspired by the method used in previous literature on bond default predictability (e.g., Huffman and Ward, 1996; Grammenos et al., 2008; Khan and O'Rawe, 2016). Default is a limited dependent variable, meaning that a bond could either default or not default. This can be expressed as a binary indicator variable that takes the value of one if the bond defaulted or the value of zero if the bond did not default. To analyse the probability of default, we use a regression model; however, due to the binary limitation of the dependent variable, we need a function that transforms the implied probabilities from the regression model to be bounded within the interval [0,1]. The most suitable way to accomplish this is through the non-linear logit regression model.

The logit regression model utilises the logistic function, *F*, which is a function of any random variable, *z*, expressed as (Brooks, 2019):

$$F(z_i) = \frac{e^{z_i}}{1+e^{z_i}} = \frac{1}{1+e^{-z_i}}\,. \tag{2}$$

The logistic function, *F*, is the cumulative logistic distribution, implying that the estimated model would be defined as:

$$P_i = \frac{1}{1+e^{-(\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{ki}+u_i)}} \tag{3}$$

where P_i is the probability that the dependent variable is equal to one (default).

The nature of the logistic function $F(z_i)$ implies that as z_i tends towards infinity, the function tends to the value of one and as z_i tends towards negative infinity, the function tends to the value of zero (Brooks, 2019). This means that for any value of our regression, the probability of default, P_i, will be bounded to a probability between zero and one.

Due to the non-linearity of the model, the parameters are estimated through maximum likelihood estimation (MLE). Through an iterative process, the parameters are chosen to maximise the log-likelihood function, *LLF*, which maximises the probability to observe a default in our model (Brooks, 2019). The iteration is done by a statistical software program, in our case by R and RStudio.

$$LLF = -\sum_{i=1}^{N} \left[y_i \ln(1+e^{-z_i}) + (1-y_i)\ln(1+e^{z_i})\right] \tag{4}$$

where y_i is our dependent default variable.

The non-linearity of the logit regression model also has implications for how the model should be interpreted. It is not possible to directly observe the partial effects of a one-unit change in the independent variable on the dependent variable through the coefficients; however, it is possible to interpret the direction of the effect from the independent variables through the sign of the coefficient. A positive coefficient means that the independent variable is positively correlated with the probability of default, and vice versa for a negative coefficient. To observe the partial effect of a change in the independent variable on the probability of default, we must use calculus to determine the marginal effects. Since our dependent variables consist of integers, indicators, and categorical variables, we choose to calculate the average marginal effects (AME) to derive representative marginal affects for existing bonds in our sample. When deriving the AME of a variable, it is required to take each bond's values for all variables into consideration as the partial derivative is defined as:

$$AME_v = \beta_v \frac{1}{N} \sum_{i=1}^{N} P_i \, (1 - P_i) \tag{5}$$

where β_v is the coefficient of a specific variable (v), N is the number of bonds, and P_i is the probability of default in accordance with Equation (3).

From the AME, it is possible to directly interpret the average effect of a one-unit change in the independent variable on the probability of default.

Lastly, we test for multicollinearity among our regressors by calculating the variance inflation factors (VIF). Multicollinearity appears when there is a linear relationship between the independent variables in the model, which impairs the validity of the model. As a rule of thumb, a VIF value greater or equal to five could be a cause of concern for the model while VIF values below five is considered negligible (Brooks, 2019).

4. Empirical Results

4.1. Descriptive Statistics

The descriptive statistics presented in Table III compare the characteristics of defaulted and non-defaulted bonds for our explanatory variables. From the Welch's t-test it is evident that the means of the two groups is statistically significant different from each other in several variables.

For bonds that defaulted, the average yield spread *(YIELD*) is more than four times higher than the average yield spread of non-defaulted bonds at 8.9 percent and 2.0 percent, respectively. Additionally, the average amount raised (*AMOUNT)* among non-defaulted bonds is approximately four times larger than for defaulted bonds, with the average issue size among non-defaulted bond issues amounting to approximately 480 million SEK and to approximately 122 million SEK for defaulted bonds. Furthermore, the shares of defaulted bonds that are fixed rated bonds *(FIXED)*, have an embedded call option *(CALL)*, or are a first-time issue *(FIRST)* are significantly higher than for non-defaulted bonds. Moreover, the categorical variable currency of issue *(CURRENCY)* illustrates that there is a higher share of bonds issued in a foreign currency in the defaulted bonds group than in the non-defaulted bonds group. The categorical variable risk classification *(RISK)* shows that defaulted bonds, on average, have a higher seniority and priority of security than non-defaulted bonds. Lastly, the sample consists of 13 bonds distributed between nine issuers with a time to maturity exceeding 10 years. All these bonds are in the non-defaulted group, which could impair the mean of the group. The adjusted mean for time to maturity (*MATURITY)* in the non-defaulted group decreases to 3.77 years if the maturity of these bonds would be excluded, compared to a mean of 4.15 years among defaulted bonds.

In terms of company variables, return on assets *(ROA)*, dividend payout ratio *(DPR)*, company size *(COMPSIZE)*, and company age *(COMPAGE)* are the most significant distinguishers between the groups. Both the return on assets and dividend payout ratio are on average negative for defaulted bonds at -2.7 percent and -8.1 percent, respectively, while for the non-defaulted group the return on assets is on average 5.0 percent and the dividend payout ratio is on average 35.4 percent. This supports the theory that companies in financial distress are less efficient in generating operational profit, which in turn leads to a lower capability to generate return to the shareholders. Furthermore, the average company size in the defaulted group is substantially smaller than for the non-defaulted group. Among the defaulted bonds, the average company size is approximately 920 million SEK measured as total assets while for the non-defaulted group the equivalent number is approximately 71 billion SEK. As expected, larger firms have also been established longer. The average company age among defaulted bonds is 10 years and for non-defaulted bonds it is over 50 years, which supports the theory that younger firms have higher risk of default than more mature firms.

Table III – Descriptive Statistics and Welch's t-test

Variable	Defaulted Bonds (96) Mean	Defaulted Bonds (96) Std. Dev.	Non-defaulted Bonds (1,845) Mean	Non-defaulted Bonds (1,845) Std. Dev.	t-statistic
Bond Variables					
YIELD	8.90	3.59	2.01	2.49	18.57 * * *
FIXED	0.82	0.38	0.39	0.49	10.7 * * *
CALL	0.85	0.35	0.25	0.43	16.12 * * *
CONV	0.07	0.26	0.03	0.18	1.46
ZEROC	0.05	0.22	0.00	0.07	2.07 *
FIRST	0.58	0.50	0.16	0.36	8.31 * * *
RISK	0.40	0.61	0.94	0.31	-8.74 * * *
MATURITY	4.15	2.17	4.14	4.47	0.04
AMOUNT	18.62	1.37	19.99	1.20	-9.62 * * *
CURRENCY	0.72	1.04	0.29	0.75	3.96 * * *
Company Variables					
EBITMARG	-281.25	1,051.93	-22.46	749.53	-2.38 *
ROE	-7.75	223.51	10.29	110.56	-0.79
ROA	-2.74	15.63	5.03	12.98	-4.78 * * *
RETAIN	17.19	33.78	22.34	28.70	-1.47
ICR	-93.61	630.16	-3.58	476.01	-1.38
DPR	-8.06	70.07	35.42	137.26	-5.55 * * *
ISSIZE	140.55	439.86	24.03	228.47	2.58 *
WCTA	1.42	25.37	-1.26	18.71	1.02
CACL	136.15	152.55	123.16	403.78	0.71
CASH	5.36	9.18	4.80	6.26	0.59
REVTA	37.63	52.99	50.46	55.29	-2.31 *
SOLVENCY	36.98	24.59	35.87	17.60	0.44
COMPSIZE	20.64	2.89	24.99	3.24	-14.31 * * *
COMPAGE	10.02	10.93	52.54	53.49	-25.44 * * *

Note: * p<0.1, ** p<0.05, *** p<0.01. The table displays an overview of descriptive statistics between defaulted bonds and non-defaulted bonds, together with the t-statistic from Welch's t-test. The number of observations in each group is presented in parentheses (·). Descriptive statistics for control variables INDUSTRY and QUART are not presented in the table due to the absence of intuitive interpretational value of these categorial variables.

4.2. Probability of Default Logit Regression Models

When performing the logit regressions, we build our models on different datasets, namely the full sample (FS) and a subset including only high-yield bonds (HY). Since the

Swedish corporate bond market has inadequate coverage of official credit ratings, Stamdata provides an indicative rating classification for each bond as either IG or HY. In our sample, 537 bonds are labelled as HY bonds according to Stamdata and these constitute our HY dataset. In addition, we apply two different models on each dataset, where we first study the effects of only bond variables (B) and thereafter add the effects of both bond variables and company variables (B & C). When testing for multicollinearity among the variables in the different models, all bond and company variables have a VIF value below five and thus all these variables are included in the models (Appendix AIV). However, the VIF values for the control variables significantly exceeded the threshold in a preliminary VIF test and are thus excluded from the logit regression models to not impede the validity of the models due to multicollinearity. The statistically significant results from the logit regressions are presented in Table IV below, and a full overview of the results covering all variables is presented in Appendix AII.

Table IV – Statistically Significant Coefficients of Logit Regression of Default

	Model Specification			
Variable	FS – B	FS – B & C	HY – B	HY – B & C
YIELD	0.36***	0.33***	0.26***	0.24***
	(0.06)	(0.06)	(0.06)	(0.06)
FIXED	0.63*	0.78**	0.80**	0.88**
	(0.36)	(0.39)	(0.37)	(0.40)
CALL	1.37***	0.86*	0.29	-0.02
	(0.48)	(0.51)	(0.51)	(0.55)
ZEROC	2.78***	2.97***	2.81***	2.71**
	(0.92)	(1.09)	(1.06)	(1.11)
SUBORDINATED	-1.22*	-0.85	-1.33*	-1.12
	(0.74)	(0.83)	(0.71)	(0.80)
MATURITY	0.05**	0.07***	0.10**	0.12***
	(0.02)	(0.03)	(0.04)	(0.05)
NOK	2.74***	2.85***	2.81***	2.82***
	(0.56)	(0.67)	(0.61)	(0.71)
USD	1.45**	1.55**	1.43**	1.55**
	(0.62)	(0.70)	(0.60)	(0.69)
DPR		-0.01***		-0.005**
		(0.002)		(0.002)
CACL		-0.002*		-0.002*
		(0.001)		(0.001)
CASH		-0.04*		-0.04*
		(0.02)		(0.02)
COMPAGE		-0.04***		-0.03**
		(0.01)		(0.02)
Constant	-3.81*	-3.20	-2.66	-1.97
	(2.12)	(2.43)	(2.04)	(2.33)
Observations	1,941	1,941	537	537
Psuedo-R^2	0.51	0.55	0.31	0.36
H-L statistic	14.92 [0.38]	18.16 [0.92]	17.39 [0.14]	31.99 [0.19]
Mean VIF	1.41	1.62	1.42	1.61

Note: * p<0.1, ** p<0.05, *** p<0.01. The table displays the statistically significant coefficients for the logit regression models. Standard errors are in the parentheses (·). H-L statistic refers to the Hosmer-Lemeshow test statistic, with corresponding p-values in brackets [·].

Several bond variables are statistically significant at varying confidence levels. Notably, the yield spread at issuance *(YIELD)* is positively significant at the one percent significance level across all models, implying that a bond with higher yield spread has a higher probability of default. Also, time to maturity *(MATURITY)* is significantly positively correlated with probability of default across all models. This contrasts the findings of Cotter and Peck (2001) that longer maturities would lead to lower likelihood of default and supports that the risk of financial distress increases with the unpredictability of the future state of the economy but is in line with studies finding an aging effect for corporate bonds. Further, zero coupon bonds *(ZEROC)* tend to increase the probability of default, which would support the hypothesis that more risky firms are willing to issue debt at a discounted price and that the possibility to avoid interest payments does not lead to a lower probability of default. However, the impact of zero coupon bonds could be exaggerated due to few observations in the sample. Moreover, in accordance with the descriptive statistics, defaulted bonds tend to be issued in a foreign currency which is supported by the positive direction of the coefficient for NOK *(NOK)* and USD *(USD)*, indicating that riskier firms potentially turn to foreign markets to secure funding. Additionally, this could be due to companies issuing bonds in a foreign currency being exposed to currency risk, which could have an impact on its financial health and in extension its ability to service its debt. Furthermore, as expected, fixed coupon bonds *(FIXED)* have a higher probability of default, especially in the HY sample. The additional interest risk, that is neutralised for floating rate bonds, is proved to be a burden for companies with fluctuations in interest rates. Interestingly, to have a call option *(CALL)* embedded in the loan agreement has a positive effect on default for the full bond sample; however, the effect is not evident in the HY sample. A potential explanation for this is that call options are more common among HY bonds which dilutes the effect within that sample group.

For the full models including both bond data and financial ratios, a few additional variables show significant effects on default probability. The dividend payout ratio *(DPR)* is expectedly negatively correlated with default, as the dividend is distributed from funds available after interest payments have been settled or from a buffer of accumulated profits. Furthermore, company age *(COMPAGE)* is found to be significantly negatively related to default, which confirms our hypothesis that established firms with proven record of performance and more experience are more prone to avoid financial distress.[8]

4.3. Marginal Effect of Yield Spread on Probability of Default

To address our research question on how the yield spread at issuance can predict probability of default, we need to study the impact of the yield spread on default probability. The positive coefficient for the yield spread in all logit regression models confirm that the yield spread is a good predictor for default probabilities; however, to understand the actual effect on default probability we must compute the marginal effects. For the full bond sample with both bond and company variables, the AME of the yield spread is 0.008 and for the same variable selection on HY bonds the AME is 0.022. This means that, on average, a one percentage point increase in yield spread increases the probability of default among all bonds with 0.8 percentage points and among HY bonds with 2.2 percentage points, ceteris paribus. The difference between the samples is plausibly an effect of the relatively higher share of defaulted bonds in the HY sample than in the full sample. Since the AME is dependent on all bonds in the sample, non-defaulted bonds with lower yield spreads dampens the AME in the full sample.

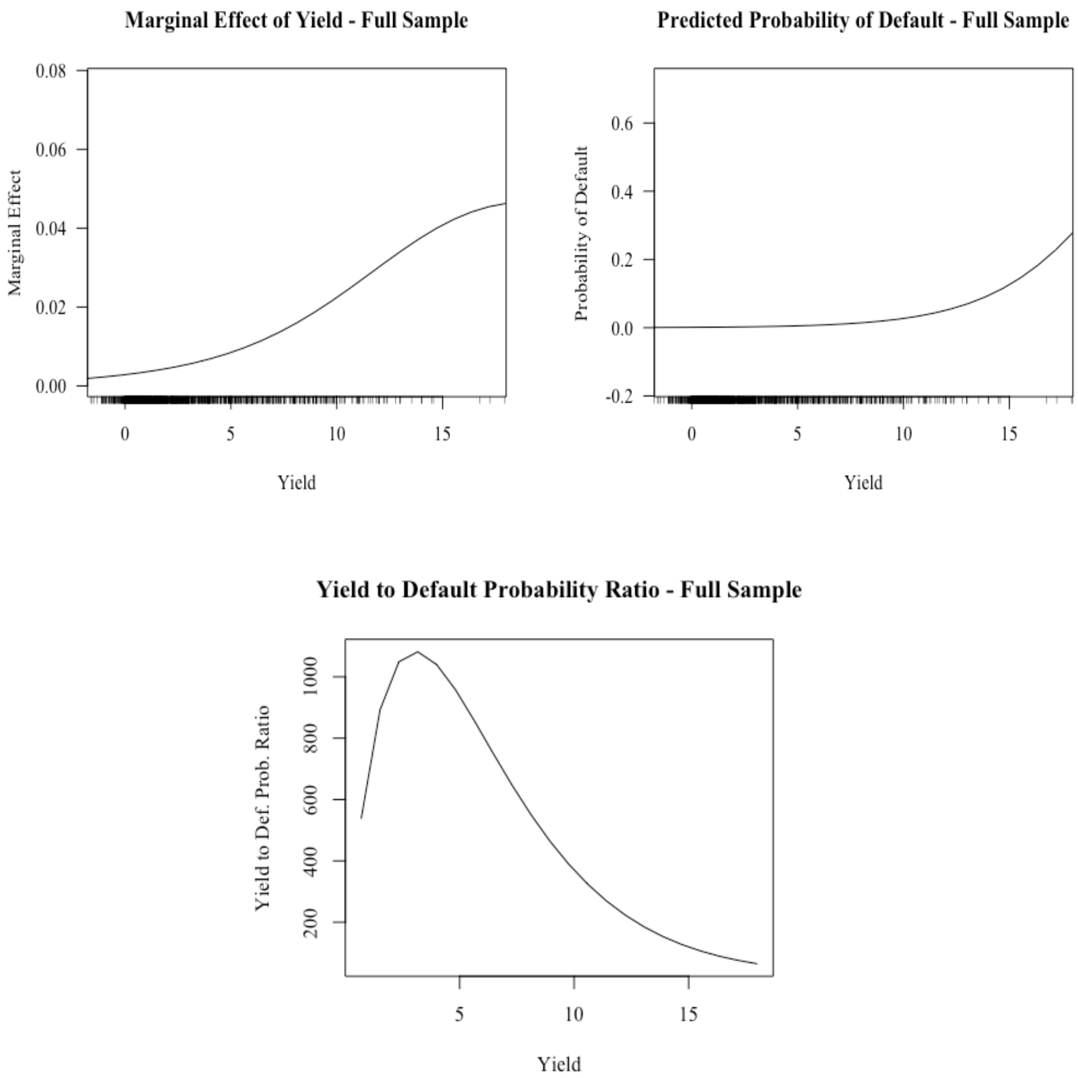
However, due to the non-linearity of the logit regression model, it can be misleading to report one sole marginal effect as the marginal effect changes along the curve. In Figure I, the dynamics of the marginal effect is illustrated for different values of the yield spread for the full bond sample. When studying the marginal effect curve, it becomes evident that the AME of 0.8 percentage points is representative for bonds with low yield spreads. However, for bonds with a yield spread above approximately five percent, the marginal effect on default probability increases significantly. Additionally, for bonds with a yield spread of 10 percent at issuance, the marginal effect is slightly above two percentage points and for yield spreads of 15 percent the marginal effect increases to over four percentage points, on average, ceteris paribus. Consequently, the predicted probability of default illustrates a similar pattern. For yield spreads below 10 percent, the probability of

[8] As a robustness test, we also estimate our models with winsorized versions of our financial ratio variables at the 95 percent level to see if the model yields different results when adjusting for outliers. These results show no significant difference to our original models which supports that our models are not driven by outliers, and to not lose potentially valuable data points we elect to use the original data.

default is almost negligible; however, ceteris paribus, for yield spreads between 10 and 15 percent, the predicted probability of default increases from approximately three percent to 12 percent and for bond issues with a yield spread above 15 percent the predicted probability of default increases significantly to slightly below 28 percent for a yield spread equivalent to 18 percent.

From the marginal effects, we can derive an optimal yield spread where the ratio of yield spread to default probability is highest. For the full bond sample, the optimal yield spread at issuance is estimated to 3.2 percent. Beyond this point, the compensation to bondholders relative the increase in default probability is decreasing.

Figure I – Marginal Effects of Yield Spread (Full Sample)

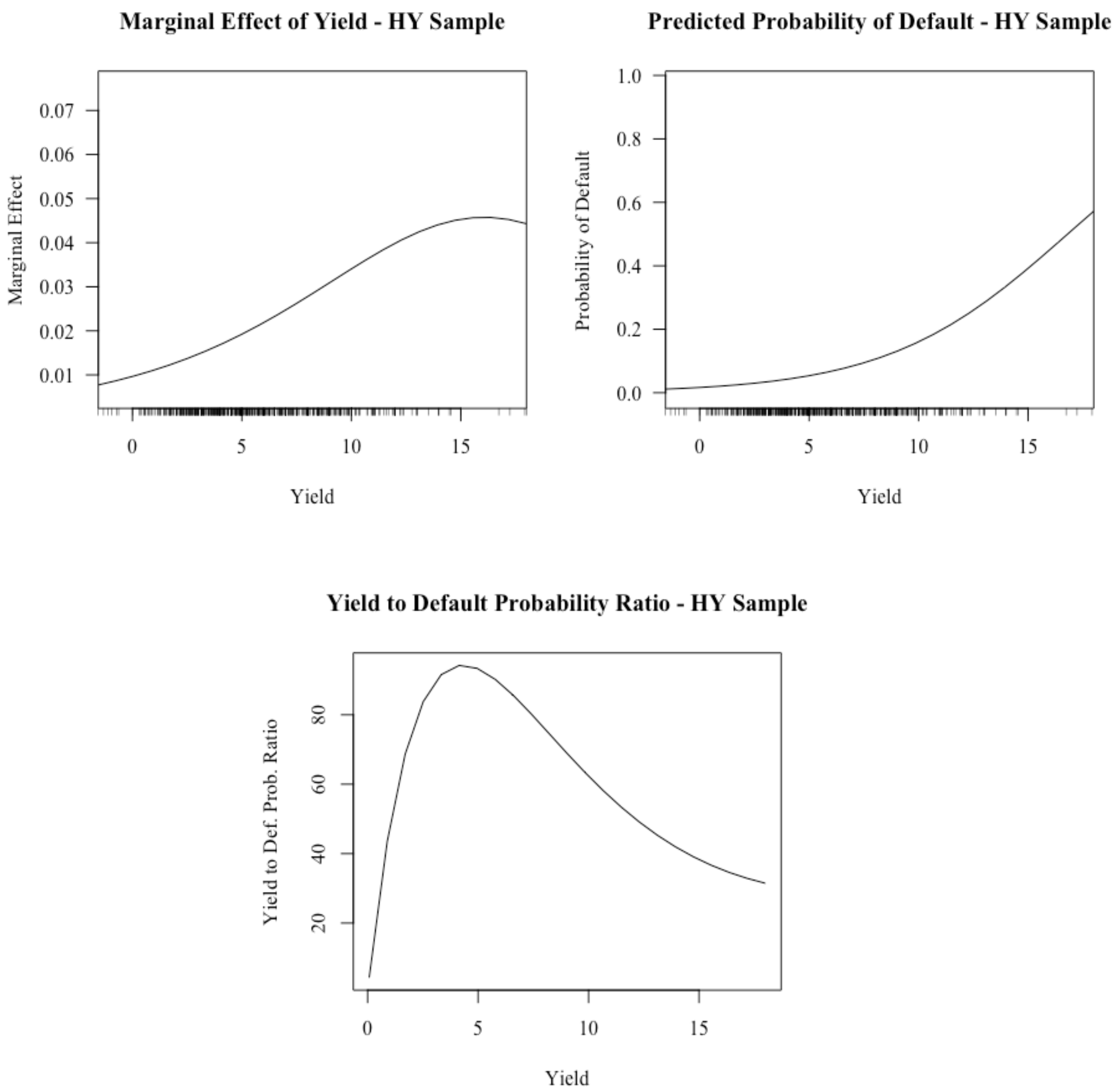


Note: The figures illustrate the marginal effect of the yield spread on probability of default for the full bond sample. The figure to the top left shows how the marginal effect evolves for different yield spreads and the figure to the top right shows the corresponding effect on default probability. The horizontal axes also display the yield spread frequency. The bottom figure illustrates the ratio between yield spread to default probability for various levels of yield spread.

The HY sample illustrates a similar pattern, shown in Figure II. However, for HY bonds the results are amplified compared to the full bond sample. The AME of 2.2 percentage points represents a corresponding yield spread of approximately six percent, which is substantially higher than for the full sample. Thereafter, the increase in marginal effect increases up to a yield of 15 percent where the marginal effect is approximately four percent, in line with the effect of the full bond sample. In addition, the predicted probability of default in the HY sample illustrates a more distinct effect of the yield spread than in the full bond sample. The predicted probability of default is still relatively low for lower yield spreads; however, the default probability accelerates at a higher rate among HY bonds. All else equal, a HY bond with a yield spread of five percent has a probability of default of five percent, a yield spread of 10 percent equals a probability of default of over 16 percent, and yield spread of 15 percent is equivalent a probability of default up to approximately 40 percent, on average, which is substantially higher than for the full bond sample.

The yield spread to default probability ratio for the HY sample also mimic the pattern of the full bond sample. For the HY sample, the optimal yield spread at issuance level is 4.2 percent, which is one percentage point higher than for the full sample. The difference between the two samples indicates that HY bond issuers have to compensate creditors for more than the default risk up to a higher level of yield spread, implying a higher cost of debt related to a weaker risk profile.

Figure II – Marginal Effects of Yield Spread (HY Sample)



Note: The figures illustrate the marginal effect of the yield spread on probability of default for the HY sample. The figure to the top left shows how the marginal effect evolves for different yield spreads and the figure to the top right shows the corresponding effect on default probability. The horizontal axes also display the yield spread frequency. The bottom figure illustrates the ratio between yield spread to default probability for different levels of yield spread.

4.4. Model Evaluation

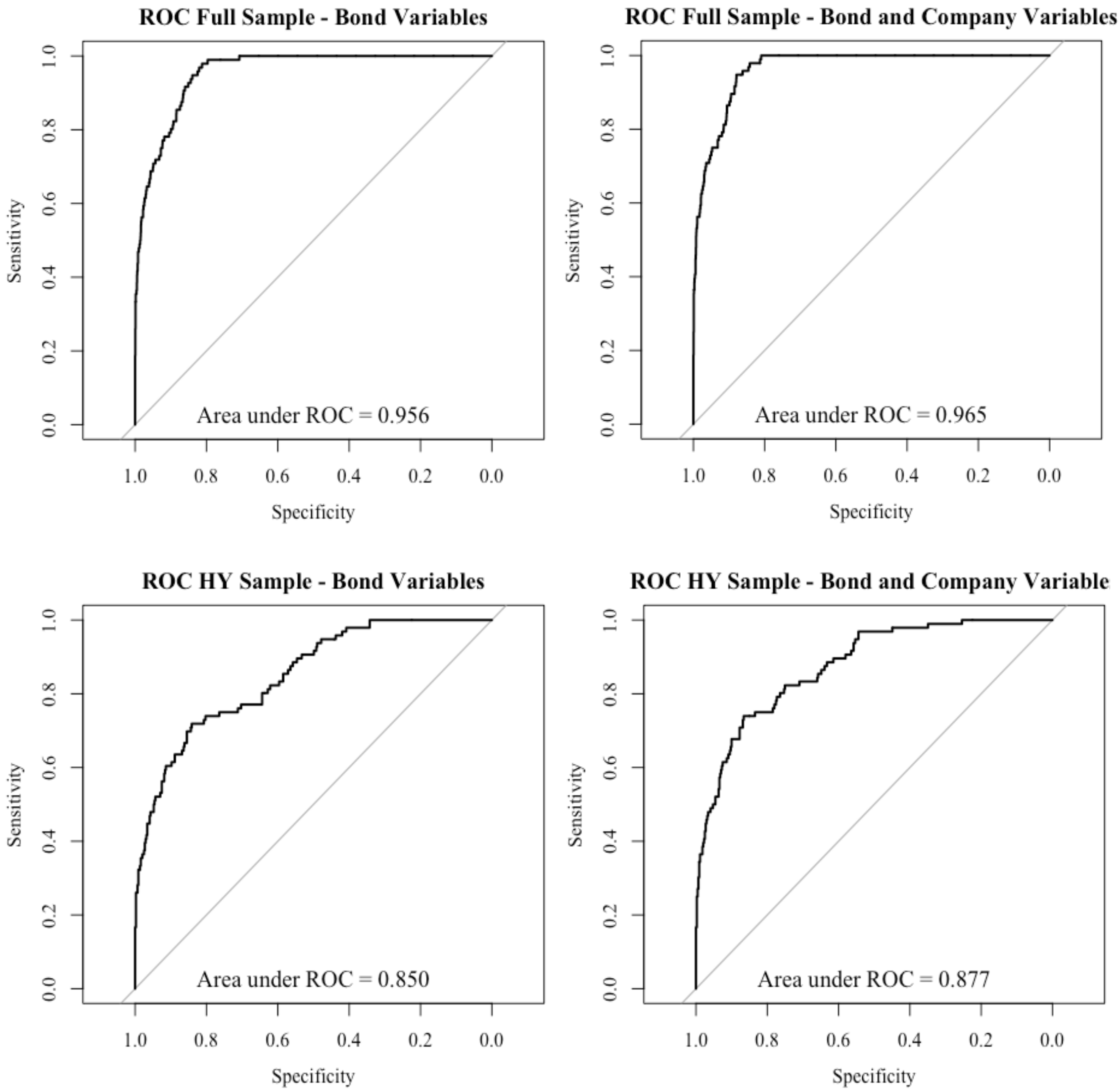
To assess the accuracy of the models and validate our results, we perform a selection of goodness of fit tests, namely Pseudo-R^2, the Hosmer-Lemeshow (H-L) goodness of fit test, and the Area under the Receiver Operating Characteristic (ROC) curve.

The Pseudo-R^2, sometimes referred to as McFadden's R^2, is a measure of the explained variation in a logit regression model. For each model, the Pseudo R^2 tests the explanation value of the model compared to a model only consisting of the intercept without covariates. Unlike the well-known R^2 for OLS linear regression models, the Psuedo-R^2 does not measure the proportion of the variation explained by the model. Instead, Pseudo-R^2 is used to compare models on the same dataset, where a higher value is considered a better fit. In general, Pseudo-R^2 values between 0.2 and 0.4 are considered as an excellent model fit. For our full sample models, the reported Pseudo-R^2 is 0.51 for bond variables only and 0.55 for the model including both bond and company variables. For the HY sample, equivalent reported Pseudo-R^2 values are 0.31 and 0.36, respectively, indicating that all four models have a good fit. Notably, the models including both bond and company data report higher values than the models using only bond data, which proves that the company data contributes to predict default probability but not to the same degree as expected ex-ante.

Furthermore, H-L's goodness of fit test assesses the logit regression models' abilities to correctly predict the binary outcomes. By ranking the predicted probabilities from lowest to highest and divide the sample into subgroups, the H-L test compares if the observed events (default) match the expected events for the subgroups. The model is considered a poor fit if the p-values are lower than 0.05. For our models, the p-values of the H-L statistics range between 0.14 to 0.92, which is well-above the threshold, and we consider all our models as having a good fit.

Lastly, to assess the predictive power of our models, we assess the models' abilities to accurately classify defaulted bonds as defaults ($Y = 1$) and non-defaulted bonds as non-defaults ($Y = 0$). The ROC curve is a probability curve that plots true positives (sensitivity) and false positives (specificity) and the Area Under the Curve (AUC) represents a measure of separability by explaining how capable the model is to distinguish between defaulted and non-defaulted bonds. The higher AUC, the better the model is at performing accurate predictions. Generally, the possible AUC values range between 0.5 and 1.0, and values above 0.8 and 0.9 are considered as excellent and outstanding, respectively. Figure III presents the ROC with corresponding AUC values for each of our four models. For the full sample models, the AUC values are 0.956 and 0.965 and for the HY sample the AUC values are 0.850 and 0.877, for the bond variables only and both bond and company variables, respectively. A potential reason for the higher predictive power for the models on the full bond sample is the relatively lower share of defaults (Y = 1) compared to the HY sample, implying that the risk of misclassification is relatively lower. However, the AUCs for the HY sample with a relatively higher share of defaults still imply high predictive power of our model. As for the previous tests, the model provides higher accuracy of predictability with company variables included in the model. Consequently, we conclude that our models have high predictive power.

Figure III – ROC and AUC Values



Note: The figures illustrate the ROC and the corresponding AUC values for all four models. The AUC value ranges between 0.5 and 0.1, with values above 0.8 and 0.9 considered as excellent and outstanding classification accuracy, respectively.

5. Discussion

The most essential finding in our paper is the strong relationship between the yield spread at issuance on corporate bonds and the probability of default, which is an indication of the efficiency of the participants in the primary Swedish corporate bond market to assess risk, in line with the findings of other papers applying a similar methodology in other markets (Rosengren, 1993; Khan and O'Rawe, 2016). This result is also in line with conventional theory that bond yield spreads should primarily be governed by default risk. Our empirical results regarding the yield spread at issuance are also conforming with the findings of Becker and Josephson (2016), who studied the impact of bankruptcy procedures on the mix of bank and bond financing for companies. Becker and Josephson (2016) derive both theoretical and empirical evidence for that less efficient bankruptcy proceedings lead to less bond issuance for risky firms but not for safe firms, which is in line with the structure of the Swedish corporate bond market and our sample. In explaining their results, they state that out of court resolutions favours concentrated lenders, e.g., banks, over bondholders due their stronger bargaining power. Rational bond investors anticipating this raises their interest rate required to fund risky firms to compensate for being expropriated by banks when distress occurs, which significantly increasing the cost of bond financing for risky firm.

The strong ability of yield spreads to forecast default in our data and the relatively small improvements in Pseudo-R^2 and Area under the ROC curve from adding company-specific variables are in line with the literature on primary issuance of bonds. The evidence points toward that yield spreads at issuance are primarily composed of a default component or credit risk, being already adjusted for the risk-free rate, which in turn would mean that a liquidity component in the primary market is of less importance than in the secondary market, similar to the findings in previous literature on the primary market (e.g., Fridson and Garman, 1998; Gabbi and Sironi, 2002). This interpretation is further supported by the findings of Duan and Rimsaite (2016) who found that credit ratings alone, which is an estimator of creditworthiness, makes up a substantial part of the explanatory power for yield spreads at issuance of bonds issued both in Norway and Sweden. However, for Sweden, the effect is twice as large, indicating that risky issuers need to pay a premium in the less developed HY market in Sweden compared to Norway.

Several other interesting findings are uncovered related to bond characteristics and default probability. Firstly, fixed rate bonds are found to have a statistically significant positive relationship with probability of default relative to floating rate bonds. This relationship is also corroborated by the descriptive statistics which shows that 82 percent of the defaulted sample is made up of fixed rate bonds compared to 39 percent of the non-defaulted sample. There could be both exogenous and endogenous explanations for this relationship. For instance, it could be due to investor preferences and risk mitigation strategies, imagining investors in the HY bond segment, ceteris paribus, preferring fixed rate bonds over floating rate bonds to hedge themselves from additional interest rate risk in addition to the higher credit risk associated with risky borrowers. Consequently, this would incentivise high risk issuers to issue fixed rate bonds to be able to attract funding. This result is also in line with the reasoning of Cooper and Mello (1988) who found that firms with an inherent exposure to inflation risk in their operations are more likely to issue fixed rate bonds as a hedge. This inherent risk would also impact the firm's probability of default and rational investors would demand different compensation for these different types of debt.

Furthermore, longer maturity of bonds is found to be statistically significantly associated with default probability which is supported by the studies finding an aging effect for bonds (Asquith et al., 1989; McDonald and Van De Gucht, 1996; McDonald and Van De Gucht, 1999; Moeller and Molina, 2003), but stands in contrast with the results of Cotter and Peck (2001). An explanation for this finding could be related to the borrowing firm's perception of its own ability to refinance its debt on the market in the long-term. A risky borrower with bleak expectations of its ability to refinance would prefer to issue long-term bonds to secure financing and prolong the time to repayment or refinancing. In addition, the association between default and maturity could also be due to the mechanical relationship that both endogenous and exogenous adverse events are more likely to take place the longer the maturity of a bond.

Another interesting result is the relation of issues in currencies other than SEK and probability of default, with both NOK *(NOK)* and USD *(USD)* being found to be significantly positive. We hypothesise that this result stems from, with the limited domestic market for HY issues, issuers hunting for lower yields in markets more accessible to risky borrowers as discussed by Black and Munro (2010). Norway and the

U.S. would be examples of this as they are among the largest HY bond markets globally (Duan and Rimsaite, 2016). This behaviour is also conforming with the discussion regarding the unfavourable environment experienced by HY issuers in Sweden related to the findings of Becker and Josephson (2016) and Duan and Rimsaite (2016). Given the higher return requirements imposed on risky bond issues from investors, it would be logical for risky borrowers to be motivated to search for lower cost of debt in other markets. Siegfried et al. (2007) similarly finds that the choice to issue bonds in foreign currency by non-financial companies is motivated by decreasing cost, hedging, and a desire to establish a wider investor base. Another explanation could also be that foreign investor are less informed about Swedish issuers than local investors. These results support the notion that HY issuers in Sweden need to adjust their bond offerings based on investor preferences and the market structure of the Swedish corporate bond market rather than necessarily their own corporate financing requirements.

The result that some bond characteristics are not found to have a statistically significant relationship with default probability in either direction differs from the results in other studies. For example, Rosengren (1993) found that HY convertible bonds have a lower rate of default than non-convertible bonds, also when controlling for issue size and coupon rates. However, no such indications could be found in our study.

A surprising result from our logistic models is also the comparatively low increase in Psuedo-R^2 and AUC from adding the company-specific, primarily accounting, variables. Similarly, the generally non-statistically significant results for the financial ratios are surprising. It is noticeable that variables such as profitability, retained earnings, and equity ratio having been found to be significant in other studies of both bond failure (e.g., Huffman and Ward, 1996; Grammenos et al., 2008; Khan and O'Rawe, 2016) and firm failure (e.g., Altman, 1968; Ohlson, 1980) are not significant in our sample. However, Marchesini et al. (2004) also found poor predictive power for financial ratios on bond defaults in comparison. The low degree of predictability could be because defaults of the bonds occur a long time after issuance, leading to the low explanatory power of our financial ratios far out into the future due to data instability over time. This explanation would be supported by the literature related to the aging effect of bonds, which refers to that the default probability increases the longer bonds have been outstanding (McDonald and Van De Gucht, 1996). For example, Moeller and Molina

(2003) found that bonds face constantly increasing default risk over time, with the largest increase beyond four years after issuance, at which point observable characteristics at issuance would have lower explanatory power. These results should also be contrasted to studies in the firm failure prediction literature (e.g., Altman, 1968; Skogsvik, 1990; Duffie et al., 2007) which typically applies financial ratios for default prediction at fixed time periods ahead of default, as close to one year in advance, a factor we are unable to control.

However, the relationships we do find are in line with conventional thinking and would at least be able to give some guidance to investors for their investment decision in Swedish corporate bonds. Dividend payout ratio is found to have a statistically significant negative relationship with probability of default, which is understandable as more financially sound firms are more likely to be able to payout dividends compared to financially distressed firms. Lastly, company age is also a company characteristic that is found to be significantly negatively correlated with probability of default, which is both in line with conventional thinking and previous Nordic studies (e.g., Khan and O'Rawe, 2016), as mature firms will be more stable, have a strong cash balance, and operate in more stable markets, amongst other factors.

6. Limitations and Future Research

It is worthwhile to acknowledge some limitations to the methodology and data applied in this paper to guide the interpretation of our results and future research.

One limitation to our data is the reliance on the classification of HY bonds according to that of Stamdata. Although we deem Stamdata to be a reliable source as the largest provider of bond information in the Nordics, its methodology and criteria for classifying bonds as IG or HY is opaque as no guidance is presented, which means it could differ to the rating methodology used by the official credit rating institutions. However, we chose to use the Stamdata definition for completeness as the Swedish corporate bond market is characterised by a low degree of bonds issued with credit ratings, also seen in our sample where only 65 percent of bonds had a credit rating of which 93 percent had a credit rating equivalent to IG. Furthermore, Stamdata's access to vast data inputs and experience for classification of the bonds and the fact that the same methodology has been relied on in other research papers on the Nordic HY corporate bond market, such as in Grøstad (2013) and Khan and O'Rawe (2016), support our reliance on Stamdata's classifications.

A practical limitation is the timing of the collection of accounting data. We have used the accounting data for the fiscal year prior to the issuance of a bond for the predictive data, similar to the methodology applied in other papers (e.g., Skogsvik, 1990; Khan and O'Rawe, 2016), implicitly assuming that the data is available to investors at the time of issuance. This method was chosen based on a preference to preserve the size of our sample as many issuers had limited financial history at the time of issuance, as well as maintaining statistical relevance of the data for the purpose of default prediction. Depending on the time of the year the issuance occurs, this assumption, however, would not hold true assuming investors only have access to public accounting data when evaluating an issue as annual reports in Sweden are mandated to be reported only within seven months of the end of the fiscal year. It could also be the case that a borrower has issued multiple bonds over the course of a year meaning that the data available would in practice differ between the issuances.

The number of defaulted bonds make up five percent of our full bond sample. From a statistical analysis perspective, Brooks (2019) states that a share of the binary independent outcome below 10 percent could generate different result if another limited independent variable model would be used, such as a probit regression model. However, the tendencies and patterns in our results are similar for both the full bond sample and the HY sample, for which the share of defaults is 18 percent which is well-above the 10 percent threshold stated by Brooks (2019). Combined with the widespread use of the logit regression in the bond default predictability literature, we feel confident in the use of the logit model, although the potential implication of a small default sample is acknowledged.

Furthermore, there are also other exogenous and endogenous factors that could impact a firm's probability of default not accounted for in our models. Such factors that have been found to be useful in previous studies are recent stock market performance and distance to default, which possibly could have improved the predictive power of our models. Our reasons to not include these factors are motivated by a preference for a parsimonious model and to maintain an acceptable sample size of both total number of bonds as well as defaulted bonds to preserve statistical power (Doane and Seward, 2022). This is because many of the variables not included in our model would require an issuer to be publicly listed which would have limited our sample further. Additionally, aggregate variables such as changes in GDP, stock market returns, and stock market volatility were also deemed to be less relevant for our study as we looked at company level default events rather than aggregate default rates. It could also be that the effects of some macroeconomic variables are already impounded in the accounting ratios and the yield spread variable, which takes the level of the risk-free interest rate into account.

We acknowledge that a larger sample of defaulted bonds would generate stronger statistical inference and would have allowed us to incorporate more variables. Considering the rapid development of the Swedish corporate bond market, we would encourage future research to replicate our results with a larger sample at a time when more data is available. A larger sample of defaulted bonds would also allow one to keep a holdout sample that would further provide a better indication of the model validity (Jones, 1987). Other extensions to our research could be to adopt the company variables by a specific industry, such as in Grammenos et al. (2008) and Manning (2004), to better reflect the performance of individual firms, especially since a large share of our sample and the Swedish bond market in general is made up of real estate companies.

Lastly, our definition of the Swedish corporate bond market as bonds issued by Sweden domiciled firms was, apart from input from industry practitioners, also influenced by the availability of consistent and accurate accounting data. This limits the interpretation of our results to issuers matching this criterion as there additionally are foreign firms issuing bonds in Sweden (determined by ISIN) which could also be viewed as part of the Swedish corporate bond market.

7. Conclusion

This paper investigates the relationship between bond yield spreads, and other bond or company characteristics observable at the time of issuance, and the probability of default in the Swedish corporate bond market between the years 2004 and 2023. This has been conducted by matching bond data from Stamdata with firm level accounting data from the Serrano database for the year prior to the bond issuance to approximate the information available to an investor at the time of issuance.

The empirical outcomes of our multiple logit regression models, examining both the full sample and a HY subset, reveal a statistically significant relationship between the yield spread at issuance and probability of default in both samples. The analysis indicates that a constellation of factors, primarily the yield spread but also other bond characteristics such as fixed rated bonds and longer time to maturity, were associated with a higher probability of default. Also, certain company variables such as company age and dividend payout ratio had a significant negative relationship with the likelihood of default. However, the low predictable power of financial ratios in general in our models differ from previous studies using a similar method for predicting corporate failure as well as bond default.

For our main variable of interest, the yield spread at issuance, we also investigate the marginal effect on the probability of default at various levels. In our full sample model, we find that a one percentage point increase in yield spread, on average, increases the probability of default by 0.8 percentage points, and for the HY sample the probability of default increases by 2.2 percentage points by the same metric, ceteris paribus. However, the non-linearity of the logit models imply that the marginal effect increases non-linearly with higher yield spreads. Additionally, we estimate that the optimal level of yield spread to the probability of default is reached at a yield spread of 3.2 percent for the full sample and 4.2 percent for the HY sample. Past these points in the respective samples, at higher levels of yield spread the ratio between yield spread and the default probability diminishes.

The implications of this study are twofold. Firstly, it highlights the complexity and layered nature of risk assessment in the Swedish corporate bond market, underscoring the necessity for investors and issuers to consider a comprehensive range of factors beyond traditional credit ratings due to their low prevalence in the Swedish market. Secondly, the findings in this paper contribute to the understanding of the Swedish corporate bond market by studying a new area of research in a Swedish context, namely determining factors of default. This is vital for practitioners and issuers in understanding the specific market dynamics for informed decision-making.

8. References

Altman, Edward I., 1968, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, The Journal of Finance 23, 589-609.

Altman, Edward I., and Robert G. Haldeman, and P. Narayanan, 1977, Zeta^{TM} analysis, A new model to identify bankruptcy risk of corporations, Journal of Banking & Finance 1, 29-54.

Amato, Jeffrey D., and Eli M. Remolona, 2003, The credit spread puzzle, BIS Quarterly Review, Bank of International Settlement.

Andersson, Oscar, 2021, The Smile of Corporate Bonds: Size Risk Premium on the Swedish Corporate Bond Market, Master Thesis, KTH Royal Institute of Technology.

Asquith, Paul, and David W. Mullins, and Jr., Eric D. Wolff, 1989, Original Issue High Yield Bonds: Aging Analyses of Defaults, Exchanges, and Calls, The Journal of Finance 44, 923-952.

Aziz, Abdul, and David C. Emanuel, and Gerald H. Lawson, 1988, Bankruptcy Prediction - An Investigation Of Cash Flow Based Models, Journal of Management Studies 25, 419-437.

Barr, Daniel, 2011, Den svenska företagsobligationsmarknaden en förstudie, Ackord Centralen.

Beaver, William H., 1966, Financial Ratios as Predictors of Failure, Journal of Accounting Research 4, 71-111.

Becker, Bo, and Jens Josephson, 2016, Insolvency Resolution and the Missing High-Yield Bond Markets, The Review of Financial Studies 29, 2814–2849.

Bellovary, Jodi L., and Michael D. Akers, and Don E. Giacomino, 2007, A review of bankruptcy prediction studies: 1930 to present, Journal of Financial Education 33, 1-42.

Bernhardsen, Eivind, 2001, A model of bankruptcy prediction, Working Paper, No. 2001/10, Norges Bank.

Black, Fischer, and John C. Cox, 1975, Valuing Corporate Securities: Some Effects of Bond Indenture Provisions, The Journal of Finance 31, 351-367.

Black, Susan, and Anella Munro, 2010, Why issue bonds offshore?, BIS Working Paper No 334, Bank of International Settlements.

Blum, Marc, 1974, Failing Company Discriminant Analysis, Journal of Accounting Research 12, 1-25.

Bonthron, Fredrik, 2014, The development of the Swedish market for corporate bonds, Economic Commentaries No. 7, Sveriges Riksbank.

Bonthron, Fredrik, and Tor Johansson, and Jenny Mannent, 2016, Marknadslikviditeten på den svenska obligationsmarknaden och dess betydelse för finansiell stabilitet, Economic Commentary Nr. 3, Sveriges Riksbank.

Brooks, Chris, 2019, Introductory econometrics for finance, fourth edition (Cambridge University Press, Cambridge).

Chen, Long, and David A. Lesmond, and Jason Wei, 2007, Corporate Yield Spreads and Bond Liquidity, The Journal of Finance 62, 119-149.

Cheung, Rayner, and Joseph C. Bencivenga, and Frank J. Fabozzi, 1992, Historical Return And Default Experiences, 1977–1989, The Journal of Fixed Income 2, 58-75.

Collin-Dufresne, Pierre, and Robert S. Goldstein, and J. Spencer Martin, 2001, The Determinants of Credit Spread Changes, The Journal of Finance 56, 2177-2207.

Cooper, Ian, and Antonio S. Mello, 1988, Default spreads in the fixed and in the floating interest rate markets: A contingent claims approach, Advances in Futures and Options Research 3, 269-289.

Cotter, James F., and Sarah W. Peck, 2001, The Structure of Debt and Equity Financing and the Likelihood of Subsequent Default in Leveraged Buyouts, Journal of Financial Economics 59, 101-147.

Delianedis, Gordon, and Robert L. Geske, 2001, The Components of Corporate Credit Spreads: Default, Recovery, Taxes, Jumps, Liquidity, and Market Factors, UCLA Anderson Working Paper NO. 22-01, The Anderson School at UCLA.

Doane, David, and Lori Seward, 2022, Applied Statistics in Business and Economics, (McGraw-Hill Education, New York, NY).

Driessen, Joost, 2005, Is Default Event Risk Priced in Corporate Bonds?, The Review of Financial Studies 18, 165-195

Duan, Cecilia, and Vaiga Rimsaite, 2016, A comparative study: The pricing factors of high yield corporate bond issuances in Norway and Sweden, Master Thesis, Copenhagen Business School.

Duffie, Darrell, and Leandro Saita, and Ke Wang, 2007, Multi-Period Corporate Default Prediction with Stochastic Covariates, Journal of Financial Economics 83, 635–665.

Ekman, Melker, and Andreas Tibell, 2019, Obligationens risker: En studie om kreditrisk, likviditetsrisk och ränterisk för företagsobligationer på den svenska marknaden, Master Thesis, Umeå Universitet.

Elton, Edwin J., and Martin J. Gruber, and Deepak Agrawal, and Christopher Mann, 2001, Explaining the Rate Spread on Corporate Bonds, The Journal of Finance 56, 247-277

Eom, Young H., and Jean Helwege, and Jing-Zhi Huang, 2004, Structural Models of Corporate Bond Pricing: An Empirical Analysis, The Review of Financial Studies 17, 499-544.

Fabozzi, F., 2012, Bond Markets, Analysis and Strategies Global Edition (8th ed., Global ed.). Pearson Education M.U.A.

Fridson, Martin S., and M. Christopher Garman, 1998, Determinants of Spreads on New High-Yield Bonds, Financial Analysts Journal 54, 28-39.

Frisén, Hampus, and Jacob L. Lövstedt, 2015, Underpricing in Corporate Bond Issues: An Explanation Based on Information Problems, Master Thesis, Stockholm School of Economics.

Gabbi, Giampaolo, and Andrea Sironi, 2002, Which Factors Affect Corporate Bond Pricing? Empirical Evidence from Eurobonds Primary Market Spreads, Newfin Research Center, Bocconi University.

Gentry, James A., Paul Newbold, and David T. Whitford, 1985, Classifying Bankrupt Firms with Funds Flow Components, Journal of Accounting Research 23, 146-160.

Grammenos, C.Th., and N.K. Nomikos, and N.C. Papapostolou, 2008, Estimating the probability of default for shipping high yield bond issues, Transportation Research Part E: Logistics and Transportation Review 44, 1123-1138.

Greenwood, Robin M., and Samuel G. Hanson, 2011, Issuer Quality and the Credit Cycle, Harvard Business School Working Paper No. 1734528, Harvard University.

Grøstad, Kristian. N., 2013, Predicting default in the Norwegian High Yield bond market: a study of defaults in the years 2006-2013, Master Thesis, Norwegian School of Economics.

Gunnarsdottir, Gudrun, and Sofia Lindh, 2011, Marknader för svenska icke-finansiella företags lånebaserade finansiering, Penning- och valutapolitik 2011:2, Sveriges Riksbank.

Göransson, Philip, and Markus Olofsgård, 2020, An analysis of the underlying variables on the credit spread of the Swedish corporate bond market, Bachelor Thesis, KTH Royal Institute of Technology.

Huang, Jing-Zhi, and Ming Huang, 2012, How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk?, Review of Asset Pricing Studies 2, 154-202.

Huffman, Stephen P., and David J. Ward, 1996, The prediction of default for high yield bond issues, Review of Financial Economics 5, 75-89.

Jones, E. Philip, and Scott P. Mason, and Eric Rosenfeld, 1984, Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation, Journal of Finance 39, 611-627.

Jones, F., 1987, Current techniques in bankruptcy prediction, Journal of Accounting Literature 6, 131-164.

Khan, Jamal F., and Alastair O'Rawe, 2016, Default in the Nordic High-Yield Bond Market, Master Thesis, Norwegian School of Economics.

Lin, Hai, Sheen Liu, and Chunchi Wu, 2011, Dissecting Corporate Bond and CDS Spreads, The Journal of Fixed Income, 7-39.

Longstaff, Francis A., and Sanjay Mithal, and Eric Neis, 2005, Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market, The Journal of Finance 60, 2213-2253.

Manning, Mark J., 2004, Exploring the relationship between credit spreads and default probabilities, Working paper no. 225, Bank of England.

Marchesini, R., and V. Bryan, and G. Perdue, 2004, Applying Bankruptcy Prediction Models to Distressed High Yield Bond Issues. The Journal of Fixed Income 13, 50-56.

McDonald, Cynthia G., and Linda M. Van De Gucht, 1996, The Default Risk of High-Yield Bonds, Onderzowksrapport Nr 9641, Katholieke Univesiteit Leuven.

McDonald, Cynthia G., and Linda M. Van De Gucht, 1999, High-Yield Bond Default and Call Risks, The Review of Economics and Statistics 81, 409-419.

Merton, Robert C., 1974, On The Pricing Of Corporate Debt: The Risk Structure Of Interest Rates, The Journal of Finance 29, 449-470.

Moeller, Thomas, and Carlos A. Molina, 2003, Survival and Default of Original Issue High-Yield Bonds, Financial Management 32, 83-107.

Nordic Trustee, 2023, Corporate Bond Market Report 2022.

OECD iLibrary, 2022, The Swedish Corporate Bond Market and Bondholder Rights, https://www.oecd-ilibrary.org/sites/c443352d-en/1/3/1/index.html?itemId=/content/publication/c443352d-en&_csp_=946b2b90ec19e539ebc8c8b22c33e20e&itemIGO=oecd&itemContentType=book.

Ohlson, James A., 1980, Financial Ratios and the Probabilistic Prediction of Bankruptcy, Journal of Accounting Research 18, 109-131.

Rosengren, Eric S., 1993, Defaults of Original Issue High-Yield Convertible Bonds, The Journal of Finance 48, 345-362.

Shumway, Tyler, 2001, Forecasting Bankruptcy More Accurately: A Simple Hazard Model, The Journal of Business 74, 101-124.

Siegfried, Nikolaus, Emilia Simeonova, and Cristina Vespro, 2007, Choice of Currency in Bond Issuance and the International Role of Currencies, Working Paper Series No 814, European Central Bank.

Skogsvik, Kenneth, 1990, Current cost accounting ratios as predictors of business failure: The Swedish case, Journal of Business Finance & Accounting 17, 137-160.

Statistics Sweden, SCB, Statistikdatabasen Finansräkenskaper, 2020.

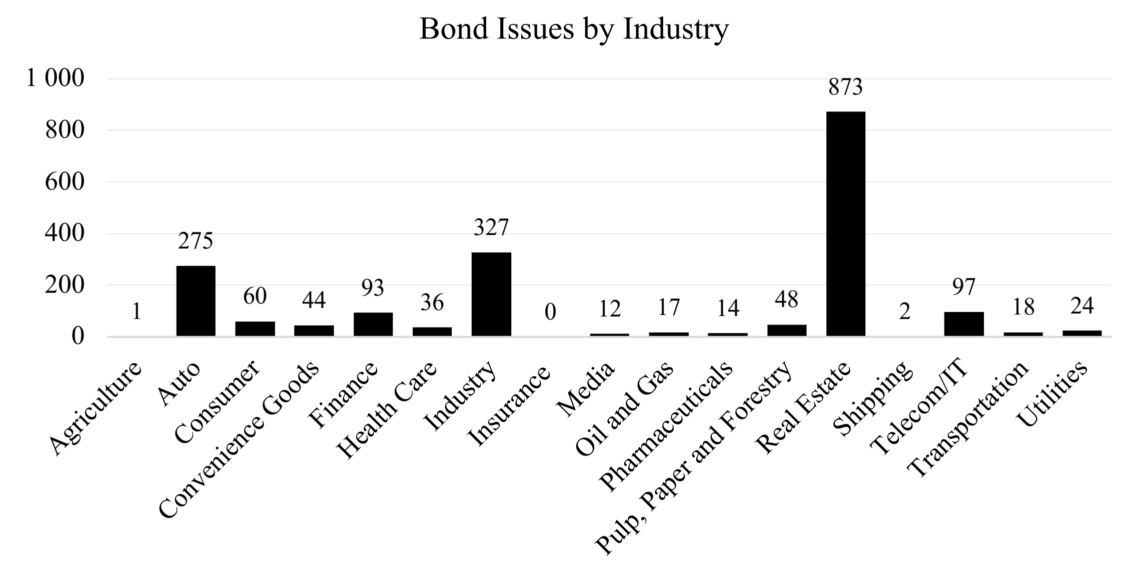
Wollert, Stephan, 2020, Svenska företagsobligationer under coronapandemin, Staff Memo, Sveriges Riksbank.

Zavgren, Christine V., 1985, Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis, Journal of Business Finance and Accounting 12, 19-45.

9. Appendix

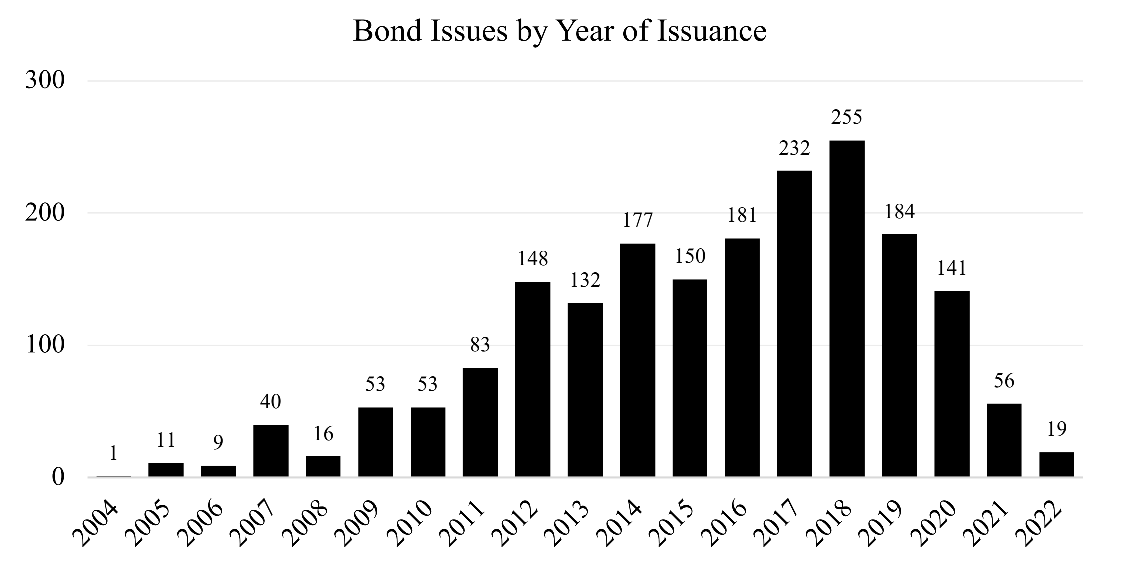
AI – Bond Sample Characteristics

Figure AI-I – Number of Bonds per Industry



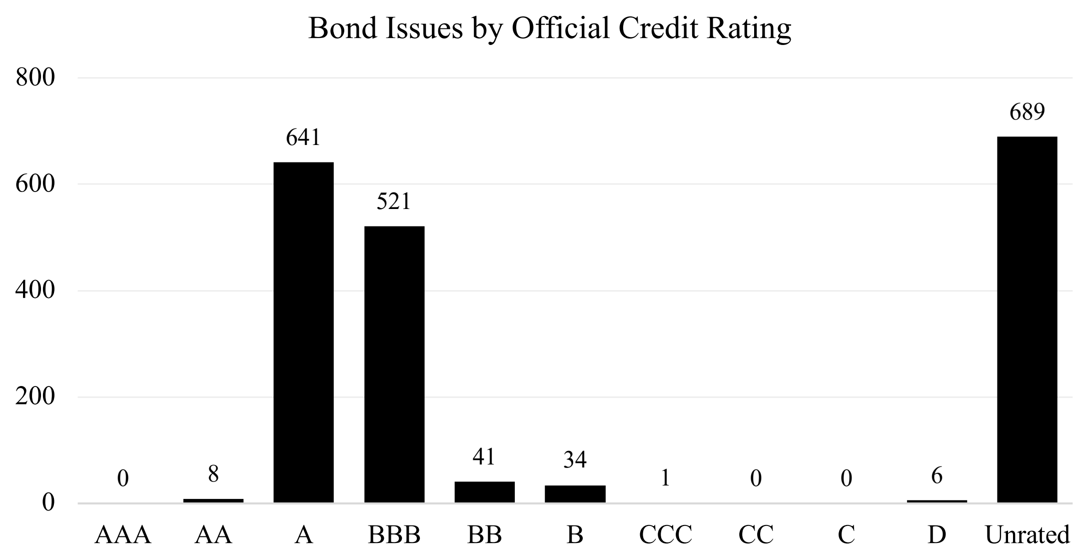
Note: The figure displays the number of bonds issued per industry in our data sample. Real Estate is the most frequent industry followed by Industry and Auto.

Figure AI-II – Number of Bond Issues per Year



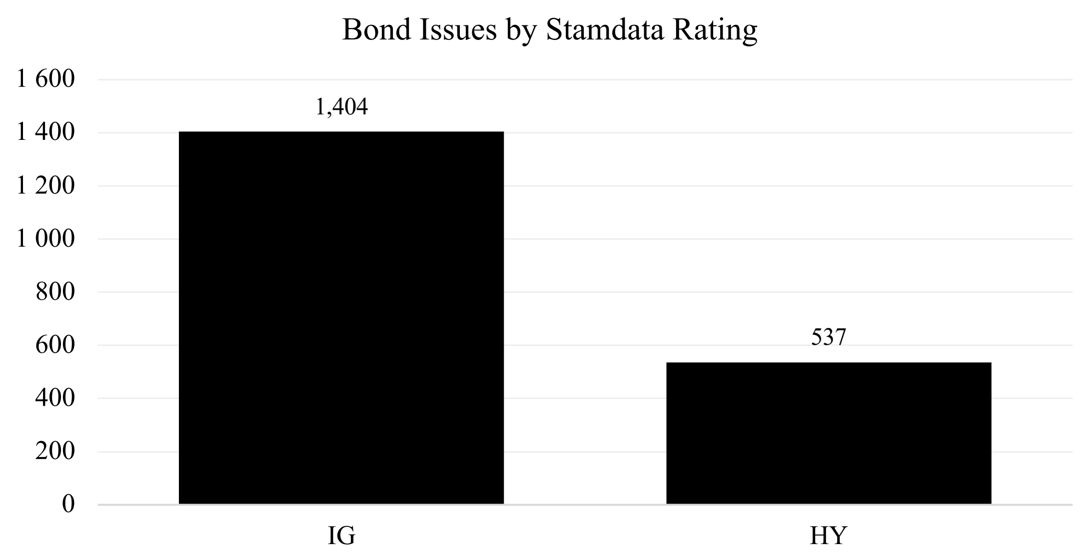
Note: The figure displays the number of bonds issued year. As expected, the number of issues is skewed towards more recent years as the development of the Swedish corporate bond market has increased with over time.

Figure AI-III – Number of Bonds per Official Credit Rating



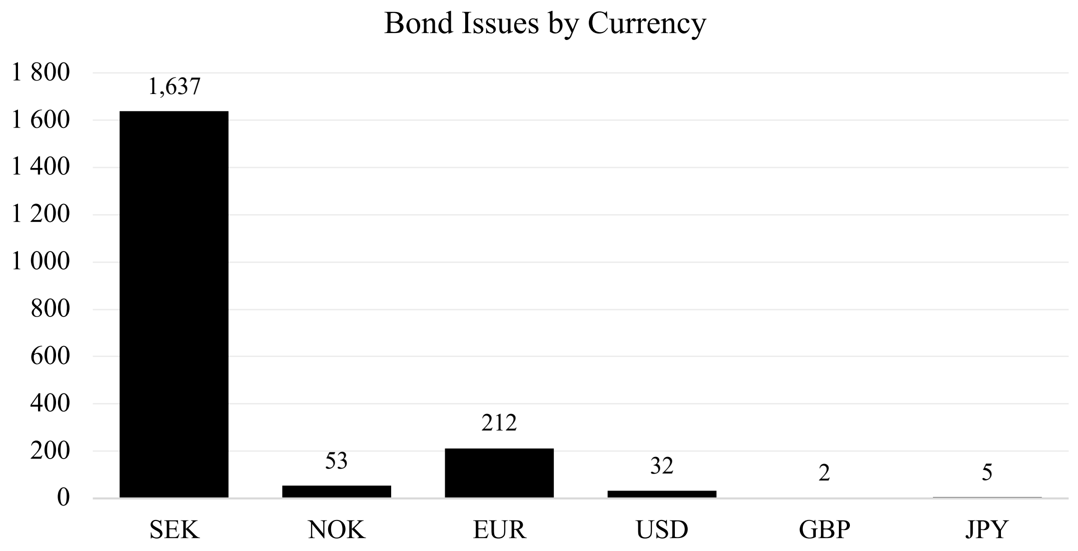
Note: The figure displays the number of bonds with an official credit rating. Notable is the large share of both A-rated and BBB-rated, and unrated bonds, as well as the low share of HY-rated bonds.

Figure AI-IV – Number IG and HY Bonds by Stamdata Classification



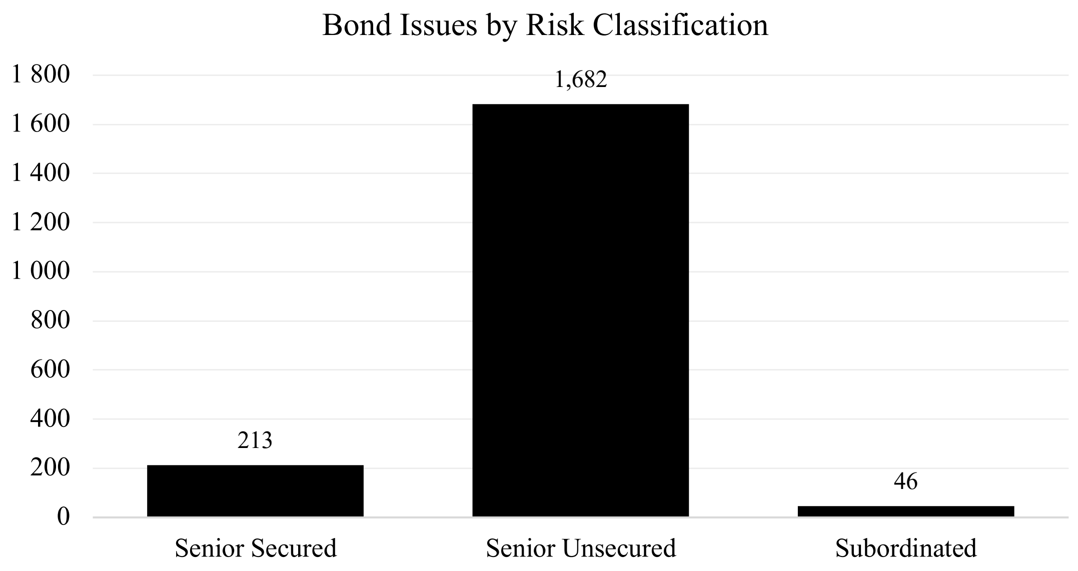
Note: The figure displays the number of bonds rated as IG and HY according to Stamdata's rating classification.

Figure AI-V – Number of Bond Issues by Currency



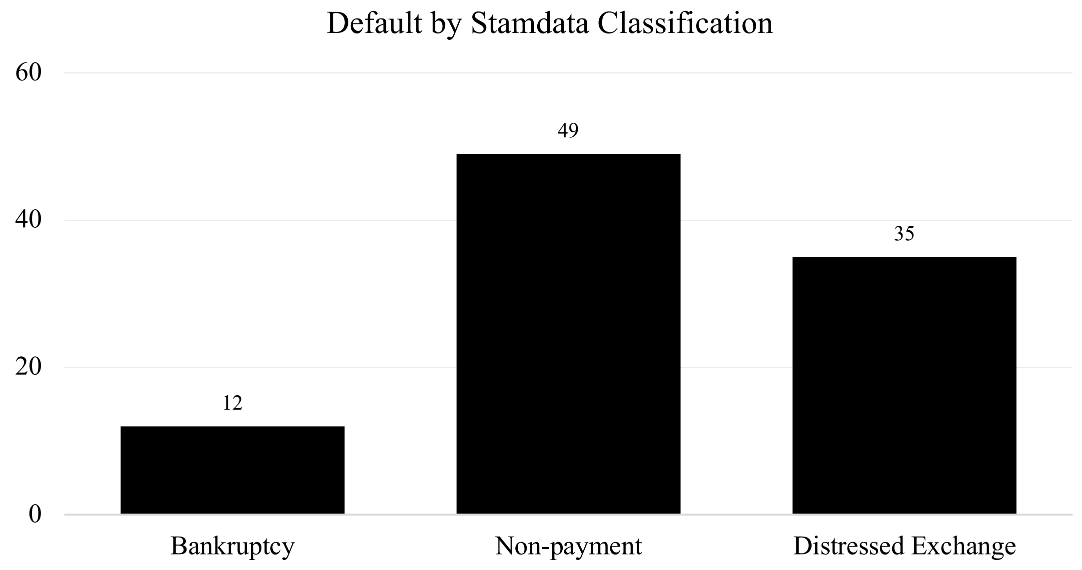
Note: The figure displays the number of bonds issued in each currency. Bonds issued in SEK dominates the data sample, followed by EUR and NOK.

Figure AI-VI – Distribution of Risk Classification



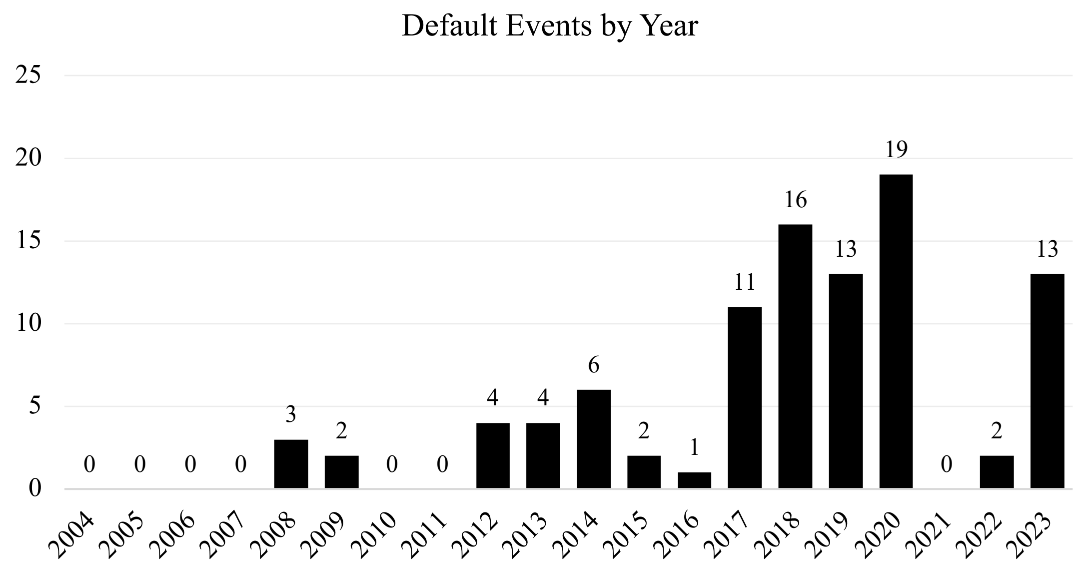
Note: The figure displays the distribution of risk classification among the bonds in the data sample.

Figure AI-VII – Distribution of Default Events



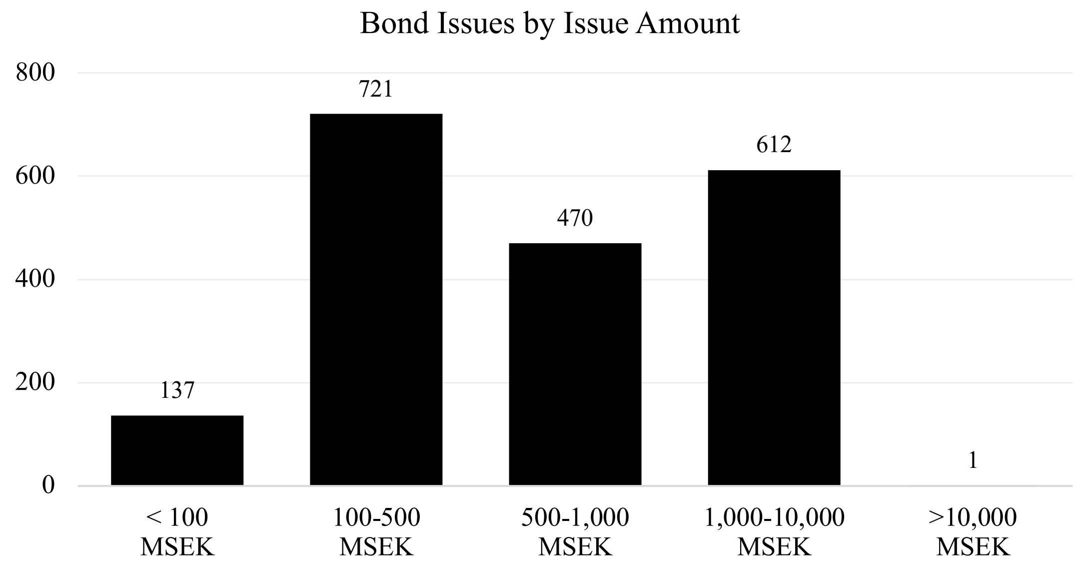
Note: The figure displays the number of bond defaults per underlying cause of financial distress, as reported by Stamdata.

Figure AI-VIII – Distribution of Default Events per Year



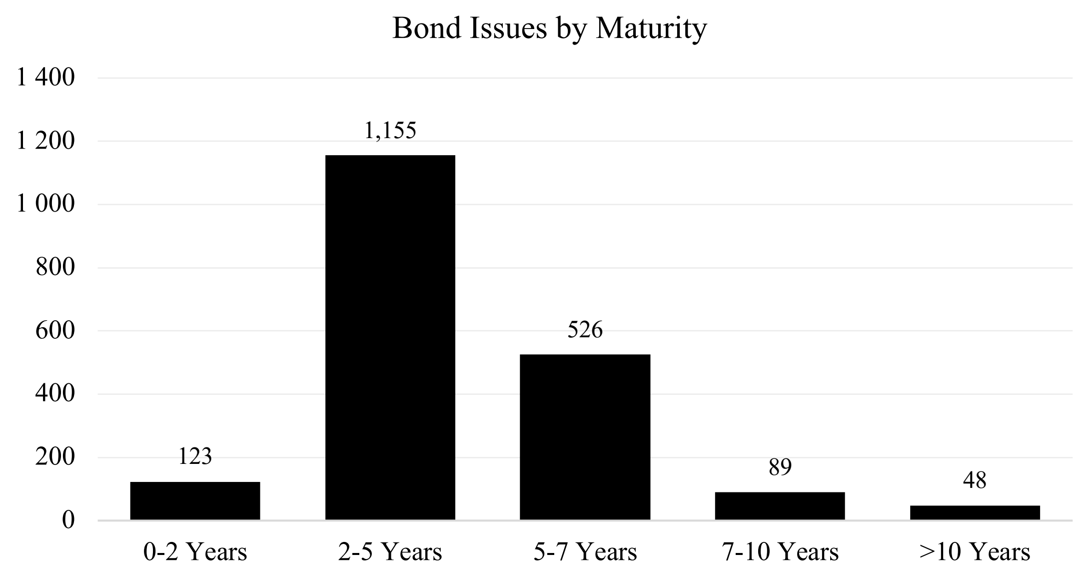
Note: The figure displays the number of bond defaults events per year.

Figure AI-IX – Distribution of Issue Amount



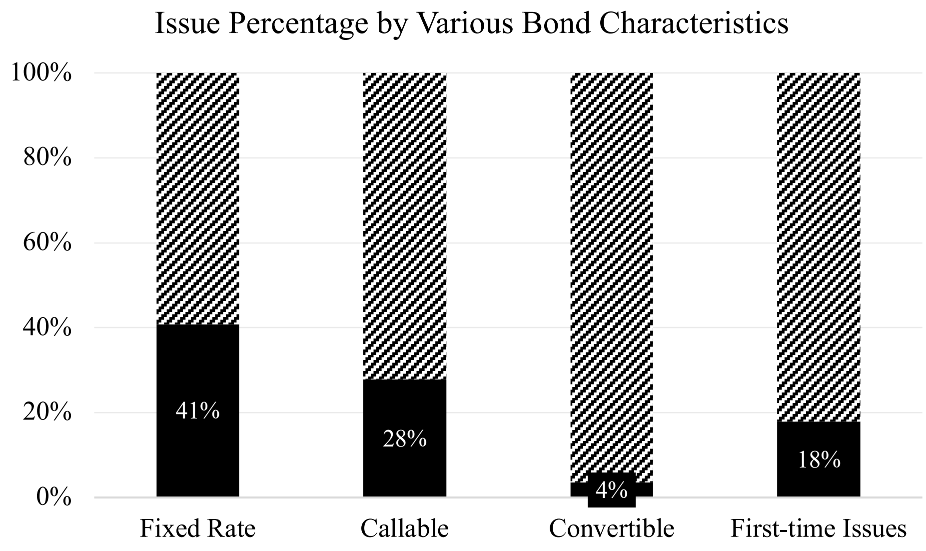
Note: The figure displays the number of bond issues clustered in different tranches dependent on the amount issued for the bond. Bond issues ranging between 100 to 500 MSEK have the highest representation in the data sample.

Figure AI-X – Distribution of Time to Maturity



Note: The figure displays the number of bond issues clustered in different tranches dependent on the time to maturity. Most bonds in the data sample matures within two to five years.

Figure AI-XI – Distribution of Various Bond Characteristics



Note: The figure displays the percentage of bonds with fixed coupon rate, a call option, a convertible option, and first-time issues, respectively. The shaded areas represent the counter-alternative for each bond characteristic, namely, floating coupon rate, no call option, no convertible option, and not a first-time issue, respectively.

AII – Logit Regression Results

Table AII presents a full overview of the results for all variables in all models of the logit regressions.

Table AII – Logit Regression of Default Estimates

	Model Specification			
Variable	FS – B	FS – B & C	HY – B	HY – B & C
YIELD	0.36***	0.33***	0.26***	0.24***
	(0.06)	(0.06)	(0.06)	(0.06)
FIXED	0.63*	0.78**	0.80**	0.88**
	(0.36)	(0.39)	(0.37)	(0.40)
CALL	1.37***	0.86*	0.29	-0.02
	(0.48)	(0.51)	(0.51)	(0.55)
CONV	0.76	0.62	-0.22	-0.20
	(0.62)	(0.69)	(0.62)	(0.69)
ZEROC	2.78***	2.97***	2.81***	2.71**
	(0.92)	(1.09)	(1.06)	(1.11)
FIRST	0.48	0.23	0.29	0.14
	(0.30)	(0.33)	(0.29)	(0.32)
SENIOR UNSECURED	-0.41	-0.25	-0.43	-0.33
	(0.37)	(0.39)	(0.36)	(0.38)
SUBORDINATED	-1.22*	-0.85	-1.33*	-1.12
	(0.74)	(0.83)	(0.71)	(0.80)
MATURITY	0.05**	0.07***	0.10**	0.12***
	(0.02)	(0.03)	(0.04)	(0.05)
AMOUNT	-0.13	0.001	-0.10	-0.01
	(0.10)	(0.13)	(0.10)	(0.12)
NOK	2.74***	2.85***	2.81***	2.82***
	(0.56)	(0.67)	(0.61)	(0.71)
EUR	-0.40	-0.63	-0.41	-0.69
	(0.46)	(0.50)	(0.47)	(0.51)
USD	1.45**	1.55**	1.43**	1.55**
	(0.62)	(0.70)	(0.60)	(0.69)
GBP	-10.96	-9.72		
	(1,683.25)	(1,562.54)		
JPY	-10.63	-8.81		
	(1,072.78)	(998.15)		
EBITMARG		-0.0000		-0.0001
		(0.0002)		(0.0002)
ROE		0.001		0.001

		(0.001)		(0.001)
ROA		0.0000		0.002
		(0.01)		(0.01)
RETAIN		0.001		0.0005
		(0.004)		(0.005)
ICR		0.0000		0.0001
		(0.0001)		(0.0001)
DPR		-0.01***		-0.005**
		(0.002)		(0.002)
ISSIZE		-0.0000		0.0001
		(0.0004)		(0.0003)
WCTA		0.01		0.01
		(0.01)		(0.01)
CACL		-0.002*		-0.002*
		(0.001)		(0.001)
CASH		-0.04*		-0.04*
		(0.02)		(0.02)
REVTA		-0.001		-0.002
		(0.003)		(0.003)
SOLVENCY		0.01		0.01
		(0.01)		(0.01)
COMPSIZE		-0.08		-0.07
		(0.07)		(0.07)
COMPAGE		-0.04***		-0.03**
		(0.01)		(0.02)
Constant	-3.81*	-3.20	-2.66	-1.97
	(2.12)	(2.43)	(2.04)	(2.33)
Observations	1,941	1,941	537	537
Psuedo-R^2	0.51	0.55	0.31	0.36
H-L statistic	14.92 [0.38]	18.16 [0.92]	17.39 [0.14]	31.99 [0.19]
Mean VIF	1.41	1.62	1.42	1.61

Note: * p<0.1, ** p<0.05, *** p<0.01. The table displays the coefficients for the logit regression models. Standard errors are in the parentheses (·). H-L statistic refers to the Hosmer-Lemeshow test statistic, with corresponding p-values in brackets [·].

AIII – Average Marginal Effects

Table AIII presents a full overview of the average marginal effects (AME) for all variables.

Table AIII – Overview of Average Marginal Effects

	Model Specification			
Variable	FS – B	FS – B & C	HY – B	HY – B & C
YIELD	0.010***	0.008***	0.025***	0.022***
	(0.002)	(0.002)	(0.005)	(0.006)
FIXED	0.017*	0.019**	0.078**	0.079**
	(0.010)	(0.010)	(0.036)	(0.036)
CALL	0.037***	0.022*	0.028	-0.001
	(0.013)	(0.013)	(0.049)	(0.050)
CONV	0.020	0.016	-0.021	-0.018
	(0.017)	(0.017)	(0.060)	(0.062)
ZEROC	0.074***	0.074***	0.274***	0.245**
	(0.024)	(0.027)	(0.101)	(0.098)
FIRST	0.013	0.006	0.028	0.013
	(0.008)	(0.008)	(0.028)	(0.029)
SENIOR UNSECURED	-0.012	-0.007	-0.044	-0.031
	(0.010)	(0.010)	(0.038)	(0.036)
SUBORDINATED	-0.029**	-0.019	-0.111**	-0.089*
	(0.015)	(0.017)	(0.049)	(0.054)
MATURITY	0.001**	0.002***	0.010**	0.011***
	(0.001)	(0.001)	(0.004)	(0.004)
AMOUNT	-0.004	0.000	-0.010	-0.001
	(0.003)	(0.003)	(0.009)	(0.011)
NOK	0.136***	0.130***	0.437***	0.401***
	(0.041)	(0.043)	(0.106)	(0.114)
EUR	-0.010	-0.014	-0.038	-0.057
	(0.011)	(0.010)	(0.040)	(0.037)
USD	0.055*	0.056*	0.192*	0.195*
	(0.031)	(0.033)	(0.098)	(0.105)
GBP	-0.043***	-0.045*		
	(0.006)	(0.024)		
JPY	-0.043***	-0.045		
	(0.006)	(0.037)		
EBITMARG		0.000		0.000
		(0.000)		(0.000)
ROE		0.000		0.000

	(0.000)	(0.000)
ROA	0.000	0.000
	(0.000)	(0.001)
RETAIN	0.000	0.000
	(0.000)	(0.000)
ICR	0.000	0.000
	(0.000)	(0.000)
DPR	0.000***	0.000**
	(0.000)	(0.000)
ISSIZE	0.000	0.000
	(0.000)	(0.000)
WCTA	0.000	0.001
	(0.000)	(0.001)
CACL	0.000*	0.000*
	(0.000)	(0.000)
CASH	-0.001*	-0.003*
	(0.001)	(0.002)
REVTA	0.000	0.000
	(0.000)	(0.000)
SOLVENCY	0.000	0.001
	(0.000)	(0.001)
COMPSIZE	-0.002	-0.006
	(0.000)	(0.006)
COMPAGE	-0.001***	-0.003**
	(0.000)	(0.001)

Note: * p<0.1, ** p<0.05, *** p<0.01. The table displays the average marginal affects for all variables. Standard errors are in the parentheses (·).

AIV – Variance Inflation Factor (VIF)

Table AIV presents the Variance Inflation Factor (VIF) for all variables included in the logit regression models. The HY sample did not consist of any bonds issued in GBP or JPY.

Table AIV – Variance Inflation Factor (VIF)

	VIF Values			
Variable	FS – B	FS – B & C	HY – B	HY – B & C
Bond Variables				
YIELD	2.00	2.25	1.64	1.82
FIXED	1.40	1.55	1.47	1.63
CALL	2.01	2.16	2.08	2.27
CONV	1.91	2.24	1.96	2.26
ZEROC	1.11	1.15	1.08	1.16
FIRST	1.15	1.28	1.07	1.20
SENIOR UNSECURED	1.74	1.74	1.61	1.62
SUBORDINATED	1.53	1.92	1.44	1.76
MATURITY	1.43	1.66	1.29	1.34
AMOUNT	1.52	2.21	1.48	2.17
NOK	1.09	1.25	1.06	1.23
EUR	1.14	1.27	1.18	1.31
USD	1.17	1.35	1.13	1.35
GBP	1.00	1.00		
JPY	1.00	1.00		
Company Variables				
EBITMARG		1.48		1.49
ROE		1.11		1.13
ROA		1.64		1.61
RETAIN		1.75		1.77
ICR		1.10		1.08
DPR		1.19		1.26
ISSIZE		1.61		1.61
WCTA		2.14		2.01
CACL		1.72		1.73
CASH		1.78		1.70
REVTA		1.34		1.36
SOLVENCY		1.99		1.96
COMPSIZE		2.66		2.41
COMPAGE		1.37		1.31

Mean VIF	1.41	1.62	1.42	1.61
Lowest VIF	1.00	1.00	1.06	1.08
Highest VIF	2.01	2.66	2.08	2.41

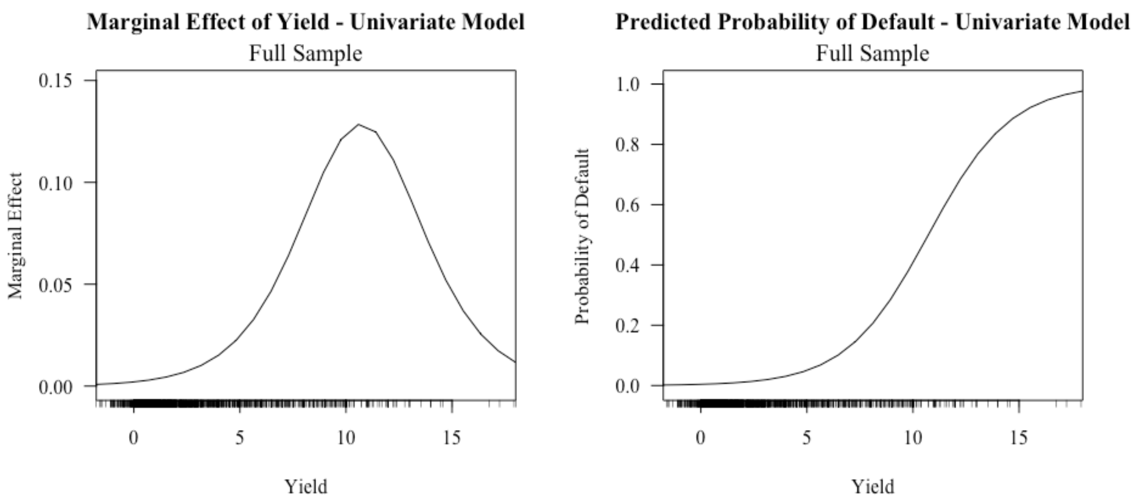
Note: The table displays the Variable Inflation Factor (VIF) for all separate models. The threshold VIF value for multicollinearity to be a concern for the model is equal to or greater than 5. The HY sample does not have any bonds issued in GBP and JPY.

AV – Univariate Marginal Effect of Yield Spread

When conducting a univariate analysis with the yield spread at issuance as the sole regressor for probability of default, the dynamic of the marginal effects of the yield spread is clearly illustrated. Note that this univariate model does not have the same predictive and explanatory power as our four models in the main analysis, yet the results provide interesting insights in how the yield spread could be indicative for a higher risk of default.

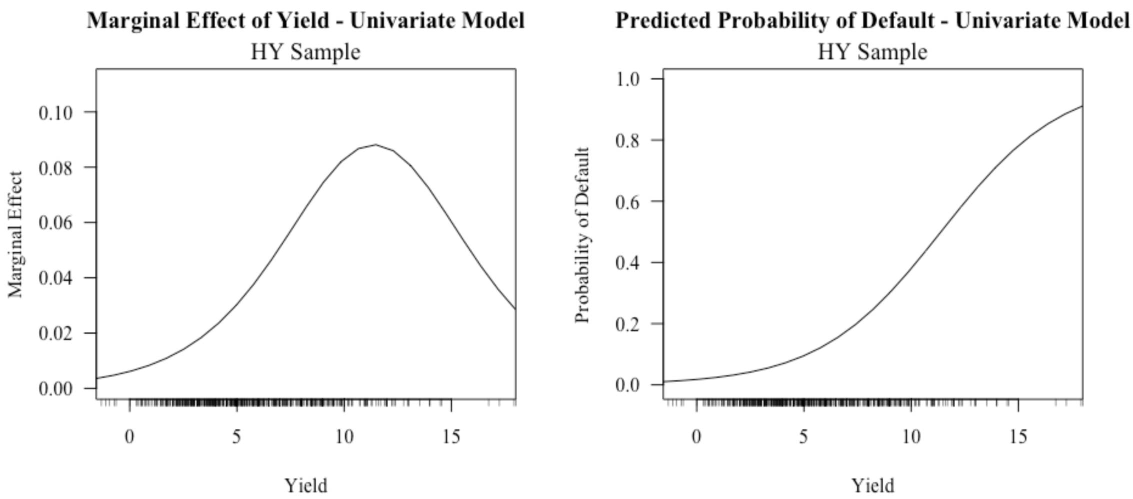
For the full sample, Figure AV-I displays a peak in the marginal effect at a yield spread of approximately 11 percent. At this point, on average, an increase of one percentage unit in yield would imply an increase in default probability of over 12 percent, ceteris paribus. This is also evident in the predicted probability of default curve which increases steeply from around seven to 12 percent yield spreads, increasing the probability of default from below 20 percent to over 80 percent. Further, for the HY sample (Figure AV-II), the marginal effect and probability of default follow a similar pattern, although with smoother effects. Also, the effects are not as distinct as for the full sample, with a maximum marginal effect below nine percent.

Figure AV-I – Univariate Marginal Effects of Yield Spread (Full Sample)



Note: The figures illustrate the univariate marginal effect of the yield spread on probability of default for the full sample. The figure to the left shows how the marginal effect evolves for different yield spreads and the figure to the right shows the corresponding effect on default probability. The horizontal axis displays the yield spread frequency.

Figure AV-II – Univariate Marginal Effects of Yield Spread (HY Sample)



Note: The figures illustrate the univariate marginal effect of the yield spread on probability of default for the HY sample. The figure to the left shows how the marginal effect evolves for different yield spreads and the figure to the right shows the corresponding effect on default probability. The horizontal axes display the yield spread frequency.