

A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics.

**Predicting Default in High Yield Bond Markets: A Comprehensive Regression Approach
Evaluating the Significance of Market Sentiment Indicators**

Julia Hilberath

Work project carried out under the supervision of:

Dr. Giorgio Ottonello

20/12/2023

Abstract

This thesis addresses the challenge of predicting defaults in high-yield bond markets by integrating forward-looking market data into regression analysis. While traditional models focus on firm-specific financial ratios, this study emphasizes the pivotal role of market sentiment data. These indicators capture real-time market dynamics and risk perceptions, offering a nuanced view of investor sentiment. In contrast, financial ratios, derived from historical financial statements, lack the timeliness to predict defaults effectively. The inclusion of market sentiment data enriches the model, providing a comprehensive and timely assessment of default risk in high-yield bond markets.

#High-Yield #Corporate Bonds #Market Sentiment #Regression #Default Prediction

Table of Contents

1. Introduction	5
2. Literature Review	6
2.1 Accounting Models for Bankruptcy Prediction.....	6
2.2 Market Data Models for Default Prediction.....	7
2.3 Bankruptcy Models for Default Prediction	8
3. Data and Methodology	9
3.1 Data Sample	9
3.2 Selection of Dependent and Independent Variables.....	11
3.2.1 Dependent Variable.....	12
3.2.2 Independent Variables	12
3.3 Methodology	15
4. Results and Discussion.....	17
4.1 Regression with Market Sentiment Indicators	17
4.2 Regression without Market Sentiment Indicators	18
4.3 Discussion	20
4.4 Scope and Limitations	24
5. Conclusion.....	26
Appendix	28
Bibliography.....	36

List of Figures

Figure 1. Regression 1 - Linearity between the Independent and Dependent Variables	28
Figure 2. Regression 1 - Mean of Residuals	28
Figure 3. Regression 1 - Check for Homoscedasticity	28
Figure 4. Regression 1 - Normality of Residuals	29
Figure 5. Regression 1 - Autocorrelation for Residuals	29
Figure 6. Regression 1 - Correlation of Variables	30
Figure 7. Regression Results with Market Sentiment Variables	31
Figure 8. Regression 2 - Linearity between the independent and dependent variables	32
Figure 9. Regression 2 - Mean of Residuals	32
Figure 10. Regression 2 - Check for Homoscedasticity	32
Figure 11. Regression 2 - Normality of Residuals	33
Figure 12. Regression 2 - Autocorrelation for Residuals	33
Figure 13. Regression 2 - Correlation of Variables	34
Figure 14. Regression Results without Market Sentiment Variables	35

1. Introduction

Credit risk is a much-discussed topic among financial economists as it affects every financial contract (Altman, Brady, et al. 2005) and can result in significant financial losses for both, creditors and stockholders (Deakin 1972). Within this context, the cyclic ascent, decline, and subsequent resurgence of high-yield (HY) bonds (Marchesini, Perdue and Bryan 2004) has sparked considerable debate. Over the years HY bonds have progressively gained wider acceptance among both investors and corporations seeking debt financing, earning a place in many portfolios for enhanced risk-return characteristics (Chaplinsky 2013; Marchesini, Perdue and Bryan 2004). Many attempts have been made to build models that forecast financial distress or business failure; as a result, a great deal of work has gone into modelling default risk to value corporate debt and the derivatives attached to it (Vassalou and Xing 2002). Whilst many established models have demonstrated reliability in predicting bankruptcy, they have not achieved success in forecasting defaults in HY bond markets (Marchesini, Perdue and Bryan 2004). Moreover, most existing models have been constructed using accounting data, therefore relying on financial statements that provide insights into a firm's historical performance rather than its prospects and default potential (Vassalou and Xing 2002). Consequently, this thesis seeks to present a new model that integrates not only financial ratios but also forward-looking market data, aiming to enhance the accuracy of potential default predictions.

The primary goal of this research is to offer a novel perspective on the determinants to consider in default prediction, thereby augmenting traditional models. This study aims to scrutinize and rectify the significance of systematic market data as opposed to the exclusive reliance on idiosyncratic data, which has been the sole focus of prior models. Therefore, this thesis endeavors to elucidate the magnitude of incorporating systematic data.

2. Literature Review

The next chapter provides a comprehensive literature review on established models used for predicting bankruptcy and default events. It is crucial to note that while bankruptcy and debt default are distinct, both indicate stages of financial distress (Marchesini, Perdue, and Bryan 2004). The literature review therefore encompasses studies on both default prediction and established models for bankruptcy forecasting. It starts by highlighting accounting-based models for bankruptcy prediction and transitions to research emphasizing forward-looking market data for default prediction. The review concludes by examining a study that integrates bankruptcy models for the specific purpose of default prediction.

2.1 Accounting Models for Bankruptcy Prediction

Altman's Z-score (1968) is a renowned model that utilizes discriminant analysis with five weighted financial ratios (Working Capital/Total Assets, Retained Earnings/Total Assets, EBIT/Total Assets, Market Value of Equity/Book Value of Total Debt, and Sales/Total Assets) to predict corporate bankruptcy for manufacturing corporations from 1946 to 1965. A Z-Score above 2.99 classifies firms as 'non-bankrupt,' below 1.81 as 'bankrupt,' and the range in-between as the 'zone of ignorance.' This model achieved 94% accuracy in predicting bankruptcy in the initial sample.

Recognizing limitations and changes in financial reporting standards, Altman, Haldeman, and Narayanan (1977) developed the ZETA™ Credit Risk Model as an enhanced version. Comprising seven variables (Return on Assets, Stability of Earnings, EBIT/Total Interest Payments, Retained Earnings/Total Assets, Current Assets/Current Liabilities, 5-year average Market Value of Equity/Total Capital, and Total Assets), it demonstrated improved accuracy. The ZETA™ model achieved accuracy ranging from 82.4% to 97% for non-

bankrupt companies and from 86.8% to 96.2% for bankrupt companies one year before default (1969-1975). Notably, accuracy declined leading up to five years before default.

Other accounting-based attempts on predicting business failure were conducted by Beaver (1966) and Deakin (1972) who found that ratio analysis effectively differentiates between failed and non-failed firms up to five years before failure.

2.2 Market Data Models for Default Prediction

Vassalou and Xing (2002) criticize the reliance on accounting models for predicting corporate default risk. They argue that accounting models, which rely on data from financial statements focused on past performance, neglect future prospects. Instead, they advocate for the Merton model (1974), which uses forward-looking market values of equity and debt. Grounded in option-pricing theory, the Merton model considers asset volatility as a crucial factor in assessing default risk, unlike accounting models. Vassalou and Xing conducted a study based on this model, finding a strong connection between a company's default risk and its size as well as book-to-market traits. Their research suggests that both size and book-to-market characteristics influence default risk, with size being particularly noteworthy. Asset pricing tests confirm that default risk is systematic. However, they highlight the research work of Opler and Titman (1994) as well as Asquith, Gertner and Scharfstein (1994) who indicated in previous studies that bankruptcy is rather related to idiosyncratic factors. Opler and Titman (1994) find a positive correlation between a firm's performance and financial health during industry downturns, especially in distressed industries. Asquith, Gertner, and Scharfstein (1994) emphasize the impact of debt composition on the restructuring of financially distressed firms, noting challenges in out-of-court restructurings and the effectiveness of strategies like asset sales and mergers in highly leveraged industries.

KMV Corporations developed the KMV model (Kealhofer 2003) as an extension of the Merton Model for assessing credit risk. Unlike the Merton Model's focus on valuing debt, the KMV Model emphasizes evaluating the entire company's probability of default. It calculates a reliable default risk indicator based on equity and asset characteristics, becoming a standard for measuring default risk (Kealhofer 2003).

2.3 Bankruptcy Models for Default Prediction

In a 2004 study, Marchesini, Perdue, and Bryan assessed four bankruptcy models for predicting default on HY debt. Altman's (1968) and Ohlson's (1980) accounting-based models showed a high overall accuracy rate but performed poorly in forecasting HY default, reaching a maximum of 39%. Cash flow models demonstrated better performance, with accuracy rates up to 61.5% one year before default. However, concerns about distortions led to the reorganization of bond data into subsets, but the models still proved unreliable. The authors introduced a new methodology, multivariate logistic regression, which showed significantly improved accuracy (79.6%, 72.6%, and 68.2%) at one, two, and three years before default, respectively, compared to existing models.

Extensive research has focused on credit risk, default prediction, and bankruptcy models. In conclusion, most studies emphasize on predicting bankruptcy using accounting data. However, successful accounting models for corporate bankruptcy do not translate to effective predictions for default on HY debt (Marchesini, Perdue, and Bryan 2004), which led to extended versions like the KMV Model which provide more robust frameworks for assessing default risk (Kealhofer 2003). Some studies attribute bankruptcy to systematic risk, while others highlight the importance of idiosyncratic risk (Opler and Titman 1994; Asquith, Gertner, and Scharfstein 1994). Shifting the focus to debt default rather than bankruptcy,

research indicates a correlation between the probability of default, the duration of HY bonds (Asquith, Mullins, and Wolff 1989), and higher coupon payments (Lehman and Fridson 1995).

Considering the implications for default and the critiques of existing models, this study seeks to encompass a range of perspectives derived from the literature review. Consequently, the proposed model will integrate both accounting ratios as well as market sentiment data, offering a forward-looking approach to default prediction.

3. Data and Methodology

The subsequent section provides a detailed overview of the data gathering process, along with an explanation of the preparation, cleaning, and structuring of the relevant findings.

3.1 Data Sample

The focus of this research centers on analyzing corporate bonds that were issued between January 1, 2015, and December 31, 2022, and subsequently experienced default within this time frame. This selection criterion ensures that only bonds defaulting after their issuance from January 1, 2015, onward are included in the sample. The chosen period intentionally excludes the financial crisis in 2008 to prevent potential distortions in the analysis, as numerous defaults occurred during that time.

The required data was retrieved from Bloomberg which defines 'default' as follows: Default is as an event that occurs when 'a firm breaches the terms of a debt contract, such as failing to make a coupon payment on time or to satisfy a debt covenant' (Bondioli, et al. 2021, p.6). This kind of default is referred to as a 'technical default'. If a firm fails to rectify the situation within a grace period, it is considered to have 'defaulted' (Bondioli, et al. 2021).

As this thesis focuses on the HY bond market, the next step involved filtering the defaulted bonds according to their credit rating. The leading global credit rating agencies define HY debt as debt with a credit rating below BBB (Standard & Poor 2023), BB (FitchRatings 2023), or Baa3 (Moody's 2023). However, not every company receives a credit rating. S&P Global Ratings indicates that this can be attributed to several factors. Potentially no rating has been requested, the available information is deemed insufficient to establish a rating, or there may be considerations related to policy matters (S&P Global Ratings 2016). Hence, these bonds carry the credit rating designation 'NR' (not rated). This analysis incorporates NR-rated bonds, along with 'fallen angels' - bonds initially issued with an investment-grade (IG) rating that were subsequently downgraded to non-investment grade or HY.

To ensure data consistency, the sample size was reduced to US Corporate Bonds. It is essential to highlight that various maturity types, such as convertible or callable bonds, were included in the dataset. The maturity type can have implications for default prediction, as convertible bonds typically feature a coupon rate lower than that of regular bond issuances (Choudhry, 2004). Although the seniority of a bond significantly influences the recovery rate of defaulted debt (Altman and Kishore, 1996), it is not pertinent to the expected default frequency, which remains constant across different levels of seniority. For this reason, the factor of seniority was not taken into consideration in the data selection process. Furthermore, the industry plays an important role in the default of issued bonds as higher sensitivity to industry factors, measured by industry beta, lead to an increased probability of default (Agrawal and Maheshwari 2019). The selection of industries was restricted to Financial, Industrial, and Utility Electronics for publicly listed companies in the United States.

The chosen dependent and independent variables for the regression, which will be described in more detail in the upcoming subsection, encompassed data gathered from various sources.

Regarding the dependent variable, the historical daily closing prices of all defaulted bonds within the specified time frame were obtained from Bloomberg. Two of the independent variables incorporate market sentiment data, specifically the volatility driven VIX Index and the Spread between HY and IG Bond Indices for the US. The respective data was retrieved from the Chicago Board Options Exchange (CBOE), a major options exchange that calculates the VIX Index, and from FRED, an economic database that issues various time series data, including the yields of the specified Bond Indices. For the remaining independent variables firm-specific data such as financial ratios were retrieved from Compustat, a comprehensive database of financial, statistical, and market information about publicly traded companies.

The final data sample was therefore reduced to 70 defaulted bonds from 21 publicly listed US companies.

3.2 Selection of Dependent and Independent Variables

In accordance with the findings from the literature review, it was concluded that the mere incorporation of accounting data is insufficient. Instead, it is advisable to also integrate forward-looking market data for a more comprehensive analysis. In this context, volatility was identified as a crucial input for predicting default. The VIX Index is used as a proxy for market sentiment, while spreads between HY and IG indexes reflect market risk perception. Financial ratios like Long-term Debt/Invested Capital, Current Ratio, and Cash Conversion Cycle (Days) provide insights into the companies' financial health. The variables were selected based on their relevance to bond default prediction, established through literature review and financial theory. The Long-term Debt/Invested Capital ratio represents leverage, the Current Ratio measures short-term liquidity, and the Cash Conversion Cycle indicates operational efficiency.

In the following a detailed description of the variables, their composition, and the economic rationale behind their selection will be provided.

3.2.1 Dependent Variable

Y_t : Default Index (Average Daily Closing Prices)

As previously described, the dataset ultimately consisted of 70 defaulted bonds for which the daily closing prices were retrieved. A 'Default Index' was then constructed based on the average daily closing prices for all bonds:

$$Default\ Index_{i,t} = \frac{\sum_{i=1}^x P_{i,t}}{x_t}$$

Let $P_{i,t}$ represent the daily closing price of bond i at time t and let x denote the total number of bonds. Due to variations in issuance and default dates among bonds, the quantity of bonds fluctuates over time t , and is not fixed at 70.

Hence, an Index for defaulted bonds between 2015 – 2022 was constructed.

3.2.2 Independent Variables

Market-Sentiment Data

Variable X_1 : Volatility / VIX Index

In response to criticism surrounding the absence of forward-looking market data, the inclusion of volatility was examined. The volatility of an asset is determined by calculating the standard deviation of annualized returns over a specified time. However, this approach relies on historical returns, which lacks a forward-looking perspective. As a result, the third ratio was defined as the 'Volatility Index' (VIX) or 'Fear Index,' a key daily market indicator, offering a measure of overall market expectations for future volatility in the U.S. The VIX

Index computes a consistent '30-day expected volatility of the U.S. stock market, derived from real-time, mid-quote prices of S&P 500® Index (SPXSM) call and put options' (CBOE 2023). After the data was retrieved, it was adjusted to the given time period.

Variable X_2 : Spread of High Yield Index over Investment Grade Index

Understanding market sentiment is essential for investors in HY bonds as it provides insights into the prevailing risk environment, potential price movements, and the overall health and economic outlook of the market. Negative sentiment, which is associated with concerns about economic conditions can increase the likelihood of defaults. To gauge market sentiment, different variables can be incorporated, including Options Volume or Put/Call Ratios. In order to specifically assess sentiment within the HY market, the focus was placed on the yield spread between HY bonds and IG bonds:

$$\text{Spread} = \text{High Yield Index} - \text{Investment Grade Index}$$

A substantial difference in yields between HY and IG bonds suggests that investors are seeking a larger compensation for the added risk inherent in HY securities compared to their safer IG counterparts. This wider spread can reflect heightened risk aversion, indicating investor concerns regarding economic conditions, the creditworthiness of companies, or potential uncertainties in the market. Additionally, an increased spread frequently signifies a perceived rise in the risk of default for HY issuers, prompting investors to seek greater compensation to mitigate the heightened probability of non-payment.

The ICE BofA US High Yield Index is widely recognized and used as a benchmark for the HY market in the United States, as is the ICE BofA US Corporate Index for IG securities. Given that the dataset encompasses only U.S. listed companies, these indices are deemed

appropriate benchmarks for reflecting market sentiment during the given time period. As an example, the spread between HY and investment-grade bonds peaked on 03/23/2020, registering at 6.7%. From an economic perspective, this signifies a heightened perception of risk and apprehensions about economic conditions. Considering the timeframe, this aligns with the onset of the COVID-19 pandemic in March 2020, during which many countries initiated border closures and there was a prevailing sense of heightened uncertainty and fear regarding the state of the global economy.

Financial Ratios

Variable X_3 : Long-term Debt/Invested Capital Average

The Long-term Debt / Invested Capital ratio is an indicator that assesses a company's reliance on debt to finance its business, which in turn sheds light on the company's potential risk of default. A high ratio suggests that the company is significantly leveraged as a substantial portion of its capital structure is financed through debt. This can result in considerable interest and principal repayment obligations. In instances where a company's earnings are insufficient to meet these obligations, the default risk escalates.

Conversely, a low ratio points to a capital structure that is more equity-driven or conservative, with less dependence on debt financing. While this may indicate a lower risk profile, it also might suggest that the company is not capitalizing on the potential advantages of debt financing, such as the tax benefits associated with interest expenses:

$$\text{Long - Term Debt/Invested Capital Average} = \frac{\text{Long - Term Debt}}{\text{Invested Capital Average}}$$

Variable X_4 : Current Ratio Averages

The current ratio is a critical liquidity metric that helps to assess a company's ability to meet its short-term obligations, which is particularly relevant in HY bond markets which are known for their elevated default risk. A higher current ratio suggests better short-term financial stability, potentially indicating a lower immediate risk of default, while bolstering investor confidence:

$$\text{Current Ratio} = \frac{\text{Current Assets}}{\text{Current Liabilities}}$$

Variable X_5 : Cash Conversion Cycle (Days) Averages

The Cash Conversion Cycle (CCC) measures the time taken by a company to convert resource inputs into cash flows. In the context of HY markets, where liquidity and cash flow are paramount, the CCC provides insights into a company's operational efficiency and liquidity risk. A shorter cycle implies a more efficient management of inventory and receivables, and an ability to quickly convert sales into cash, which is crucial for servicing HY debt. Conversely, a longer CCC could signal potential liquidity issues, which might increase the risk of default:

$$\text{Cash Conversion Cycle} = \text{DIO} + \text{DSO} - \text{DPO}$$

where, DIO = Days Inventory Outstanding, DSO = Days Sales Outstanding, DPO = Days Payable Outstanding.

3.3 Methodology

A quantitative research design is employed utilizing secondary data to perform a time-series multiple regression analysis. The aim is to identify and quantify the relationships between

the independent variables and the dependent variable. The dependent variable is the Default Index. The independent variables include the above-described market sentiment indicators, liquidity, and leverage ratios.

The model is expressed as:

$$Y_{i,t} = \beta_0 + \beta_1 X_{1,t-1} + \beta_2 X_{2,t-1} + \beta_3 X_{3,t-1} + \beta_4 X_{4,t-1} + \beta_5 X_{5,t-1} + \varepsilon$$

where:

- Y_i is the Defaulted Bonds Index at time t
- β_0 is the intercept
- $\beta_1 - \beta_5$ are the coefficients of the independent variables
- $X_{1,t-1} - X_{5,t-1}$ as described in the previous subsection at time $t-1$
- ε is the error term.

The regression model allows for a linear relationship between the forecast variable $Y_{i,t}$ and predictor variables $X_{1,t-1} - X_{5,t-1}$. The coefficients $\beta_0 - \beta_5$ denote the intercept and the slope of the line respectively. The intercept β_0 represents the predicted value of $Y_{i,t}$ when $X_{1,t-1} - X_{5,t-1} = 0$. The slopes of $\beta_1 - \beta_5$ represent the average predicted change in $Y_{i,t}$ resulting from a one unit increase in $X_{1,t-1} - X_{5,t-1}$. The independent variables were time-lagged to $t-1$, enabling the incorporation of historical information up to time t for predicting defaults.

To provide insights on the potential enhancement of the model through the inclusion of forward-looking market data, a secondary regression analysis was conducted, wherein the independent variables $X_{1,t-1}$ and $X_{2,t-1}$, which represent forward-looking market sentiment data, were omitted. Consequently, the second analysis solely relies on idiosyncratic data.

Prior to running the regressions, the data was screened for the main assumptions underlying multiple regression models that must be satisfied: (1) linearity, (2) mean of residuals, (3) homoskedasticity, (4) normality of residuals, (5) no autocorrelation, and (6) independence of independent variables. The resulting Figures can be found in the Appendix from [Figure 1-6](#) for the first regression, and [Figure 8-13](#) for the second regression.

4. Results and Discussion

The subsequent chapter describes and analyzes the outcomes of the regression analysis, complemented by an economic interpretation and rationalization. The findings of the regression analysis offer a numerical evaluation of the variables affecting the Average Closing Price of bonds.

4.1 Regression with Market Sentiment Indicators

After running the first regression, the following results ([see Figure. 7](#)) were obtained:

According to the model intercept β_0 , the expected Average Closing Price is **57.5640** when all the independent variables are at zero. After adjusting for other variables, each coefficient shows the predicted change in the Average Closing Price in response to a one-unit change in the corresponding predictor variable.

The regression model's coefficients are as follows:

The average closing price is affected by a number of factors, as shown by the regression results. The Average Closing Price decreases by 0.1566 units for every unit increase in the VIX Index, indicating a negative relationship between bond prices and market volatility. Likewise, a rise of one unit in the Spread between the HY and IG index translates into a significant reduction of 2.3659 units in the Average Closing Price, underscoring the bond market's susceptibility to variations in credit spreads. An increase of one unit in the Long-

term Debt/Invested Capital Average ratio is associated with a decrease of 2.0589 units in the Average Closing Price, indicating a negative perception of higher leverage among investors. Positively, an increase of one unit in the Current Ratio Averages is linked to a 12.6419 unit rise in the Average Closing Price, highlighting the beneficial impact of liquidity on bond prices. Furthermore, the average closing price increases by 0.0914 units for every unit increase in the cash conversion cycle. This indicates how effective working capital management affects bond valuations.

The *R-squared value of 0.485* demonstrates that the model explains a significant proportion of the variability in bond prices, although there is still room for other factors not included in the model. The statistical significance of the coefficients is confirmed by their *low P-values*, which supports the notion that these variables have a substantial effect on the Average Closing Price in the bond market. *Heteroscedasticity Robust Standard Errors (HC0)* are used to improve the robustness of the regression results by guaranteeing that the model's inference is still dependable even when the error terms exhibit non-constant variance.

4.2 Regression without Market Sentiment Indicators

The second regression, omitting the variables *VIX Index Adj Close* and *Spread HY - IG BofA* obtained the following results ([see Figure. 14](#)):

The intercept constant β_0 , is 46.0211, representing the estimated Average Closing Price when all independent variables are zero.

There is a negative correlation between leverage and bond prices; for every unit increase in the Long-term Debt/Invested Capital Average ratio, the Average Closing Price is expected to decrease by about 1.4282 units. On the other hand, an estimated increase of 12.4567 units in the Average Closing Price is linked to a one-unit increase in the Current Ratio Averages, indicating a positive impact of liquidity on bond valuations. The impact of effective working

capital management on bond prices is further demonstrated by the estimate that a one-unit increase in the Cash Conversion Cycle (Days) Averages will result in an increase of roughly 0.1002 units in the Average Closing Price. The *R-squared* value is 0.371, indicating that the model explains 37.1% of the variance in the dependent variable.

The *F-statistic* tests the overall significance of the model, and the *p-value* (*Prob (F-statistic)*) is very close to zero, suggesting that the model is statistically significant. The *p-values* for each coefficient ($P > |z|$) are all close to zero, indicating that each independent variable is statistically significant.

The following conclusions can be drawn from a comparison of the two regression outputs:

The magnitude of the coefficients for the three remaining independent variables did not change when the two market sentiment variables were removed. The second regression's *R-squared* value dropped from 0.485 to 0.371, even though it still shows statistical significance. According to the first regression's output, the independent variables in the first regression model account for about 48.5% of the variance in the dependent variable. In other words, the model accounts for almost half of the variability in the dependent variable. The results of the second regression show that the independent variables in the second model account for about 37.1% of the variance in the dependent variable. In this case, the model explains a lower proportion of the variability compared to the first model. Therefore, the results of the regression analyses demonstrate a noteworthy enhancement in the predictive capacity of the model when incorporating market sentiment indicators. The discernible constancy in the coefficients for the remaining variables across regressions suggests that the financial ratios, in isolation, did not substantially contribute to the observed reduction in explanatory power.

4.3 Discussion

The regression model's coefficients are supported by sound economic reasoning that is derived from both market behaviour and finance theory.

The Effect of Market Sentiment Data on Default Prediction

The regression outcomes underscore the pivotal role of market sentiment indicators in augmenting the model's predictive accuracy for default in HY bond markets. The higher *R-squared* value in the presence of market sentiment indicators suggests that these variables capture pertinent information, potentially associated with forward-looking market expectations, thereby enriching the overall effectiveness of the regression model.

The enhanced predictive power of the market sentiment regression model can be attributed to several economic reasons:

As this thesis has indicated, the HY-IG Yield Spread and the VIX Index both function as *forward-looking* market indicators. Investors' perceptions of risk and market volatility are reflected in the VIX Index. Yield movements may be an indicator of the market's participants' overall expectations. In order to gain understanding of economic expectations, the yield curve's shape and slope are frequently examined. Economic downturns have historically been linked to an inverted yield curve, in which short-term rates are higher than long-term rates. This illustrates how the yield curve can be forward-looking since it shows what the market anticipates will happen to interest rates and the state of the economy in the future. Conversely, financial ratios offer a static analysis of a business's financial status over a given time period or at a given point in time. The expectations and prospects of the market are subject to fluctuations and can be impacted by factors that go beyond the scope of past financial ratios. Ratios are useful for evaluating trends and a company's financial health, but they don't give clear information about what the market expects or will likely do in the future. Investors and

analysts frequently form expectations regarding a company's future performance based on additional information, such as industry trends and economic forecasts. Consequently, the prediction of default is less accurate when financial ratios are the only basis for the analysis. Furthermore, market sentiment captures *risk tolerance and investor behavior* in general. Sentiment data takes behavioral factors into account, recognizing that market players might not always behave rationally, particularly in times of crisis. Investment decisions are heavily influenced by feelings and perceptions. By adding sentiment data, the model can better understand investor behavior, which affects HY bond default risk and valuation. Compared to higher-quality bonds, HY bonds are usually more susceptible to changes in interest rates. Investors may favor more interest rate-resistant assets during times of market volatility due to increased concerns about interest rate movements. In terms of yield spreads, a higher perceived level of risk and uncertainty in the financial markets can account for a high spread economically. Investors interpret a widening spread as a signal to reassess the risk-return profile of HY bonds. It forces a review of the remuneration necessary to hold these riskier assets. A widening of the spread may be a leading indicator or a sign of worsening credit conditions, even though it does not directly cause defaults. Defaults may happen if the financial situation of HY issuers deteriorates and the economy gets worse. A greater spread can be an indication of what the market believes to be a higher default probability. As a result, adding sentiment data to the model aids in capturing changes in investor sentiment and offers insights into how risk perception affects investment choices, particularly in the highly sensitive HY segment. This is why both indicators, the VIX Index and the yield spread may prove as better predictors in contrast to the remaining variables.

Certain financial metrics and ratios may indirectly influence and shape investors' risk appetite, even though financial ratios by themselves are not reliable indicators of investors' risk appetite or behaviour in general. Leverage ratios, like the examined Long-term

Debt/Invested Capital Average, reveal details about the capital structure and financial risk of a long-term debt-financed business. Higher leverage could be seen as riskier, and investors who are less willing to take on risk might be warier of heavily indebted companies. The interest coverage ratio, earnings, and stock price are examples of additional metrics for assessing risk. Financial ratios do, however, only represent company-specific data; they do not, therefore, account for broader macroeconomic trends that could have an impact on several companies or industries. On the other hand, market sentiment data can respond rapidly to events, news, or shifts in the state of the economy. Investors and traders respond quickly to new information, which causes market prices to fluctuate rapidly. In the HY market, this volatility can determine potential bond defaults, distress, or even business failure. Thus, sentiment data can function as a precursory warning system. Abrupt changes in attitude could come before more significant market fluctuations and defaults. With the addition of sentiment indicators, the model may be able to recognise new patterns or warning signs, resulting in more accurate and proactive default forecasts. Financial ratios do not have this ability due to their reporting frequency. Generally, financial statements are published on a quarterly or annual basis. There is a delay in the availability of financial information due to this reporting frequency. Financial ratios therefore might not accurately represent the most recent changes or occurrences. Financial data from the past is compiled and analysed for reporting purposes. There is a time gap caused by the time needed to gather, handle, and examine this data. Furthermore, gathering information and carrying out the necessary calculations take time when calculating financial ratios. This procedure may take longer than expected, particularly if there are complex financial statements or if changes are required for particular analyses.

Intricately tied to investor behavior and risk tolerance factors is *liquidity*. Sentiment data can show shifts in investor flight to quality and variations in market liquidity. More liquid and

less risky assets may be preferred by investors during periods of heightened market stress. By adding sentiment indicators, the model is better able to account for the possible liquidity issues that HY bonds may encounter during times of market turbulence, which may impact default risk. Sector-to-sector variations in liquidity requirements are common. While some industries may rely more on operational liquidity metrics, others might be more sensitive to market sentiment. Generally speaking, a higher current ratio indicates that the business can pay its debts more easily, and investors might see a current ratio that is steady or rising as a sign of effective working capital management. A business that maintains a low current ratio can be seen as effectively allocating its resources, which can boost investor confidence and drive up bond prices. As a result, as indicated by the size of the regression coefficient, adding the current ratio, which accurately depicts a company's short-term liquidity position and ability to meet current obligations, greatly increased the predictive power of the model. This emphasises how crucial it is to take into account company-specific financial health data in addition to market dynamics for a more comprehensive and nuanced analysis.

In the context of financial ratios, the aspect of accounting policies followed by a company must also be considered. The ratios may be impacted by modifications to accounting procedures or adjustments, which may or may not be indicative of shifts in the company's business outlook. For example, there are difficulties when comparing the risk-averse and conservative accounting practices under the German Handelsgesetzbuch (HGB) with the IFRS accounting regulations.

However, although market sentiment data offers insightful information, it lacks specifics regarding a company's management, financial standing, or industry standing—all of which can be critical factors in default prediction. Because firm-specific financial ratios reflect aspects of liquidity, leverage, profitability, and efficiency that market sentiment data cannot, they provide crucial quantitative information for evaluating default risk. As a result, they

provide a thorough understanding of a business's financial situation, assisting analysts, investors, and risk managers in reaching well-informed conclusions regarding the likelihood of default.

The results enhance the existing literature review by offering a distinct viewpoint on the factors to consider when predicting default. Most existing models focus on predicting bankruptcy using financial ratios rather than forward-looking market data, as advocated by Vassalou and Xing (2002). While effective for business failure, these models fall short in predicting default in the HY bond market (Marchesini, Perdue, and Bryan 2004). This thesis presents an alternative framework by combining financial ratios with forward-looking market sentiment data, affirming that the inclusion of market data enhances predictive capabilities for HY bond defaults. Financial ratios, derived from historical accounting data, lack qualitative information and market sentiment. Incorporating sentiment indicators provides a more comprehensive understanding of the intricate interplay between market dynamics, investor behavior, and economic conditions, improving default predictions. Further incorporation of forward-looking data, such as implied volatilities, put/call ratios, or options volume, can enhance the model, potentially making it a reliable indicator for high-yield bond pricing. The efficacy of the model can be evaluated by implementing it in trading scenarios with specific market indicators, assessing its performance frequency and effectiveness in different market situations.

4.4 Scope and Limitations

While contributing valuable insights into default prediction in HY bond markets, this thesis is subject to several limitations that warrant consideration.

One notable constraint is the restricted time frame of the analysis, encompassing data from 01/2015 to 12/2022. A more extensive temporal coverage, including data from a more

extended period, would provide a broader perspective on bond defaults. This limitation may impact the generalizability of the findings to different economic conditions.

Furthermore, the thesis focused exclusively on U.S.-listed companies, narrowing the scope to a specific subset of the global HY bond market. As a result, the universality of the model's predictive power to international contexts or non-U.S. listed companies remains a subject for future research. The analysis was also confined to companies within three specific industries, potentially limiting the model's applicability across diverse sectors. Extending the study to include a more comprehensive array of industries would enhance the model's robustness and broaden its relevance.

The financial ratios incorporated into the regression models constitute only a fraction of the numerous ratios at disposal. Exploring additional financial ratios, capturing various dimensions of financial health, could augment the model's predictive capabilities and offer a more nuanced understanding of default risk.

Lastly, the use of the VIX index, while informative, may not capture the individualized risk profile of each examined company. Implied volatility, which represents the market's expectation of future volatility for a particular security could provide a more precise measure.

In light of these limitations, future research endeavors could expand the temporal, geographical, and sectoral scope, incorporate a more extensive array of financial ratios, and explore alternative market sentiment indicators tailored to individual companies. Such endeavors would contribute to a more comprehensive and nuanced understanding of default prediction dynamics in HY bond markets.

5. Conclusion

The impactful role of market sentiment data in predicting defaults can be attributed to its inherent characteristic of encapsulating forward-looking perspectives within financial markets. Market sentiment indicators, such as the VIX index, serve as proxies for investors' expectations and attitudes towards future market conditions. The VIX, specifically, gauges the anticipated volatility in the market. This forward-looking nature allows market sentiment data to capture timely information about evolving market dynamics and the perceived risk environment.

Similar to the VIX index, the spread between HY and IG indices provides valuable insights into market participants' risk perceptions and expectations. A widening spread often indicates increasing credit concerns and a higher demand for compensation for assuming higher risk. Conversely, a narrowing spread may suggest a more favorable credit environment and reduced concerns about default. The utilization of this spread as a market sentiment indicator contributes to the model's ability to capture shifts in investors' sentiment specific to the credit risk landscape. It complements the VIX index by offering a nuanced perspective on credit differentials between distinct bond classes.

By incorporating both the VIX index and the HY spread, the model benefits from a multifaceted view of market sentiment, encompassing expectations about market volatility (VIX) and the credit risk premium associated with HY bonds relative to IG bonds. This dual-source approach enables the model to more comprehensively evaluate the forward-looking dimensions of market sentiment, thereby strengthening its overall effectiveness in predicting defaults within the HY bond markets.

In contrast, financial ratios, being derived from a company's historical financial statements, primarily offer insights into its past performance. While financial ratios are valuable indicators of a company's financial health and stability, they inherently lack the ability to

incorporate real-time or forward-looking market expectations. As a result, financial ratios may not fully account for shifts in investor sentiment or changing market conditions that could be indicative of potential default risks.

The inclusion of market sentiment data in the regression model thus enriches the predictive capacity by providing a contemporaneous view of market participants' expectations. This enables the model to better discern emerging trends, anticipate shifts in investor sentiment, and incorporate dynamic factors that may influence the likelihood of defaults in HY bond markets. The forward-looking nature of market sentiment data complements the historical context provided by financial ratios, resulting in a more comprehensive and timely assessment of default risk.

Appendix

Figure 1. Regression 1 - Linearity between the Independent and Dependent Variables

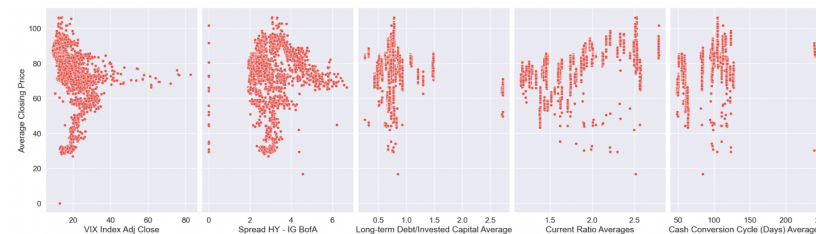


Figure 2. Regression 1 - Mean of Residuals

```

1 residuals = y_train_cleaned.values - y_pred
2 mean_residuals = np.mean(residuals)
3 print("Mean of Residuals {}".format(mean_residuals))

```

Mean of Residuals 1.594617360781743e-15

Figure 3. Regression 1 - Check for Homoscedasticity

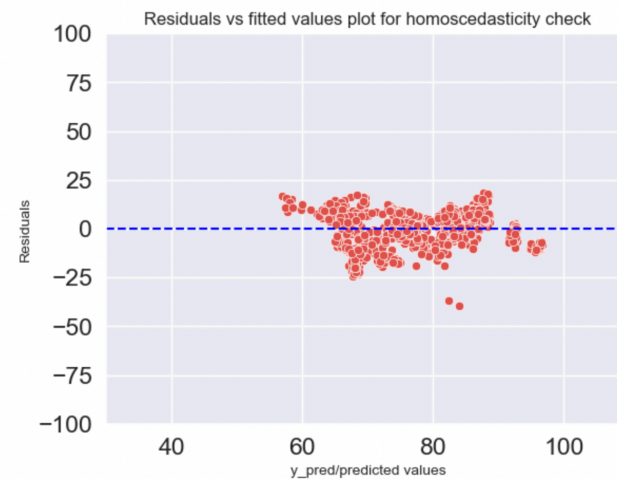


Figure 4. Regression 1 - Normality of Residuals

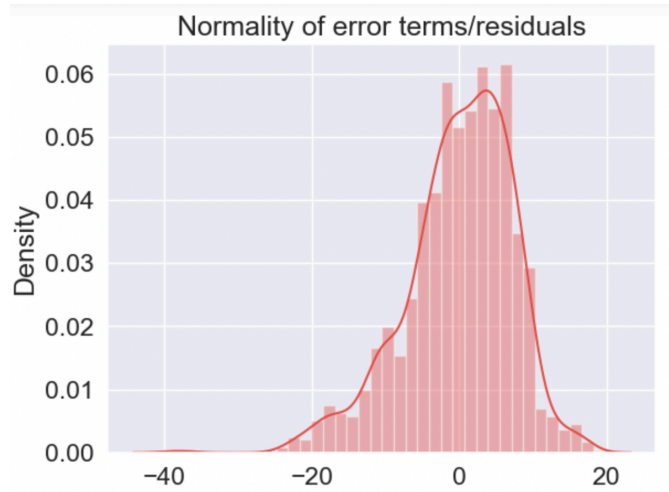


Figure 5. Regression 1 - Autocorrelation for Residuals

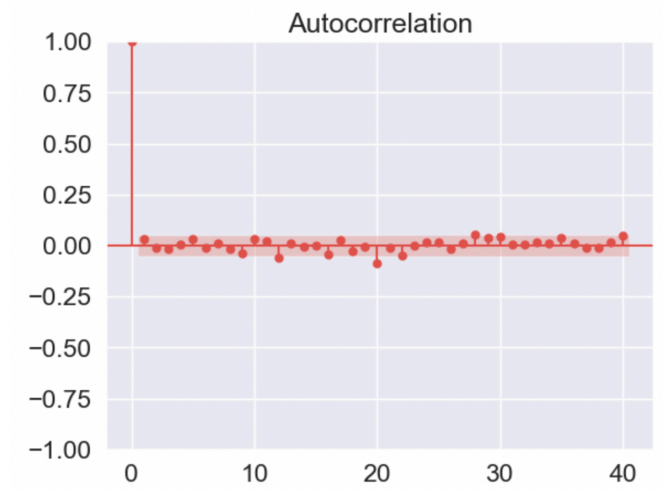


Figure 6. Regression 1 - Correlation of Variables

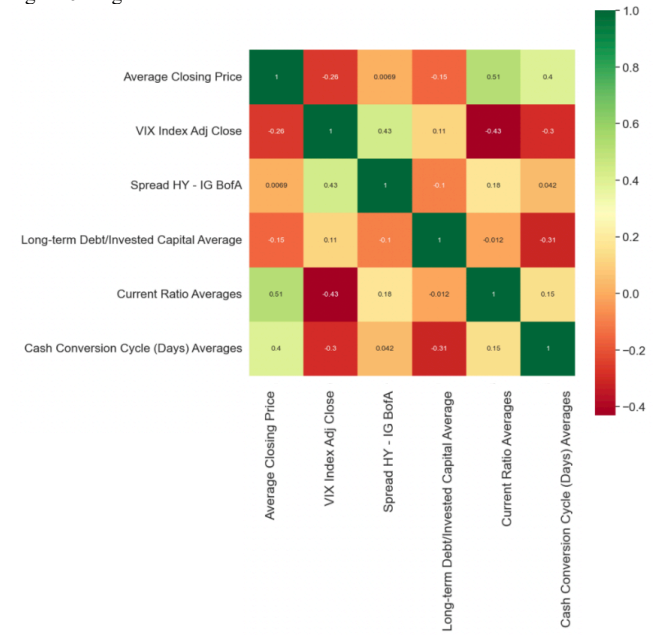


Figure 7. Regression Results with Market Sentiment Variables

Dep. Variable:	Average Closing Price		R-squared:	0.485		
Model:	OLS		Adj. R-squared:	0.484		
Method:	Least Squares		F-statistic:	288.3		
Date:	Sat, 16 Dec 2023		Prob (F-statistic):	1.05e-219		
Time:	17:18:21		Log-Likelihood:	-5598.7		
No. Observations:	1594		AIC:	1.121e+04		
Df Residuals:	1588		BIC:	1.124e+04		
Df Model:	5					
Covariance Type:	HC0					

	coef	std err	z	P> z	[0.025	0.975]
const	57.5640	1.630	35.321	0.000	54.370	60.758
VIX Index Adj Close	-0.1566	0.045	-3.463	0.001	-0.245	-0.068
Spread HY - IG BofA	-2.3659	0.290	-8.159	0.000	-2.934	-1.798
Long-term Debt/Invested Capital Average	-2.0589	0.414	-4.968	0.000	-2.871	-1.247
Current Ratio Averages	12.6419	0.514	24.605	0.000	11.635	13.649
Cash Conversion Cycle (Days) Averages	0.0914	0.008	11.459	0.000	0.076	0.107

Omnibus:	638.168	Durbin-Watson:	0.558
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4734.596
Skew:	-1.688	Prob(JB):	0.00
Kurtosis:	10.739	Cond. No.	825.

Figure 8. Regression 2 - Linearity between the independent and dependent variables

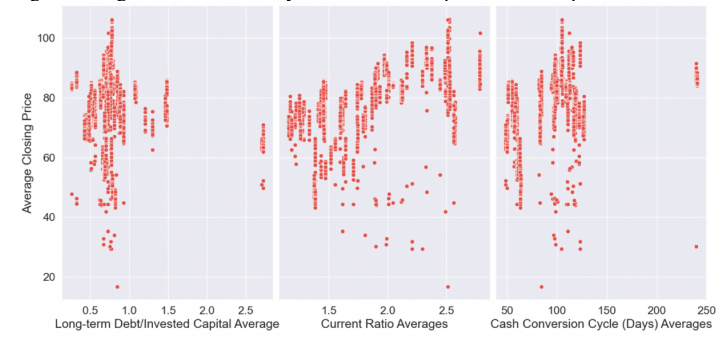


Figure 9. Regression 2 - Mean of Residuals

```
residuals = y_train_cleaned.values - y_pred
mean_residuals = np.mean(residuals)
print("Mean of Residuals {}".format(mean_residuals))
```

Mean of Residuals -1.1390730347026817e-15

Figure 10. Regression 2 - Check for Homoscedasticity

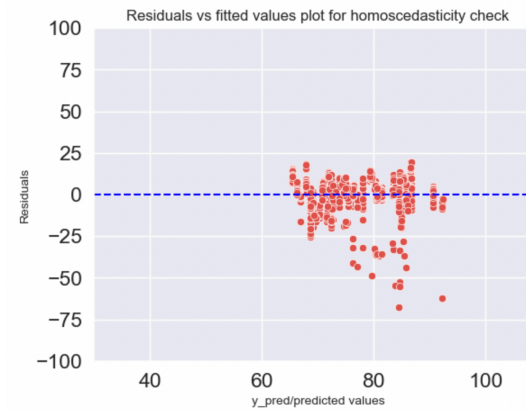


Figure 11. Regression 2 - Normality of Residuals

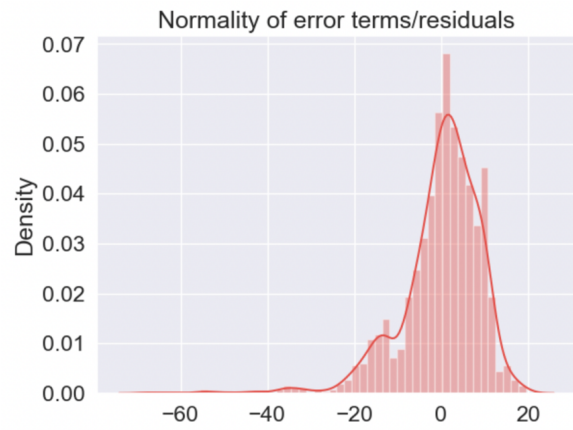


Figure 12. Regression 2 - Autocorrelation for Residuals

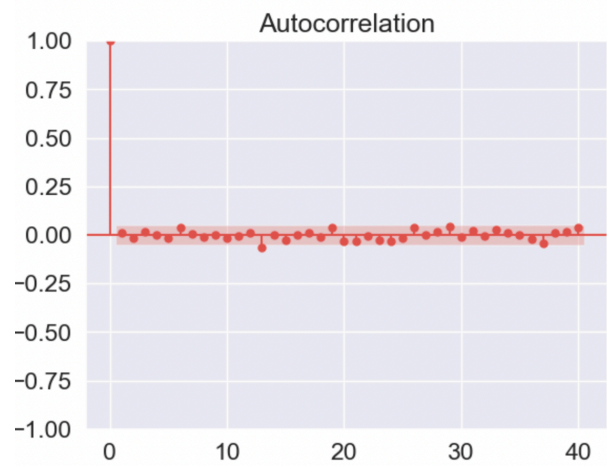


Figure 13. Regression 2 - Correlation of Variables

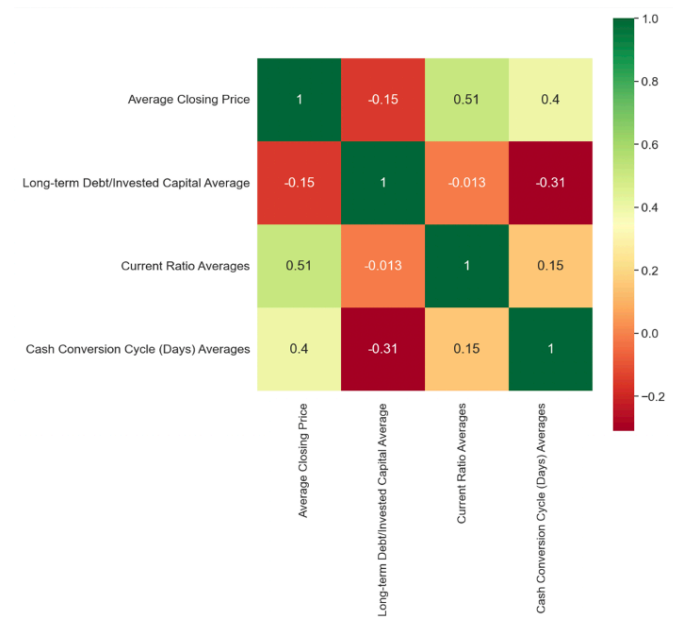


Figure 14. Regression Results without Market Sentiment Variables

DEP. VARIABLE:		Y	R-SQUARED:	0.371
MODEL:	OLS	Adj. R-squared:	0.370	
METHOD:	Least Squares	F-statistic:	397.5	
DATE:	Fri, 08 Dec 2023	Prob (F-statistic):	1.31e-203	
TIME:	18:06:09	Log-Likelihood:	-7519.6	
NO. OBSERVATIONS:	2061	AIC:	1.505e+04	
DF RESIDUALS:	2057	BIC:	1.507e+04	
DF MODEL:	3			
COVARIANCE TYPE:	HCO			

	COEF	STD ERR	Z	P> Z	[0.025	0.975]
CONST	46.0211	1.084	42.467	0.000	43.897	48.145
LONG-TERM DEBT/INVESTED CAPITAL AVERAGE	-1.4282	0.370	-3.856	0.000	-2.154	-0.702
CURRENT RATIO AVERAGES	12.4568	0.445	27.979	0.000	11.584	13.329
CASH CONVERSION CYCLE (DAYS) AVERAGES	0.1002	0.006	16.794	0.000	0.089	0.112

OMNIBUS:	773.486	DURBIN-WATSON:	0.796
PROB(OMNIBUS):	0.000	Jarque-Bera (JB):	4105.717
SKEW:	-1.695	Prob(JB):	0.00
KURTOSIS:	9.027	Cond. No.	603.

Bibliography

- Agrawal, Khushbu, and Yogesh Maheshwari. 2019. "Efficacy of industry factors for corporate default prediction." *IIMB Management Review* 71-77.
- Altman, Edward I. 1968. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *The Journal of Finance* 589-609.
- Altman, Edward I., and Brenda J. Kuehne. 2013. "Defaults and Returns in the High-Yield Bond Market: Third-Quarter 2013 Review." *Journal of Financial Management Markets and Institutions* 269-296.
- Altman, Edward I., and Vellore M. Kishore. 1996. "Almost Everything You Wanted to Know about Recoveries on Defaulted Bonds." *Financial Analysts Journal* 57-64.
- Altman, Edward I., Brooks Brady, Andrea Resti, and Andrea Sironi. 2005. "The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications." *The Journal of Business* (The Journal of Business) 2203-2228.
- Altman, Edward I., Robert G. Haldeman, and P. Narayanan. 1977. "ZETA Analysis: A new model to identify bankruptcy risk of corporations." *Journal of Banking and Finance* 29-54.
- Asquith, Paul, David W. Jr. Mullins, and Eric D. Wolff. 1989. "Original Issue High Yield Bonds: Aging Analyses of Defaults, Exchanges, and Calls." *The Journal of Finance* 923-952.
- Asquith, Paul, Robert Gertner, and David Scharfstein. 1994. "Anatomy of Financial Distress: An Examination of Junk-Bond Issuers." *The Quarterly Journal of Economics* 625-658.
- Beaver, William H. 1966. "Financial Ratios as Predictors of Failure." *Journal of Accounting Research* 71-1111.
- Berk, Jonathan, and Peter DeMarzo. 2016. *Corporate Finance: The Core*. Harlow: Pearson.

- Black, Fischer, and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *The Journal of Political Economy* 637-654.
- Bondioli, Mario, Martin Goldberg, Nan Hu, Chengrui Li, Olfa Maalaoui, and Harvey J. Stein. 2021. *The Bloomberg Corporate Default Risk Model (DRSK) for Public Firms*. The Bloomberg Corporate Default Risk Model, New York: Bloomberg.
- CBOE. 2023. *VIX Volatility Products*. 23 November. Accessed November 23, 2023. https://www.cboe.com/tradable_products/vix/.
- Chaplinsky, Susan J. 2013. *The High-Yield Debt Market*. Virginia, 14 November.
- Choudhry, Moorad. 2004. *Corporate Bonds and Structured Financial Products*. Amsterdam: Elsevier.
- Deakin, Edward B. 1972. "A Discriminant Analysis of Predictors of Business Failure." *Journal of Accounting Research* 167-179.
- FitchRatings. 2023. *FitchRatings: Rating Definitions*. Accessed October 23, 2023. <https://www.fitchratings.com/products/rating-definitions>.
- Hong, Yongmiao, Hai Lin, and Chunchi Wu. 2012. "Are corporate bond market returns predictable?" *Journal of Banking & Finance* 2216–2232.
- Kealhofer, Stephen. 2003. "Quantifying Credit Risk I: Default Prediction." *Financial Analysts Journal* 30-44.
- Lehman, Richard, and Martin S. Fridson. 1995. "Coupon Rates as Indicators of Default Risk." *Extra Credit* 4-8.
- Marchesini, Roberto, Grady Perdue, and Vicki Bryan. 2004. "Applying Bankruptcy Prediction Models to Distressed High Yield Bond Issues." *The Journal of Fixed Income* 50-56.
- Merton, Robert C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *The Journal of Finance* 449-470.
- Moody's. 2023. *Moody's Rating Scale*. Accessed October 23, 2023. https://www.moodys.com/sites/products/productattachments/ap075378_1_1408_ki.pdf.
- Ohlsen, James A. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research* 109-131.
- Opler, Tim C., and Sheridan Titman. 1994. "Financial Distress and Corporate Performance." *The Journal of Finance* 1015-1040.
- S&P Global Ratings. 2016. *S&P Global Ratings Definitions*. Rating Definitions, New York: S&P Global Ratings.
- Standard & Poor. 2023. *S&P Global Ratings: Intro to Credit Ratings*. Accessed October 23, 2023. <https://www.spglobal.com/ratings/en/about/intro-to-credit-ratings>.
- Vassalou, Maria, and Yuhang Xing. 2002. "Default Risk in Equity Returns." *The Journal of Finance* 2-57.