ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business

Corporate Bond Liquidity

Analysis of liquidity effects on corporate bond current yields and yield spreads

Author:	Lars Mackenbach
Student number:	569922
Thesis supervisor:	Dr. R. de Bliek
Second reader:	prof. dr. S. van Bekkum
Finish date:	8 July 2023

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second reader, Erasmus School of Economics or Erasmus University Rotterdam.

ABSTRACT

In this paper, I studied the relationship between liquidity and U.S. corporate bond yields. The bond yields I studied are the current yields as well as the yield spreads between 2010 and 2020. I analysed the relationship by performing multiple regressions applied on multiple panels such as different credit ratings, time to maturities and time windows. I found that the relationship between liquidity and corporate bond current yields is negative, and its magnitude differs between different credit ratings and time to maturities. I also found that the relationship between liquidity and corporate bond yield spreads is not always negative. I found a positive relationship for A-rated U.S. corporate bonds between 2010 and 2020. The magnitude of the relationship differs between different credit ratings and time periods. Previous literature suggested a negative relationship between liquidity and yield spreads, so this paper proves that this negative relationship is not always the case. Following this paper, future research should not always assume a negative relationship between liquidity and corporate bond yield spreads. Investors could make use of the fact that the liquidity premium differs between credit ratings, time to maturities and time periods.

Keywords:
Liquidity
Corporate Bond Spreads
Corporate Bond Yields
Regression Analysis

JEL codes:
C20
C23
G00
G12

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 Introduction

Over the last 10 years the corporate bond market became more and more popular (OECD, 2020). This can be the result of lower interest rates and improved bond quality. Companies issue more debt, and investors are seeking for a higher yield in corporate bonds. In the same period, corporate bond liquidity declined (Adrian et al., 2015). Liquidity is very important for investors since it is associated with risk. When a bond is illiquid, it can be bought and sold less easily on the market. Investors often want to earn a premium for this risk. In this paper we will look at the relationship between the liquidity of corporate bonds and its yield spreads.

Bond liquidity is studied already earlier in time. In 2007, Chen, Lesmond and Wei observed that illiquid bonds are associated with higher yield spreads. When this illiquid bond improved liquidity, their yield spread would go down. In 2015, Helwege, Huang and Wang observed that the liquidity has a significant effect on corporate bond yields. However, they say that there's also a part of the credit risk in this liquidity measure. They also conclude that the risk premia are time varying and are related to systematic risk. The previously mentioned papers only use U.S. bond data from 1995 until 2003 and 2002 until 2010 respectively. Since the popularity and the liquidity of corporate bonds changed in the period after 2010, this paper will provide newer insights on the relationship between liquidity and corporate bond yield spreads.

To provide newer and more recent insights, this paper will use bond data from the period 2010 until 2020. The U.S. bond prices and bond yields will come from Trade Reporting and Compliance Engine database. Other U.S bond data regarding face value, issuer and credit rating comes from the Mergent Fixed Income Securities Database. Both databases are accessible via Wharton. Of course some adjustments will be made to the dataset. I will for example delete defaulted and convertible bonds in order to create an even playing field. When the time to maturity of a bond exceeds 40 years, the bond will be deleted. These actions are in line with the data adjustment methods of previous literature. We will also use interest rates in order to calculate the corporate bond spreads. This data will come from the Investing.com database, where the CBOE Volatility Index (VIX) data will also be from extracted. Lastly, I will also use the FRED database, made available by the Federal Reserve. From here the Consumer Price Index (CPI) data is extracted.

First, I will look at overall liquidity, corporate bond spreads, inflation and volatility over time. In this way we will get a general overview of the economic situation in our period. Afterwards I will extend the database and make some additional variables ourselves. These variables include the age of the bond and the current yield. When all the variables are collected, it is time to perform some regressions. I will perform multiple regressions on the current yield as well as the yield spread. In both regressions

I will use multiple control variables in order to control for differences in bond characteristics as well as controlling for other risks like inflation risk and market risk. I will also divide the data in multiple panels based on credit ratings and time to maturities. I do this because I believe that liquidity differs a lot within these variables.

I expect to find different results compared to previous research, since previous literature suggests that the liquidity premia are time varying (Helwege, Huang and Wang, 2015). A different economic environment in the way of lower interest rates could have an impact on the popularity ant the liquidity of corporate bonds. When I assume a lower liquidity, I expect to find higher yields. My view is that investors need to be compensated for the higher liquidity risk. Especially interesting will be what the signs and magnitudes of the liquidity effect are when we use multiple panels. The outcome of this paper will be very relevant for investors and researchers. They will gain new insights on the liquidity premia on corporate bonds in the beginning of the twenty-first century which can help them make better investment decisions in different economic environments.
The remainder of this paper is structured as follows. Section 2 will discuss relevant literature and previous research. Section 3 will explain the data and will show summary statistics of the relevant variables. Section 4 will discuss the method used in this research. Following the method, the results will be discussed in section 5. Section 5 will also include a comprehensive discussion on this research. Lastly, there we be a conclusion in section 6.

CHAPTER 2 Theoretical Framework

2.1.1 Bond yield spreads

Yield spread is defined as the difference in yield between risky debt and risk-free debt, supposing the same maturity (Balasubramnian & Cyree, 2011). The risky debt will be the corporate bond yield that we are studying. This corporate debt is debt issued by companies and is considered risky because it is facing different risks such as credit and market risks. The risk-free debt is originally defined as a treasury bond with the same maturity as the risky debt (Liu, Ning & Davidson, 2010). Treasury bonds have no default risk and no reinvestment risk, which are clear conditions for a risk-free asset stated by Damordan (1999). In my research I will use U.S. interest rates instead of U.S. government bonds. I use interest rates because I don't have access to all the precise matching time to maturity government bond data that is needed to calculate the original government bond to corporate bond yield spread.

2.1.2 Bond liquidity premium

Since I want to study the effects of corporate bond liquidity, liquidity should also be defined. Liquidity refers to the ability to liquidity a position in a timely manner at a reasonable price (Muranaga & Ohsawa, 1997). When a bond can be bought and sold easily on the market, it is considered as a high liquidity bond. When selling or buying becomes more difficult due to for example a market sell-off, the bond will become more illiquid. This means that bond liquidity is dynamic and differs for every kind of bond. The risk of not being able to buy or sell the bond is referred to as the liquidity risk. Investors often want a premium on riskier bonds for bearing this liquidity risk. These liquidity premia are proven to change over time, as studied by Helwege, Huang and Wang (2015).

2.2 Empirical studies on the bond liquidity premium

The relationship between bond yield spreads and bond liquidity is studied already in 2007 by Chen, David, Lesmond and Wei. They started studying this relationship because previous literature did not explain the liquidity premium in a correct manor. A lot of early studies break down the corporate bond spreads into different elements, and just assume that the unexplained part is the liquidity premium. An example is the paper from Duffee (1999). He found that there is a liquidity premium present in the bond yield spreads, while they had no liquidity proxies or measures incorporated in their model.
Other studies just assume aggregate liquidity proxies, such as the paper from Campbell and Taksler (2003). They do incorporate liquidity in their model, however they only look at the difference between the 30-day Eurodollar rate and the treasury yields. This proxy only describes the demand for liquidity and not the implied liquidity.
The study that probably comes closes to the study of Chen, David, Lesmond and Wei is the study from Longstaff, Mithal and Neis (2005). They first divide the spread into a default and a non-default premia

part. Afterwards they conclude that a large part of the non-default part can be explained by multiple liquidity proxies. Proxies they use are the average bid-ask spread, amount outstanding, age, time to maturity, a dummy for financial firms and a dummy for credit rating. A warning is the fact that they only study a limited number of bonds which all have liquid default swap trading data. They do however find a negative relationship between liquidity and the corporate bond yields. Which means that when a bond becomes more liquid, its liquidity premia decreases and consequently the yield decreases. One might ask if the liquidity effects will be the same for a wider variety of corporate bonds. The paper of Chen, David, Lesmond and Wei (2007) studies the relationship between corporate bond yield spreads and liquidity using three liquidity proxies. They regress these proxies on the yield spreads of more than 4000 different corporate bonds. By studying such a large amount of bonds, they try to improve previous literature. The study of Ericsson and Renault (2006) studies liquidity spreads. The liquidity spread is the difference in yield between bonds with the same liquidity. They found a decreasing and convex term structure for the liquidity spreads. This means that the differences in liquidity premia decreases when the yield to maturity rises.

The study from Houweling, Metlink and Vorst (2005) also compares multiple proxies of corporate bond liquidity. This study is especially useful since we will also use proxies for corporate bond liquidity in our research. The proxies that are tested and compared are Issued amount, listed, euro, on-the-run, age, missing prices, yield volatility, number of contributors and yield dispersion. They tested the liquidity proxies without a regression but with a portfolio-based testing method, which was a new way compared to previous literature. The study contained 1190 European corporate bonds from the period 1999-2001. A popular proxy is the age of the bond. Sarig and Warga (1989) already observed that the older a bond gets, the less liquid it becomes. This results in a higher yield since investors want a higher liquidity premium. The issued amount also reflects liquidity. This statement is supported by both Howeling, Metlink and Vorst (2005) as by Fisher (1959). Both of their reasoning was that the higher the issued amount would be, the more trading would take place, and the more liquid a bond would be. Howeling, Metlink and Vorst (2005) conclude that all their proxies except “listed” where significant and yielded a liquidity premium. These significant observations match the results of earlier studies that used a regression to test the proxies.

Another study that relates to our research is the study from Ang, Bhansali and Xing (2014). They analyze the municipal bond spreads and conclude that the municipal bond spread can be divided into liquidity, credit and tax components. The liquidity component has an average significant positive effect of 2.14% on the spread after 2008. Important to mention is that municipal bonds are in general slightly less risky compared to corporate bonds. Although municipal bonds are not the same as corporate bonds, it’s important to observe that there’s a significant liquidity effect present on the spread.

All previous literature found that liquidity effects the bond yields and thus is part of the bond yield spread. Most of them also found a negative relationship between these two factors. Which means a lower liquidity implies a higher yield because of a liquidity premium. However, the way they come to their conclusions differ. Most papers decompose the whole spread into multiple elements. I try to do it differently. I only want to extract the liquidity part, keeping all the other elements such as a default or tax element fixed. Like most researchers I will use U.S. corporate bonds. However, I will investigate the yield spreads from 2010 until 2020. This makes my research also different in comparison to previous research. In my period a lot happened economically. The interest rates as well as the bond liquidity showed a decreasing pattern. Also, corporate bonds became more and more popular. I will also divide the bonds into different panels based on credit ratings and time to maturities.
Lovo, S., Raimbourg, P., & Salvadè, F. (2022) show in their research that when a bond gets an up- or downgrade by a rating agency, the liquidity also changes because of information asymmetries. I believe that the liquidity effect on bond yields indeed differ between different credit ratings.

However, like previous literature, I still expect a negative relationship between liquidity and bond yield spreads. Investors want to earn a premium for facing lower liquidity, driving up the yield spread. In comparison to the paper of Ang, Bhansali & Xing (2014) I think our liquidity effect will be even larger. Municipal bonds are generally more liquid and thus less risky compared to corporate bonds. This will probably result in a larger liquidity premium. My research studies four hypothesis. Hypotheses one and two have a focus on current yields, where hypotheses three and four have a focus on yield spreads.

Hypothesis 1: *There is a negative relationship between liquidity and corporate bond current yields in the period 2010-2020 for U.S. corporate bonds*.
Hypothesis 2: *The magnitude of the relationship between liquidity and corporate bond current yields in the period 2010-2020 for U.S. corporate bonds differs across multiple credit ratings and time to maturities.*

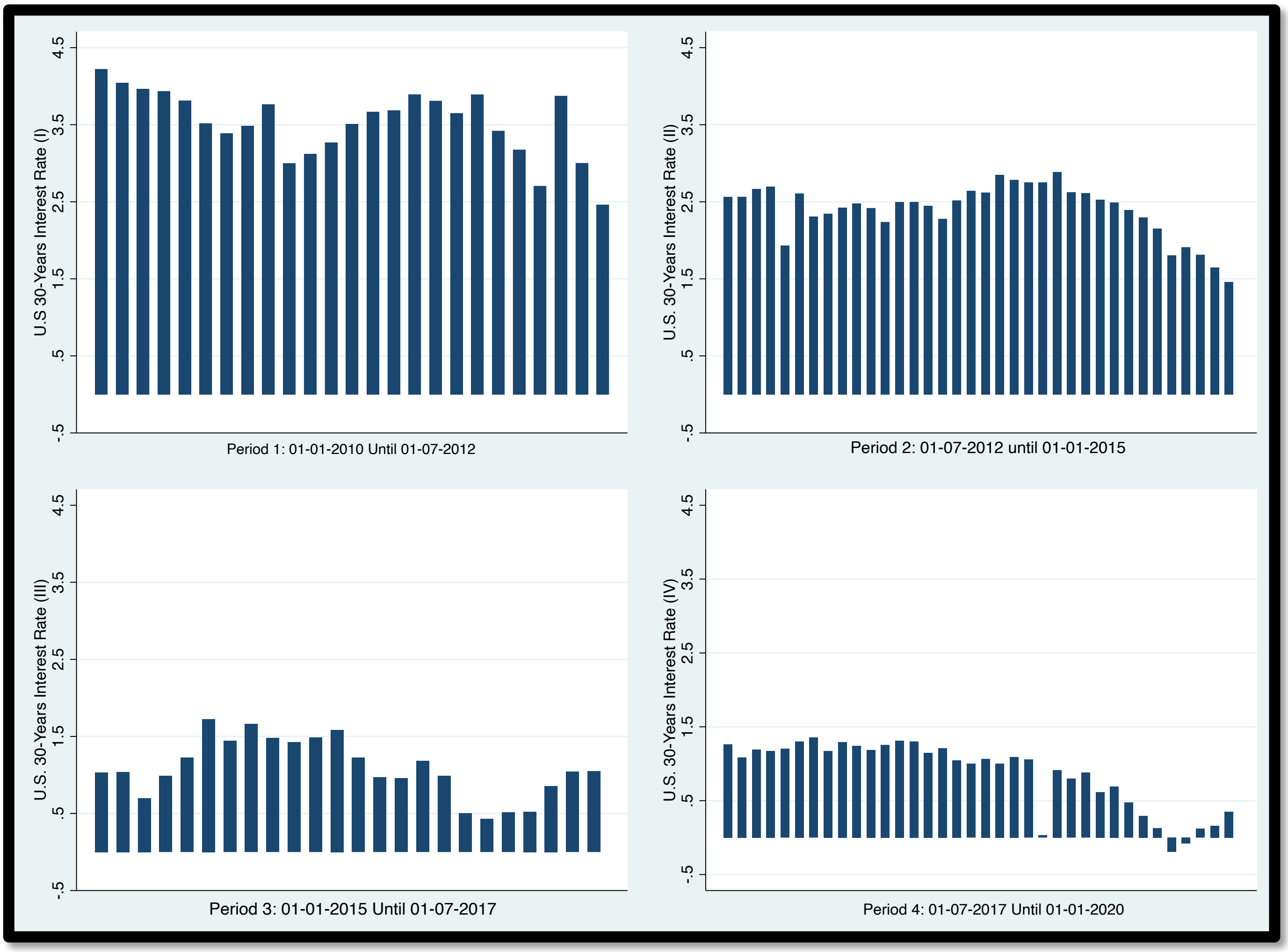
Hypotheses 3: *There is a negative relationship between liquidity and corporate bond yield spreads in the period 2010-2020 for U.S. corporate bonds.*
Hypothesis 4: *The magnitude of the relationship between liquidity and corporate bond yield spreads in the period 2010-2020 for U.S. corporate bonds differs across multiple subperiods and credit ratings.*

CHAPTER 3 Data

For this study I choose to investigate corporate bonds issued in the United States. Most of the previous literature also studies United States bonds, so an advantage is that we can compare our conclusions to previous literature in a better way. Another reason for choosing the United Stated is the dynamic economic situation between 2010 and 2020. In this period liquidity on the corporate bond market showed a decreasing trend in the United States. (Adrian, Fleming, Shachar, Stackman, & Vogt , 2015.) The U.S. 30-years interest rates also showed a decreasing trend. This is visible in figure 1. The combination of both economic trends makes the United States an interesting field to study.

Figure 1: U.S. 30-year Interest Rates

The period 2010-2020 is divided in 4 panels. Each bar represents the interest rate at the end of the month. The interest rates are shown in percentages. The minimum interest rate is -0.194% and the maximum interest rate is 4.223%. The mean interest rate over all 4 the periods is 1.800%. The monthly 30-years interest data is extracted from the investing.com database. A clear downtrend in the 30-year interest rate is visible. We can even see two months with a negative 30-year interest rate at the end of 2019.



I also choose to use United States data because there's a lot of high quality and trustworthy data available. All my bond data comes from the WRDS Bond Returns database. This is a combination of FINRA's TRACE database for bond transactions and Mergents FISD database for bond characteristics. I choose to analyze data from January first 2010 until January first 2020. I limited the data to ten years in order to make the data workable. Also, a lot of different economic situations regarding inflation and interest rates occurred in our period, which makes it especially interesting. Without deleting any variables or missing values my database contains 1,172,326 observations, but some modifications need to be made. All the data in my database is monthly. I choose for this because I want to limit the amount of datapoints in some extend, the dataset also needs to be workable. Although I use monthly data instead of daily data, I am confident that we have enough observations.

First of all, we have the variable $Convertible$, which is a dummy variable for convertible bonds. When a bond is convertible it gets the value 1 and subsequently all these convertible bonds are deleted. This is done in order to create an even level playing field among the corporate bonds. The variable $Amount\ Outstanding$ states the amount of the bond that is currently in circulation. This variable does not directly refer to the liquidity of a bond, but it does indirectly. This is also in line with the conclusion from Houweling, Metlink and Vorst (2005), who stated that the amount outstanding gives an indirect indication of liquidity. Since the values of this variable are very large, I choose to divide the amount outstanding by 100.000.

In the dataset there are three different credit ratings which are all from different rating agencies. These ratings are very important because it refers to the credit risk of a corporate bond, for which we need to control for in our regressions. The three types of credit rating agencies in our dataset are S&P, Moody's and Fitch. When analyzing the dataset, I conclude that Moody's credit rating has the least missing values for our observations. Therefore I will use Moody's credit rating as the default proxy variable for credit risk. All the observations with missing values for this credit rating are deleted. This variable is called $Credit\ Rating$. It has integer values ranging from 1 up until 21. These values match with the well-known alphabetical credit ratings ranging from AAA until C. In appendix A, a conversion table for transforming the numerical rating to an alphabetical rating can be found. Next there is the variable $Bid - Ask\ Spread$ which gives the average bid-ask spread of the corporate bond at the month of the observation. The bid-ask spread is the difference between the bid price of the bond and the ask price of the bond. The unit of the bid-ask spread is in dollars. This spread can be seen as cost for both the buyer and the seller. The bid-ask spread is closely related to the liquidity of the corporate bond, which we want to study. A narrower bid ask spread refers to a more liquid market where transaction prices are closer to each other, which is more efficient. Observations which have a missing value for the bid-ask spread are also deleted, since this is essential for measuring the liquidity.

The bid-ask spread also has some very high outliers. All the observations above 0.05 are deleted, which are only 754 observations. In this way the $Bid - Ask\ Spread$ follows a normal distribution.

Figure 2: Average Bid-Ask-Spread per Credit Rating.

Averages per credit rating are calculated over all observations. The credit rating refers to the numerical credit rating of Moody's. Appendix A includes a conversion table to alphabetical ratings. In figure 2 we can clearly see an upward trend, which indicates that liquidity decreases when a bonds credit rating goes down.

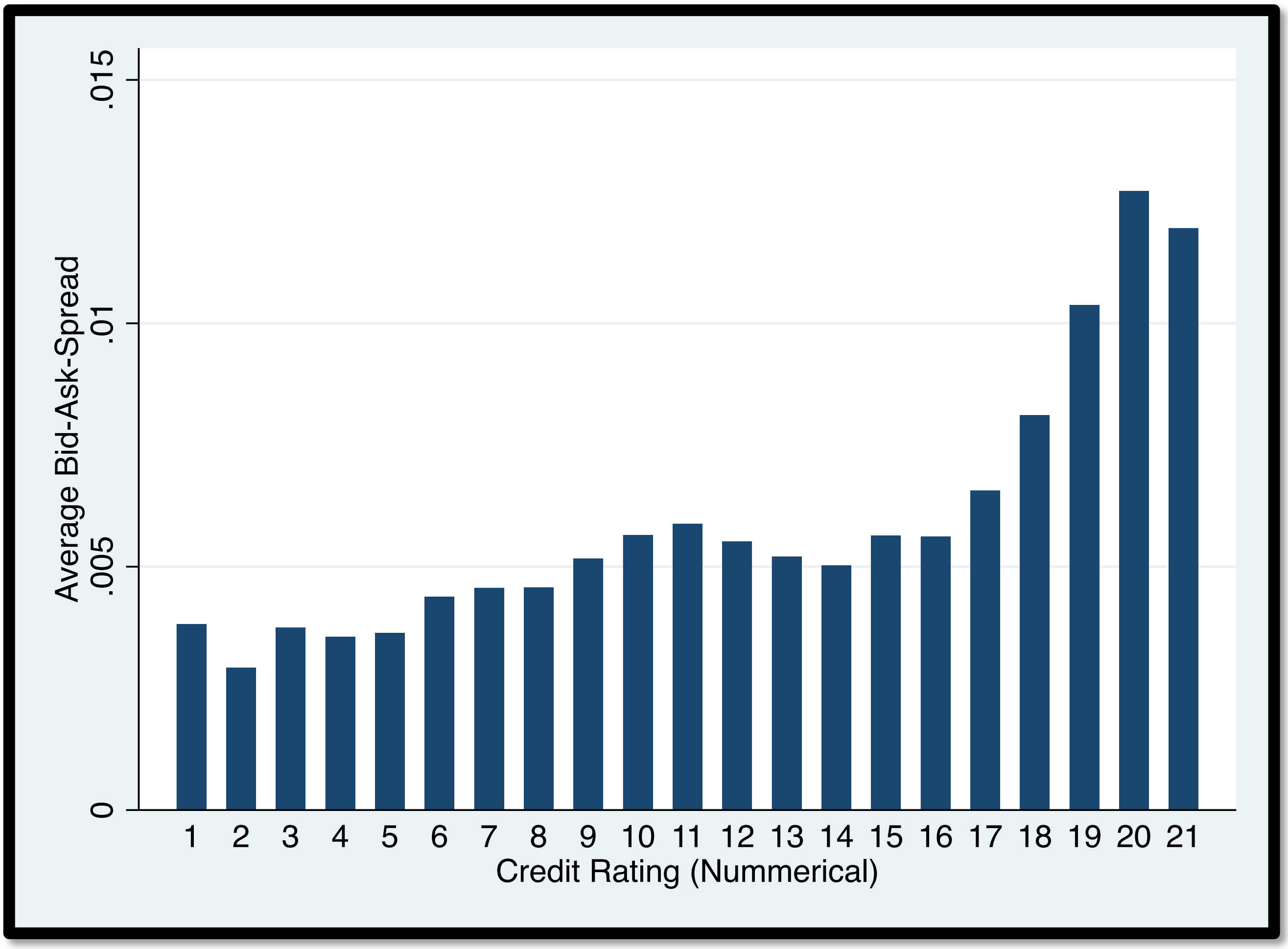


Figure 3: Average Bid-Ask-Spread over Time.

Averages of the bid-ask spread are calculated per observation date, in this way a visual representation of liquidity over time is visible. Every bar represents one month. The average bid-ask spread is given in U.S. dollars ($). A downward trend is clearly visible. The mean bid-ask spread over the full period of 2010 until 2020 is $ 0.0049.

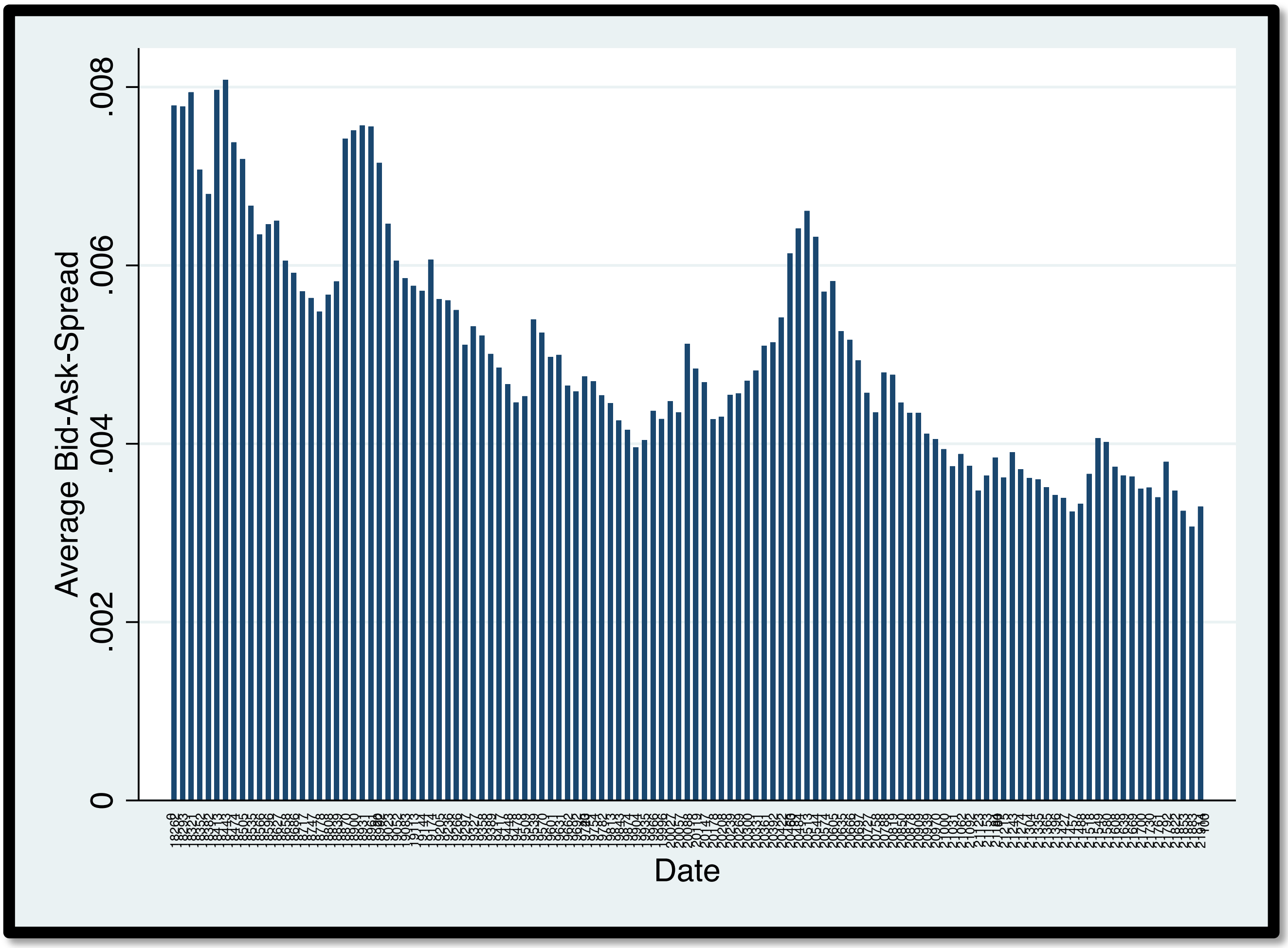


Table 2 shows the average bid-ask-spread per credit rating. One can observe that the bid-ask spread increases, and thus the liquidity decreases on average when bonds become less credit worthy. Table 3 shows the average bid-ask-spread over time. When we look at the average bid-ask spread over time we can see that it has a downward trend, which implies the bond market became more liquid over the period 2010-2020. This is contrary to the findings of Adrian et al. (2015), who looked at a slightly different period.

Next to the bid-ask spread we also have the par value volume, this variable is called $Tot\ Volume$. This variable is calculated by multiplying the par value of each bond by the number of bonds issued. This gives an idea of the size of the bond issuance. I will transform this value using a log function, which will be called $Log\ Volume$. In this way the variable follows a normal distribution. Longstaff, Mithal and Neis (2005) used an about equal proxy for liquidity, which gave them significant results. Following this paper, I will also incorporate the $Log\ Volume$ in the regressions later on.

The variable YTM states the yield to maturity of the bond. This is the total return that the investor earns when holding the bond until maturity. I delete observations that have a missing value for the yield to maturity. We also have the variable $Current\ Yield$. This variable is not given automatically in the database, but generated using a relatively easy formula:

$Current\ Yield = \frac{Coupon}{Price\ (end\ of month)}$.

The current yield states how much return an investor would receive based on the current market price. When analyzing the current yield, we observe some outliers. I choose to delete every observation that has a current yield above 0.2 (20%), which are only 1334 observations. In this way the current yield follows a normal distribution. It also seems quite high to have a current yield above 20% for a corporate bond.

Next there is the variable TMT, which is the time to maturity of the bond. The time to maturity is the time in years until the bond matures. When analyzing the dataset, we conclude that 99,6% of the bonds have a time to maturity below 40 years. The other 0.4% are way higher than 40 years. In order to create an even level playing field we delete all the bonds that have a time to maturity higher then 40 years. I expect that liquidity will differ a lot between bond with different time to maturities. This expectation is bases on the findings from Sarig and Warga (1989).

Another helpful variable is the $Duration$. The duration gives the change in a bond's price when the interest rate rises with 1%, so it measures the sensitivity of the bond towards interest rate fluctuations. This is a very useful variable as a proxy in order to control for the interest rate risk. I also use the variable $Coupon$, which gives the yearly coupon of the bond in percentage. In previous literature the coupon is a proxy for liquidity and tax. The sign of this relationship with the yield of a bond seems ambiguous. Since we also want to control for some of the tax effect, I will also include the coupon as a control variable. I also calculated the Age of the bond. I did this by subtracting the offering date from the date of the observation, and divide that by 365 days. Using this proxy is in line with previous literature from Houweling, Metlink and Vorst (2005). Finally, defaulted bonds are deleted since they do not make interest or principal payments anymore and can't be seen as the standard corporate bond that I want to analyze. Defaulted bonds have the value 1 for the variable $Defaulted$.

Next to the variables from the TRACE and FISD database I also added some extra variables. First I added the CPI variable. This is the Consumer Price Index, extracted from the FRED (Federal Reserve Economic Database). This CPI is a price index for which included all prices. I use this variable as a proxy for inflation. I deliberately choose to use the all-items CPI and not the core CPI since I want an overall impression of the inflation. When I observe the CPI trough time, an increasing pattern can be found between 2010 and 2020. I also added the VIX variable. This is the CBOE Volatility Index, which gives the expected volatility on the S&P 500. This gives an impression of the expected volatility in the market. I extracted this proxy from the investing.com database and it will be used as a proxy for market volatility.

In order to calculate the spread between the yield of a corporate bond and the interest rates, I made the variable $Yield\ Spread$. First every bond is matched with either the 10-year or the 30-year interest rate, based on the total duration of the bond. I extracted the U.S. 10-year and U.S. 30-year interest rates from the investing.com database. The spread is then calculated by subtracting the 30-year or 10-year interest rate from the current yield of the bond. Bonds with a TMT lower than 20 years are matched with the 10-year interest rate. Bonds with a TMT higher than 20 years are matched with the 30-year interest rate, Normally the spread is calculated towards an equal like treasury bond, but since that will be comprehensive for this study, I will use the spread towards the interest rate. So to conclude, this spread gives the excess return of the bond in comparison to an about even duration interest rates.

After modifying the dataset, there are still 784,708 observations left. These are observations from 19,208 unique corporate bonds issued by 2,603 unique companies. The number of observations in this research is relatively high compared to previous research. The reason is that my research includes a relatively long period and a high amount of bonds. Chen, Lesmond and Wei (2007) analyzed for example 4000 bonds for a period of 8 years, which yields considerably less observations.

Table 1: Summary Statistics

All variables are rounded to 2 decimals, except the average bid-ask spread variable. I choose to round this variable to 4 decimals in order to give a better representation of this important proxy variable. Panel 1 only includes bonds with credit rating A, panel B includes bonds with credit rating B and panel C included bonds with credit rating C.

	Mean	Sd.	Min	Max
Full Sample (n=784,708)				
Coupon	5.15	2.11	0.00	14.00
Amount Outstanding (100.000)	6.83	6.53	0.00	150
Age	5.27	5.28	0.00	89.90
TMT (Years)	8.81	8.40	0.00	40.00
Duration	5.94	4.34	0.00	30.00
Log Volume	16.32	2.10	6.06	24.50
Bid/Ask Spread	0.0049	0.0049	0.0000	0.0500
YTM	0.04	0.04	-1.00	1.00
Current Yield	0.05	0.02	0.00	0.20
Yield Spread	0.03	0.02	-0.04	0.20
Panel 1: Credit Rating A (n=295,658)				
Coupon	4.21	1.97	0.00	12.38
Amount Outstanding (100.000)	81.54	76.41	0.00	114.07
Age	5.24	5.38	0.00	82.64
TMT (Years)	9.34	9.21	0.00	40.00
Duration	6.34	4.96	0.00	30.00
Log Volume	16.22	2.24	6.06	23.07
Bid-Ask Spread	0.0041	0.0044	0.0000	0.05
YTM	0.03	0.05	-1.00	1.00
Current Yield	0.04	0.02	0.00	0.15
Yield Spread	0.02	0.02	-0.04	0.11
Panel 2: Credit Rating B (n=442,015)				
Coupon	5.48	1.83	0.00	14.00
Amount Outstanding (100.000)	61.46	56.76	0.00	150
Age	5.32	5.19	0.00	89.90
TMT (Years)	8.81	8.09	0.00	40.00
Duration	5.90	4.00	0.00	30.00
Log Volume	16.34	2.01	6.22	24.50
Bid-Ask Spread	0.0052	0.0049	0.0000	0.05
YTM	0.04	0.02	-1.00	1.00
Current Yield	0.05	0.02	0.00	0.20
Yield Spread	0.03	0.02	-0.04	0.18
Panel 3: Credit Rating C (n= 47,035)				
Coupon	8.01	1.82	0.00	14.00
Amount Outstanding (100.000)	49.57	48.64	0.00	67.50
Age	4.96	5.46	0.00	88.73
TMT (Years)	5.48	4.06	0.00	37.67
Duration	3.83	1.94	0.00	15.05
Log Volume	16.86	1.86	7.60	22.82
Bid-Ask Spread	0.0069	0.0064	0.0000	0.0499
YTM	0.10	0.07	-0.86	1.00
Current Yield	0.09	0.02	0.00	0.20
Yield Spread	0.06	0.03	-0.03	0.20

Table 1 shows summary statistics of the full sample as well as the summary statistics from three panels based on credit ratings. An important observation is that the Age of the bonds has some high values. Only 534 observations are higher than 40 years, where 786,201 observations are lower or equal than 40 years. Despite this, I don't see a reason to delete the higher age observations.
Another remarkable observation is that the number of observations in panel C is a lot smaller compared to the observations in panel A and B. This means that the sample contains relatively less bonds with credit rating C.
What can also be observed is that the maximum value of the $Yield\ Spread$ increases when we go from panel A to panel C and thus go down in credit rating. Based on the maximum values, a higher spread can be potentially achieved by taking on bonds with a weaker credit rating.
We can also see the mean of the $Coupon$ rise from panel A to panel C. This seems logical since investors want a higher compensation for the higher risk associated with a lower credit rating.

CHAPTER 4 Method

After modifying the dataset, the analysis can start. First of all we will analyze some variables trough time in order to get a general overview of the economic situation in our timespan. I will first make new variables which give the monthly averages of $Current\ Yield$ and the $Bid - Ask\ Spread$. They are called $Average\ Current\ Yield$ and $Average\ Bid - Ask\ Spread$. We do not need these variables in the regressions, they are only useful for making graphs that show the monthly averages over time. After making the variables I will make multiple charts in order to create a visual overview of the variables.

For analyzing the effect of liquidity on corporate bond yields I will use STATA. With the use of this statistical software, I will perform multiple regressions.

$$Current\ Yield = \beta_0 + \beta_1 * Bid - Ask\ Spread_i + \beta_2 * \boldsymbol{Control\ Variables_i} + \varepsilon_i$$

This will be the starting point. I will use the $Current\ Yield$ as the dependent variable because I first want to see the effect on the "normal" yield so without the comparison to the interest rates. I want to measure the effect of a change in liquidity on the current yield, so I will use $Bid - Ask\ Spread$ as my proxy for liquidity. First of all I will control for some bond differences, this creates an even level playing field among the bonds. I will use control variables for differences in bonds by using $Coupon$, $Amount\ Outstanding$, TMT, $Log\ Volume$, Age and $Duration$. Afterwards I will control for the inflation and market risks by using the CPI and the VIX proxies respectively. When using all the control variables I formulate to the following equation:

$$\begin{aligned} CurrentYield = {} & \beta_0 + \beta_1 * \ Bid - ASk\ Spread_i + \beta_2 * Coupon_i + \beta_3 * Amount\ Outstanding_i \\ & + \beta_4 * TMT_i + \beta_5 * Log\ Volume_i + \beta_6 * Age_i + \beta_7 * CPI_i + \beta_8 * VIX_i + \varepsilon_i \end{aligned}$$

The next thing that needs to be done is checking for collinearity and multicollinearity. This is done by making a correlation table. I than formulate the following regression to estimate the effect of a liquidity change on the current yield:

$$\begin{aligned} CurrentYield = {} & \beta_0 + \beta_1 * \ Bid - ASk\ Spread_i + \beta_2 * Coupon_i + \beta_3 * Amount\ Outstanding_i \\ & + \beta_4 * TMT_i + \beta_5 * LogVolume_i + \beta_6 * Age_i + \beta_7 * Duration_i + \beta_8 * CPI_i \\ & + \beta_9 * VIX_i + \varepsilon_i \end{aligned}$$

The last thing that needs to be done is controlling for the credit risk. This is very important, since liquidity can differ very much between credit ratings. I will control for the credit risk by creating 3 panels. This method is in line with previous literature such as the paper from Helwege, Huang and Wang (2014). Panel A only includes observations with a $Credit\ Rating$ lower or equal then 7, this corresponds to type A credit ratings. Panel B only includes observations with a $Credit\ Rating$ higher then 7 and lower or equal then 16, this corresponds to type B credit ratings. Panel C only includes observations with a $Credit\ Rating$ higher than 16, which corresponds to type C credit ratings.

Afterwards, every credit rating panel will be divided in three categories based on the time to maturity of the bond at observation. Every credit panel will be divided in parts with time to maturities (TMT) of 0 up until 7 years, 8 up until 15 years and time to maturities larger than 15 years. This is done because I have the expectation that liquidity also plays a big role across bonds with different time to maturities. This last step gives thus a total of 9 different regressions.

Finally, we will look at the effect of liquidity on the $Yield\ Spread$. As mentioned earlier, this is the difference between the current yield and the corresponding interest rate. The regression will be in line with the research from Chen, David, Lesmond and Wei (2007). The main differences are the proxies that are used. I will also correct the standard errors for time series, firm fixed effects and heteroscedasticity. This is in line with the regression method of Helwege et al. (2014).

$$Yield\ Spread = \beta_0 + \beta_1 * \ Bid - ASk\ Spread_i + \beta_2 * Coupon_i + \beta_3 * Amount\ Outstanding_i$$
$$+ \beta_4 * TMT_i + \beta_5 * Age_i + \beta_6 * CPI_i + \beta_7 * VIX_i + \beta_8 * Log\ Volume\ \varepsilon_i$$

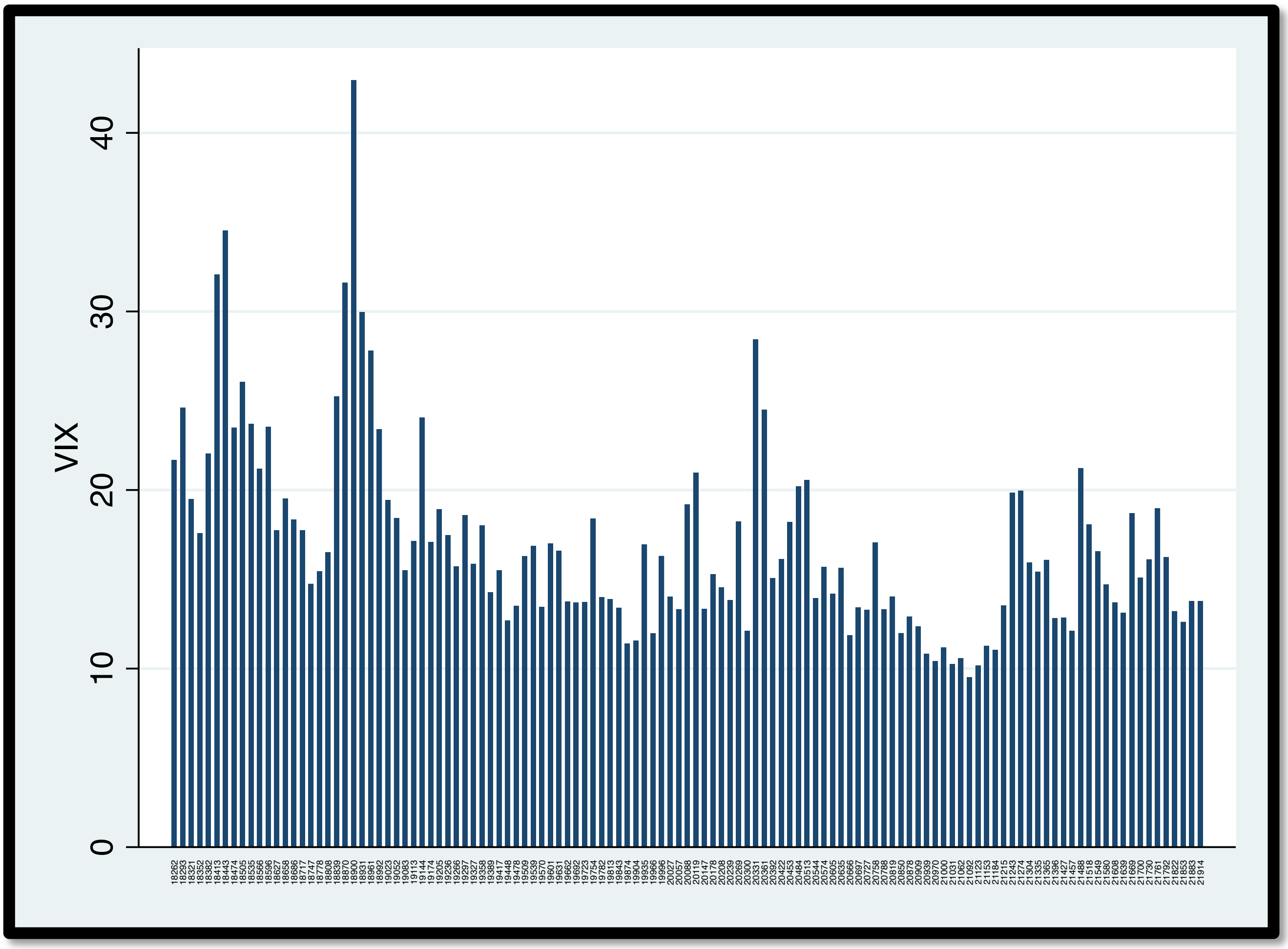
This regression will be performed on different models. We will use models that differentiate time and that differentiate credit ratings. To differentiate in time, I will make 2 panels. One which only includes observations before 01-01-2015, and one that only includes observations after 01-01-2015. To differentiate in credit ratings, I will use the same method as used earlier in this paper.

CHAPTER 5 Results & Discussion

The first step is to create a visual overview of the VIX, Average Bid-Ask Spread and the Average Current Yield over time.

Figure 4: VIX over Time.

The VIX is the expected S&P 500 market volatility. The applied period is from 2010 until 2020. Every bar represents one month. The mean of the VIX over the whole period is 16.676. We can see an unstable and slightly downward sloping trend, which indicates that the expected volatility is volatile itself.



In figure 4 we can see that the VIX fluctuates a lot in the period 2010-2020. This means that the expected volatility of the S&P 500 differs a lot depending on the date in our time frame. We will use the VIX as a proxy for market volatility.

Figure 2 already showed the $Average\ Bid - Ask\ Spread$ over the period from 2010 until 2020. This is an indication of the corporate bond market liquidity for our period. A clear downward sloping trend can be observed, which means the average bid-ask spread on average decreases over time. This means the corporate bond market becomes on average more liquid over our time frame.

Figure 5: Average Current Yield over Time.

The average current yield is an average of all the corporate bonds in the sample. The period used is from 2010 until 2020, every bar represents one month. The average current yield is given as a fraction of 1. The mean of the current yield is 0.048, or 4.8%.

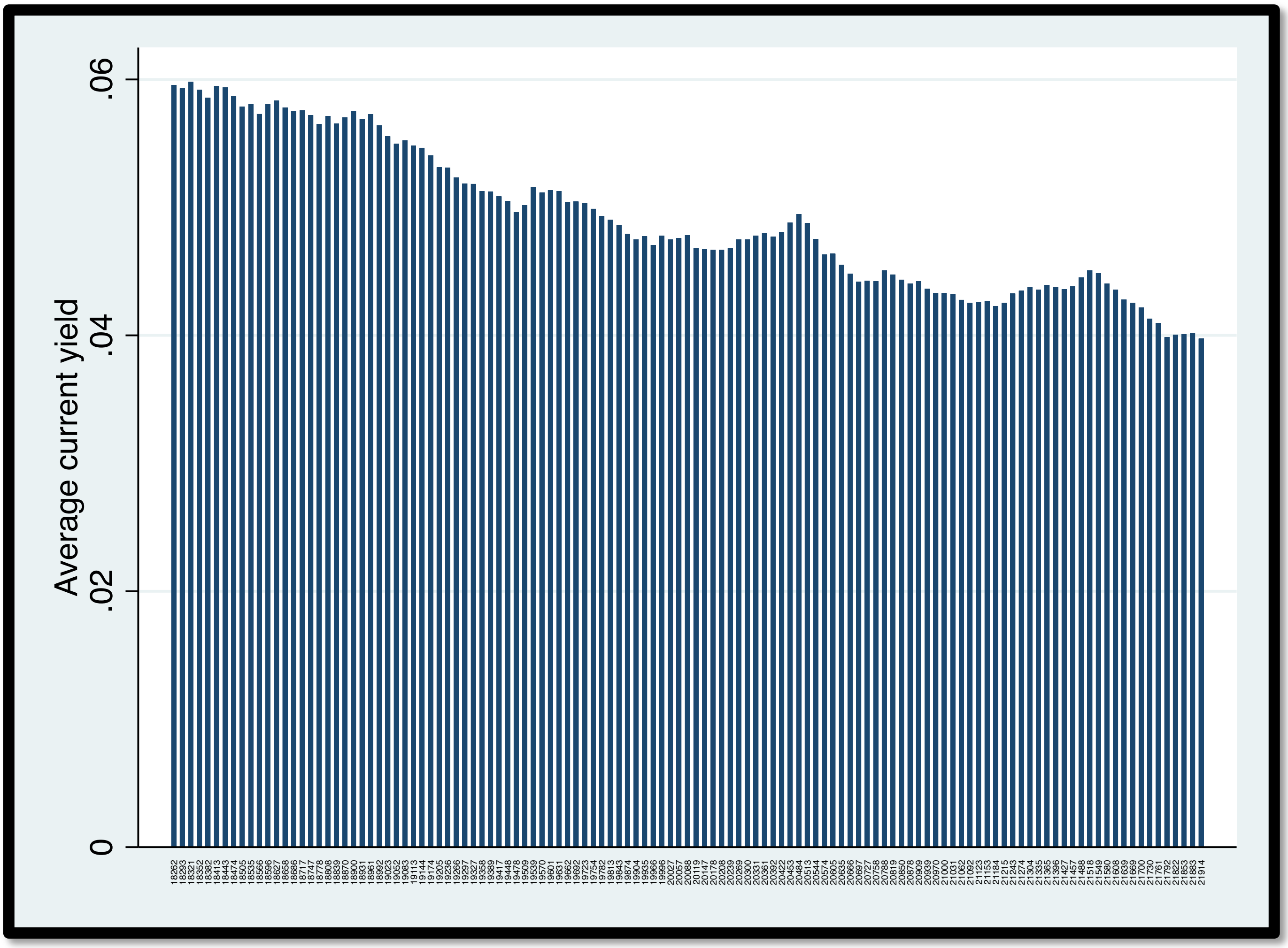


Figure 5 shows the $Average\ Current\ Yield$ over time. There's clearly a downward sloping trend visible for the period 2010 until 2020. This means that the $Current\ Yield$ decreases on average over the chosen time frame.

To conclude, we can see that the liquidity increases over time. Next to that we can see that the $Average\ Current\ Yield$ decreases over time. The next step is to look at the relationship between these two variables, using multiple regressions. Important to notice is that all standard errors are corrected for heteroskedasticity. I observed the presence of heteroskedasticity trough a scatter plot of the residuals against the fitted values, as well as the white test. Appendix B shows a scatterplot between the residuals against the fitted values. When a White Heteroscedasticity Test was performed, the p-value resulted in 0.000. When a 5% significance level is used, the null hypothesis which says there is homoscedasticity can be rejected. Because the White test as well as the scatter plot indicate heteroscedasticity, I will correct for it using robust standard errors.

Table 2: Correlation Table

Table 2 shows regression results for all the variables in our dataset. The correlation values are rounded to two decimals.

	Coupon	Amount Outstanding	Age	TMT	Duration	Log Volume	Bid-Ask Spread	YTM	Current Yield	Yield Spread
Coupon	1.00									
Amount Outstanding	-0.16	1.00								
Age	0.41	-0.23	1.00							
TMT	0.16	0.02	0.03	1.00						
Duration	0.11	0.03	-0.00	0.97	1.00					
Log Volume	-0.14	0.56	-0.46	-0.02	-0.02	1.00				
Bid-Ask Spread	0.25	-0.19	0.19	0.28	0.27	-0.19	1.00			
YTM	0.33	-0.06	0.03	0.15	0.12	0.05	0.24	1.00		
Current Yield	0.93	-0.15	0.31	0.11	0.05	-0.07	0.29	0.46	1.00	
Yield Spread	0.87	-0.14	0.36	0.21	0.17	-0.08	0.27	0.42	0.92	1.00

Table 2 shows a correlation table of all the variables we will use in the coming regressions. We see that TMT and $Duration$ have a very high correlation of 0.97, which can lead to collinearity. As a result, I remove the variable $Duration$ from the regressions. We can also see that $Coupon$ and $Current\ Yield$ have a very high correlation. This could lead to multicollinearity. In previous literature, like the paper from Chen, Lesmond and Wei (2007), the $Coupon$ was however used in all of the regressions. I will not use the $Coupon$ in the main model since it will interrupt the regression results by giving an extremely high R-squared. The regressions which include the $Coupon$ can however be found in appendix B.

Table 3: Regressions on Current Yield I

Table 3 starts with the regression results for the full sample as well as the regression results for 3 panels based on Moody's credit ratings. Model 1 represents the full sample. Model 2 only includes bonds with a $Credit\ Rating$ of 1-7 (A-rated bonds). Model 2 includes bonds with a credit rating of 8-16. (B-rated bonds) Model 3 only includes bonds with a credit rating of 17-21. (C-rated bonds)

	(1) Current Yield	(2) Current Yield	(3) Current Yield	(4) Current Yield
Bid-Ask Spread	0.1405^{***}	0.0649^{***}	0.0575^{***}	0.8876^{***}
	(0.0050)	(0.0071)	(0.0050)	(0.0296)
Amount Outstanding (100.000)	0.0002^{***}	0.0002^{***}	0.0000^{***}	0.0004^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
TMT (Years)	0.0004^{***}	0.0005^{***}	0.0004^{***}	0.0009^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log Volume	0.0002^{***}	0.0006^{***}	-0.0003^{***}	-0.0004^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
VIX	0.0002^{***}	0.0001^{***}	0.0002^{***}	0.0003^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CPI	-0.0004^{***}	-0.0003^{***}	-0.0004^{***}	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Credit Rating	0.0037^{***}	0.0018^{***}	0.0043^{***}	0.0067^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0002)
Age	0.0013^{***}	0.0018^{***}	0.0011^{***}	-0.0011^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	0.0909^{***}	0.0803^{***}	0.0980^{***}	-0.0295^{***}
	(0.0004)	(0.0007)	(0.0005)	(0.0057)
Observations	784708	295658	463419	25631
R^2	0.5999	0.5069	0.5448	0.1735
Adjusted R^2	0.5999	0.5069	0.5448	0.1733

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For the full-sample model we observe that the R-squared is equal to 59.99%. When all the control variables would be left out, the R-squared for the full-sample model is equal to only 8.64%. We can also see that all the variables are significant. When we look at the $Bid - Ask\ Spread$, we can see that it has a significant positive effect on the $Current\ Yield$. We need to be careful with specifying what this relationship actually means. When the $Bid - Ask\ Spread$ increases, the market becomes less liquid. Following the regression results, the $Current\ Yield$ will go down ceteris paribus. For the full sample we can state that when the average $Bid - Ask\ Spread$ increases with 1, and all else saying equal, the $Current\ Yield$ increases with 0.1405 or 14.05%. This seems high, but the reason is that an increase in the $Bid - Ask\ Spread$ of 1 is abnormal high.

When we look at models 2, 3 and 4 a couple of things are visible. One is that the R-squared drops significantly in model 4 compared to model 2 and 3. This means that the variables have less explanatory power when bonds with a bad credit rating are analyzed. A possible reason for this can also be the number of observations. Model 4 has relatively less observations compared to models 2 and 3. Another possible reason could be that the $Current\ Yield$ of bonds with a bad $Credit\ Rating$ is mostly influenced by factors other than we uncounted for. When we look at the regression results, another important trend is that the liquidity effect, proxied by the $Bid - Ask\ Spread$, increases its effect on the $Current\ Yield$ when a bond becomes less credit worthy. So when a bond has a bad $Credit\ Rating$, a change in the $Bid - Ask\ Spread$ can cause a relatively large increase in the $Current\ Yield$. With these findings there can be concluded that hypothesis 1, which stated that there is a negative relationship between liquidity and the $Current\ Yield$, cannot be rejected. For models 1 up until 4 we can see that the $Log\ Volume$, VIX, and CPI all have significant positive effects on the $Current\ Yield$. Age has a slightly negative but significant effect on the $Current\ Yield$.

Table 4: Regressions on Current Yield II

Table 4 shows the regressions results when we divide the sample into even more sub-panels. Models 1, 2 and 3 show regression results for bonds with a $Credit\ Rating$ 1 up until 7. Model 4, 5 and 6 show regression results for bonds with a $Credit\ Rating$ of 8 up until 16. Models 7, 8 and 9 show regression results for bonds with a $Credit\ Rating$ of 17 up until 21. Next to that, models 1,4 and 7 only include bonds with a TMT up until 7 years. Models 2, 5 and 8 only include bonds with a TMT of 8 up until 15 years. Models 3,6 and 9 only include bonds with a TMT greater than 15 years.

	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield
Bid-Ask Spread	-0.1290***	-0.0107	0.0298***	-0.0795***	0.0973***	0.2082***	0.9298***	0.5613***	0.8045***
	(0.0169)	(0.0099)	(0.0051)	(0.0109)	(0.0075)	(0.0059)	(0.0348)	(0.0614)	(0.0804)
Amount Outstanding (100.000)	0.0002***	0.0002***	0.0001***	0.0000***	0.0001***	0.0000***	0.0002***	0.0009***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)
TMT (Years)	0.0014***	0.0003***	-0.0003***	0.0010***	0.0002***	-0.0002***	0.0018***	-0.0006***	-0.0007***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)	(0.0002)
Log Volume	0.0012***	-0.0001***	-0.0002***	-0.0003***	-0.0002***	0.0001***	-0.0002	-0.0007***	-0.0007*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)	(0.0004)
VIX	0.0001***	0.0003***	0.0001***	0.0002***	0.0004***	0.0002***	0.0003***	0.0006***	0.0006***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
CPI	-0.0003***	-0.0004***	-0.0004***	-0.0005***	-0.0004***	-0.0003***	0.0000	-0.0003***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Age	0.0028***	0.0012***	0.0003***	0.0017***	0.0009***	0.0004***	-0.0014***	0.0001*	-0.0004*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0003)
Credit Rating	0.0020***	0.0010***	0.0009***	0.0046***	0.0038***	0.0036***	0.0062***	0.0107***	0.0127***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0006)	(0.0008)
Constant	0.0614***	0.1147***	0.1394***	0.1064***	0.0914***	0.0946***	-0.0287***	-0.0239	-0.1987***
	(0.0010)	(0.0010)	(0.0007)	(0.0007)	(0.0009)	(0.0009)	(0.0066)	(0.0160)	(0.0265)
Observations	170739	53463	71456	267619	111530	84270	21240	3059	1332
R^2	0.5321	0.6648	0.5989	0.5464	0.6483	0.5772	0.1726	0.3542	0.3276
Adjusted R^2	0.5321	0.6648	0.5989	0.5464	0.6482	0.5772	0.1723	0.3525	0.3235

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Remarking is that all the variables in table 4 have significant effects on the current yield within the 5% significance level. When we take a look at the R-squared, we can observe multiple trends. Firstly, we observe that models with a bad $Credit\ Rating$ got a lower R-squared. We observed the same from the regression results in table 3, so this shows consistency. Another important and new observation is that the R-squared decreases when the time to maturity increases. A possible reasoning could be that the longer the horizon and thus the time to maturity is, the more other factors play a role in determining the $Current\ Yield$.

From table 4 , a pattern in the $Bid - Ask\ Spread$ can also be observed. An increase in the $Bid - Ask\ Spread$, has a greater effect on the $Current\ Yield$ for bonds with a longer time to maturity (TMT). With these findings hypothesis 2, which stated that the magnitude of the relationship between liquidity and the current yield differs across credit ratings and time to maturities, cannot be rejected.

For models 5 up until 13 we can see that the VIX has a significant positive effect on the $Current\ Yield$. Age, CPI and $Log\ Volume$ do not have consisted positive or negative effects on the $Current\ Yield$, the effects are however significant in every single case.

Table 5: Regressions on Yield Spread I

Table 5 shows the regressions on the $Yield\ Spread$. Model 14 gives represents the full sample. Model 15 only includes bonds with a $Credit\ Rating$ of 1 up until 7. (A-bonds) Model 16 only includes bonds with a $Credit\ Rating$ of 8 up until 16. (B-bonds) Model 17 only includes bonds with a $Credit\ Rating$ of 17 up until 21. (C-bonds)

	(14) Yield Spread	(15) Yield Spread	(16) Yield Spread	(17) Yield Spread
Bid/Ask Spread	0.0795***	-0.0611***	0.0206***	0.9017***
	(0.0055)	(0.0081)	(0.0059)	(0.0309)
TMT (Years)	0.0007***	0.0008***	0.0007***	0.0011***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
VIX	0.0001***	0.0001***	0.0002***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CPI	-0.0001***	0.0000*	-0.0001***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Credit Rating	0.0038***	0.0017***	0.0044***	0.0063***
	(0.0000)	(0.0000)	(0.0000)	(0.0002)
Age	0.0015***	0.0020***	0.0013***	-0.0010***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log Volume	0.0002***	0.0007***	-0.0004***	-0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Amount Outstanding	0.0002***	0.0002***	0.0001***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-0.0091***	-0.0265***	0.0036***	-0.1347***
	(0.0005)	(0.0008)	(0.0006)	(0.0062)
Observations	784708	295658	463419	25631
R^2	0.5492	0.5162	0.4697	0.1591
Adjusted R^2	0.5492	0.5162	0.4697	0.1589

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The R-squared of the full model is equal to 54.92%, which is means the model has a good explanatory power. When all the control variables would be left out, the full-sample model would only have a R-squared of 7.12%. Again a relatively low explanatory power at the C-Bonds model in terms of a relatively low R-squared can be observed. We also observe that all the variables have a significant effect on the $Yield\ Spread$ within the 5% significance level.

When we look at the TMT, we see it has a larger effect on the $Yield\ Spread$ when the bonds have a lower $Credit\ Rating$. A possible reason could be that the uncertainty of a longer time to maturity in combination with the uncertainty of a lower $Credit\ Rating$, causes a higher $Yield\ Spread$. When the liquidity effects on the $Yield\ Spreads$ are analyzed, I observed a remarkable coefficient. For the full sample an increase in the $Bid - Ask\ Spread$ results in a higher $Yield\ Spread$. This positive relation is also the case for model 16 and 17. (B- and C-bonds) Model 15 (A-bonds) show however a negative relationship. When the $Bid - Ask\ Spread$ increases with 1, the $Yield\ Spread$ decreases on average with -0.0611. (-6.11%) So when bonds with a "good" A-rating face a higher bid-ask spread and become less liquid, their yield goes down on average as well. This seems to be contrary to the fact of a liquidity premium and also is contrary to the liquidity regression coefficients in tables 3 and 4. When we look at the proxy variables we observe that the VIX, CPI and $Amount\ Outstanding$ all have significant positive effects on the $Yield\ Spreads$. Age only has a significant negative effect on the $Yield\ Spread$ of C-rating bonds. With these findings I can only find partial support for hypothesis 3, which stated that there is a negative relationship between liquidity and yield spreads. We do see a negative relationship between liquidity and yield spreads in models 14, 16 and 17. However, we also observe a remarkable positive relationship in model 15, where only "good" A-rated bonds are included.

Table 6: Regressions on Yield Spread II

Model 18 includes observations before 01-01-2015. Model 19 includes observations after 01-01-2015. We choose this particular date because it's the exact middle of our time frame.

	(18) Yield Spread	(19) Yield Spread
Bid-Ask Spread	0.0679***	0.2933***
	(0.0061)	(0.0091)
TMT (Years)	0.0002***	0.0010***
	(0.0000)	(0.0000)
VIX	0.0004***	0.0001***
	(0.0000)	(0.0000)
CPI	0.0004***	-0.0003***
	(0.0000)	(0.0000)
Credit Rating	0.0040***	0.0034***
	(0.0000)	(0.0000)
Age	0.0010***	0.0019***
	(0.0000)	(0.0000)
Log Volume	0.0001***	0.0003***
	(0.0000)	(0.0000)
Amount Outstanding	0.0003***	0.0001***
	(0.0000)	(0.0000)
Constant	-0.1054***	0.0349***
	(0.0012)	(0.0008)
Observations	356860	427848
R^2	0.5280	0.6567
Adjusted R^2	0.5280	0.6567

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Lastly, I divided the full sample into 2 sub periods. Important to notice is that the R-squared of both models is about equal and about the same as the previous R-squares in models 1 and 14. In both models all the variables have significant effects within the 5% significance level. Remarkable is that the effect of an increase in the $Bid - Ask\ Spread$ has a much larger effect on the $Yield\ Spread$ in model 19 compared to model 18. In other words; a change in liquidity has on average a stronger effect on the yield spreads in 2015-2020 compared to 2010-2014. In conclusion, hypothesis 4 which stated that the magnitude of the relationship between liquidity and yield spreads differs across multiple sub periods and credit ratings, cannot be rejected.

My results show that the relationship between liquidity and U.S. corporate bond current yields are negative for the period 2010-2020. With this result, we cannot reject hypothesis 1, which stated that liquidity will have a negative relationship with a bonds current yield. The magnitude of this relationship also differs between bond characteristics such as credit rating and time to maturity. Given this, I cannot reject hypothesis 2, which stated that the relationship between current yields and liquidity differs across different credit ratings and time to maturities. These findings are similar to previous studies that studied a comparable context, such as the papers from Duffee (1999) and Chen, Lesmond and Wei (2007). My results however also show that the relationship between liquidity and US corporate bond yield spreads are not always negative. For A-rating corporate bonds I found a positive relationship between liquidity and bond yield spreads for the period 2010-2020. This finding differs from previous literate such as Longstaff, Mithal and Neis (2005) who always found a negative relationship between liquidity and bond yield spreads. The difference between my outcome and the outcomes of previous literature could be a result of using a different method for calculating the yield spreads. For calculating the yield spreads, I used the 10-years and 30-years interest rates, where previous literature most of the time used even duration governments bonds. Another possible reason that explains the difference is the time period and number of observations. I studied a relatively large sample for a different period in time.

I also found that the magnitude of the relationship differs between bond characteristics such as credit ratings and multiple sub time periods. This is partly in line with the study of Helwege, Huang and Wang (2015) who showed a negative relationship, but also concluded that this relationship was time varying. My study shows that hypothesis 3, which stated that the relationship between liquidity and U.S. corporate bond yield spreads is negative, can be rejected. I found a positive relationship for A-rated bonds. My study also shows that hypothesis 4, which stated that the relationship between liquidity and U.S. corporate bond yield spreads differs across time and credit ratings is different, cannot be rejected.

CHAPTER 6 Conclusion

In this paper, I studied the effects of liquidity on corporate bond current yields as well as corporate bond yield spreads. Therefore, I used U.S. corporate bonds and used a study period of 2010 until 2020. Previous research all found negative relationships between liquidity and corporate bond yields. However, this relationship hasn't been studied over the period 2010-2020. It also remained unclear how the magnitude of the relationship would differ over different credit ratings, time to maturities and time periods. Therefore, the question that was studied in this paper was: "What is the direction and the magnitude of the relationship between liquidity and corporate bond yields, and does this differ across different credit ratings, time to maturities and time periods."

To answer the research question, I have studied 19,209 different corporate bonds from 2,603 unique companies. First, I looked at the economic situation in the selected time frame, which is from 2010-2020. Afterwards I performed multiple regressions on the current yield as well as on the yield spread. Within these regressions, I used proxies in order to control for various risks and differences in bond characteristics. The most important variable is the bid-ask spread, which is used as a proxy for liquidity. The regressions showed that there is a negative relationship between liquidity and current yields. The magnitude of this relationship also differs between different credit ratings and time to maturities. The regressions also showed that there is not always a negative relationship between liquidity and yield spreads. I found that there is a positive relationship between liquidity and yield spreads within A-rated corporate bonds. A warning is the fact that the way of calculating the yield spreads differs in my study, compared to previous literate. I also found that the magnitude of the relationship differs between different credit ratings and time periods.

This paper concludes that although previous research found a negative relationship between liquidity and bond yields, a positive relationship can also exist. We have evidence for this with A-rated bonds in the period 2010 until 2020. I do however conclude that the relationship most of the time is negative, as previous literature suggests. I also conclude that the magnitude of the relationship differs between credit ratings, time to maturities and time periods. Previous literature already suggested that the liquidity premium is time varying. The explicit observation that it also differs between different credit ratings and time to maturities is an extension to previous literature.

A potential limitation for this study is the frequency of the data points. For the analysis, I used monthly data in order to make the data workable. Future researchers are encouraged to use daily data, in order to make the data even more detailed and increase the statistical power of the analysis even more. Another potential limitation for this study is in the calculation of the yield spreads. In this research I only used the 10-years and 30-years interest rates as risk free rates and compared this to the yields in the spread calculation. A preferable method would be to match the risk-free rates with the exact corresponding time to maturities to each observation. Unfortunately, I couldn't use this method because I didn't have access to all those precise data. It would be interesting to see how the results would differ when using this method compared to the method I used in this research.

REFERENCES

Adrian, T., Fleming, M., Shachar, O., Stackman, D., & Vogt, E. (2015). Has Liquidity Risk in the Corporate Bond Market Increased? *Liberty Street Economics*. https://libertystreeteconomics.newyorkfed.org/2015/10/has-us-corporate-bond-market-liquidity-deteriorated/

Ang, A., Bhansali, V., & Xing, Y. (2014). The Muni Bond Spread: Credit, Liquidity, and Tax. *Columbia Business School Research Paper*, *14*(37).

Balasubramnian, B., & Cyree, K. B. (2011b). Market discipline of banks: Why are yield spreads on bank-issued subordinated notes and debentures not sensitive to bank risks? *Journal of Banking and Finance*, *35*(1), 21–35. https://doi.org/10.1016/j.jbankfin.2010.07.015

Campbell, J., & Taksler, G. B. (2003). Equity Volatility and Corporate Bond Yields. *Journal of Finance*, *58*(6), 2321–2350. https://doi.org/10.1046/j.1540-6261.2003.00607.x

Chen, L. Q., Lesmond, D. A., & Wei, J. Z. (2007). Corporate Yield Spreads and Bond Liquidity. *Journal of Finance*, *62*(1), 119–149. https://doi.org/10.1111/j.1540-6261.2007.01203.x

Damodaran, A. (1999). Financing Innovation and Capital Structure. *Journal of Applied Corporate Finance*, *12*(1), 28–39. https://doi.org/10.1111/j.1745-6622.1999.tb00657.x

Duffee, G. R. (1996). Estimating the Price of Default Risk. *Review of Financial Studies*, *12*(1), 197–226. https://doi.org/10.1093/rfs/12.1.197

Ericsson, J. L. E., & Renault, O. (2006). Liquidity and Credit Risk. *Journal of Finance*, *61*(5), 2219–2250. https://doi.org/10.1111/j.1540-6261.2006.01056.x

Fisher, L. (1959). Determinants of Risk Premiums on Corporate Bonds. *Journal of Political Economy*, *67*(3), 217–237. https://doi.org/10.1086/258172

Helwege, J., Huang, J., & Wang, Y. (2014). Liquidity effects in corporate bond spreads. *Journal of Banking and Finance*, *45*, 105–116. https://doi.org/10.1016/j.jbankfin.2013.08.018

Houweling, P., Mentink, A., & Vorst, T. (2005). Comparing possible proxies of corporate bond liquidity. *Journal of Banking and Finance*, *29*(6), 1331–1358. https://doi.org/10.1016/j.jbankfin.2004.04.007

Liu, Y., Ning, Y., & Davidson, W. N. (2010). Earnings Management Surrounding New Debt Issues. *The Financial Review*, *45*(3), 659–681. https://doi.org/10.1111/j.1540-6288.2010.00265.x

Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *Journal of Finance*, *60*(5), 2213–2253. https://doi.org/10.1111/j.1540-6261.2005.00797.x

Lovo, S., Raimbourg, P., & Salvadè, F. (2022). Credit rating agencies, information asymmetry and US bond liquidity. *Journal of Business Finance & Accounting*, *49*(9–10), 1863–1896. https://doi.org/10.1111/jbfa.12610

Muranaga, J., & Ohsawa, M. (1999). Measurement of liquidity risk in the context of market risk calculation. *Institute for Monetary and Economic Studies*.

OECD. (2020). Corporate bond debt continues to pile up. *oecd.org*. https://www.oecd.org/newsroom/corporate-bond-debt-continues-to-pile-up.htm

Sarig, O., & Warga, A. (1989). Some Empirical Estimates of the Risk Structure of Interest Rates. *Journal of Finance*, *44*(5), 1351–1360. https://doi.org/10.1111/j.1540-6261.1989.tb02657.x

Appendix A

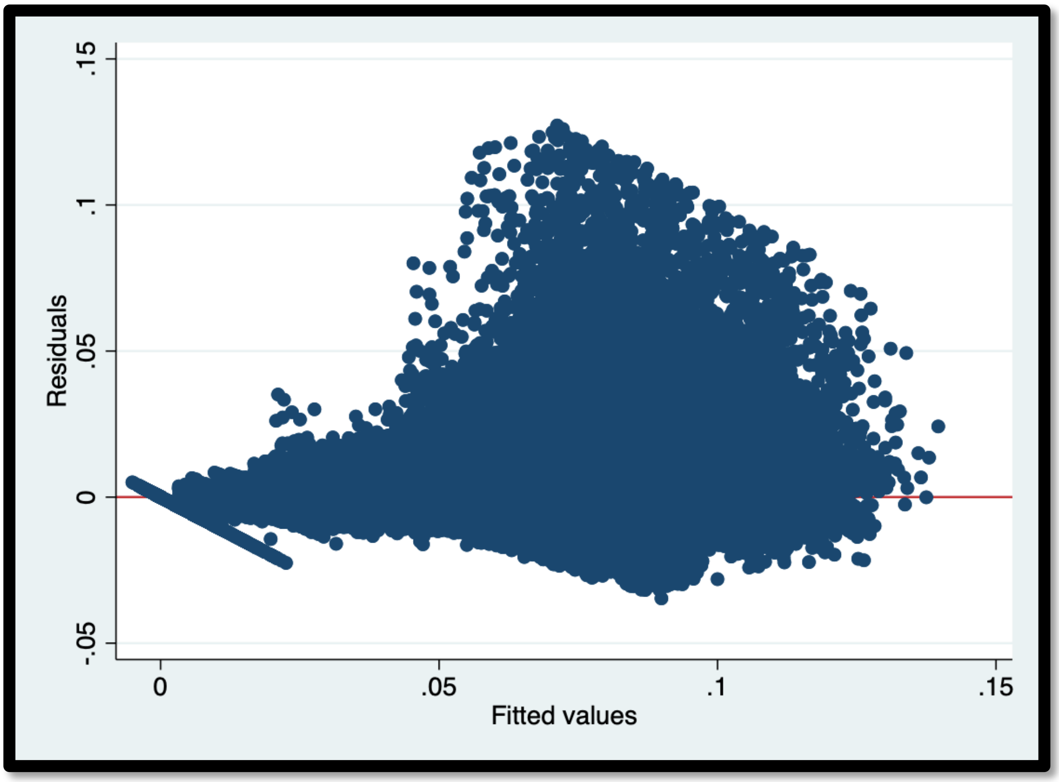
Table A: Conversion Table Moody's Credit Ratings
Table A gives the numerical Moody's credit rating as well as the corresponding alphabetical Moody's credit rating.

Numerical Moody's credit rating	Alphabetical Moody's credit rating		Numerical Moody's credit rating	Alphabetical Moody's credit rating
1.	AAA		12.	BA2
2.	AA1		13.	BA3
3.	AA2		14.	B1
4.	AA3		15.	B2
5.	A1		16.	B3
6.	A2		17.	CAA1
7.	A3		18.	CAA2
8.	BAA1		19.	CAA3
9.	BAA2		20.	CA
10.	BAA3		21.	C
11.	BA1			

Appendix B

Figure B: Scatter plot of the Residuals against the Fitted Values.

Figure B shows a scatter plot of the residuals against the fitted values. The graph clearly shows that the residuals increase when the fitted values increase. This trend is a sign of heteroscedasticity, for which we need to correct using robust standard errors.



Appendix C

Table C1: Regressions on Current Yield I (Including Coupon)

Table C1 starts with the regression results for the full sample as well as the regression results for 3 panels based on Moody's credit ratings. Model 1 represents the full sample. Model 2 only includes bonds with a credit rating 1-7 (A-rated bonds). Model 2 includes bonds with a credit rating of 8-16. (B-rated bonds) Model 3 only includes bonds with a credit rating of 17-21. (C-rated bonds) Where the variable Coupon is excluded from all regressions in the main models, the variable Coupon is included in all regressions given in the appendixes.

	(1) Current Yield	(2) Current Yield	(3) Current Yield	(4) Current Yield
Bid-Ask Spread	0.3057***	0.0740***	0.2270***	1.1713***
	(0.0038)	(0.0028)	(0.0036)	(0.0247)
Amount Outstanding	0.0000***	-0.0000***	-0.0000***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
TMT (Years)	-0.0001***	-0.0001***	0.0000***	0.0010***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log Volume	0.0003***	0.0002***	0.0002***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
VIX	0.0001***	0.0000***	0.0001***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CPI	0.0000***	0.0000***	0.0000***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Credit Rating	0.0011***	0.0002***	0.0013***	0.0069***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Age	-0.0001***	-0.0002***	-0.0000***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Coupon	0.0079***	0.0083***	0.0077***	0.0094***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Constant	-0.0155***	-0.0063***	-0.0136***	-0.2259***
	(0.0003)	(0.0002)	(0.0003)	(0.0046)
Observations	784708	295658	463419	25631
R^2	0.9019	0.9463	0.9098	0.5561
Adjusted R^2	0.9019	0.9463	0.9098	0.5559

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C2: Regressions on Current Yield II (Including Coupon)

Table C2 shows the regressions results when we divide the sample into even more sub-panels. Models 1, 2 and 3 show regression results for bonds with a credit rating 1 up until 7. Model 4, 5 and 6 show regression results for bonds with a credit rating of 8 up until 16. Models 7, 8 and 9 show regression results for bonds with a credit rating of 17 up until 21. Next to that, models 1,4 and 7 only include bonds with a time to maturity up until 7 years. Models 2, 5 and 8 only include bonds with a time to maturity of 8 up until 15 years. Models 3,6 and 9 only include bonds with a time to maturity greater than 15 years. Where the variable Coupon is excluded from all regressions in the main models, the variable Coupon is included in all regressions given in the appendixes.

	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield	Current Yield
Bid-Ask Spread	0.0258***	0.0777***	0.1021***	0.2706***	0.1983***	0.2488***	1.2448***	0.7667***	0.8188***
	(0.0031)	(0.0044)	(0.0040)	(0.0071)	(0.0052)	(0.0053)	(0.0279)	(0.0619)	(0.0801)
Amount Outstanding	0.0000***	0.0001***	0.0000***	-0.0000***	-0.0000***	0.0000*	0.0001*	0.0006***	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)
TMT (Years)	-0.0005***	-0.0000***	-0.0000***	-0.0005***	-0.0000***	0.0001***	0.0015***	0.0005***	-0.0008***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)	(0.0002)
Log Volume	0.0002***	-0.0001***	-0.0001***	0.0002***	0.0001***	0.0001***	0.0006***	-0.0008***	-0.0007*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)	(0.0004)
Coupon	0.0089***	0.0071***	0.0038***	0.0085***	0.0066***	0.0043***	0.0097***	0.0059***	0.0055***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)	(0.0006)
VIX	0.0000***	0.0001***	0.0001***	0.0001***	0.0002***	0.0002***	0.0004***	0.0006***	0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
CPI	0.0001***	-0.0000***	-0.0002***	0.0001***	-0.0000***	-0.0001***	0.0004***	0.0001**	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Age	-0.0001***	-0.0002***	0.0000***	-0.0000***	-0.0000***	0.0001***	-0.0002***	-0.0002**	-0.0007***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0002)
Credit Rating	0.0001***	0.0002***	0.0005***	0.0009***	0.0016***	0.0032***	0.0061***	0.0111***	0.0131***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0005)	(0.0008)
Constant	-0.0128***	0.0166***	0.0687***	-0.0217***	-0.0022***	0.0210***	-0.2291***	-0.1840***	-0.2537***
	(0.0001)	(0.0006)	(0.0008)	(0.0003)	(0.0005)	(0.0009)	(0.0050)	(0.0159)	(0.0255)
Observations	170739	53463	71456	267619	111530	84270	21240	3059	1332
R^2	0.9862	0.9506	0.7617	0.9545	0.9099	0.7007	0.5928	0.4653	0.3640
Adjusted R^2	0.9862	0.9506	0.7617	0.9545	0.9099	0.7007	0.5927	0.4637	0.3596

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C3: Regressions on Yield Spread I (Including Coupon)

Table C3 shows the regressions on the yield spread. Model 14 gives represents the full sample. Model 15 only includes bonds with a credit rating of 1 up until 7. (A-bonds) Model 16 only includes bonds with a credit rating of 8 up until 16. (B-bonds) Model 17 only includes bonds with a credit rating of 17 up until 21. (C-bonds) Where the variable Coupon is excluded from all regressions in the main models, the variable Coupon is included in all regressions given in the appendixes.

	(14) Yield Spread	(15) Yield Spread	(16) Yield Spread	(17) Yield Spread
Bid-Ask Spread	0.2471***	-0.0522***	0.1944***	1.1966***
	(0.0043)	(0.0040)	(0.0044)	(0.0260)
TMT (Years)	0.0002***	0.0002***	0.0003***	0.0012***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
VIX	0.0001***	0.0000***	0.0001***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CPI	0.0003***	0.0004***	0.0003***	0.0008***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Credit Rating	0.0011***	0.0001***	0.0013***	0.0065***
	(0.0000)	(0.0000)	(0.0000)	(0.0002)
Age	0.0001***	0.0001***	0.0002***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log Volume	0.0003***	0.0003***	0.0001***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Coupon	0.0080***	0.0081***	0.0078***	0.0097***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Amount Outstanding	0.0000***	-0.0000***	-0.0000***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-0.1171***	-0.1116***	-0.1108***	-0.3389***
	(0.0003)	(0.0004)	(0.0004)	(0.0050)
Observations	784708	295658	463419	25631
R^2	0.8275	0.8487	0.8058	0.5305
Adjusted R^2	0.8275	0.8487	0.8058	0.5303

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C4: Regressions on Yield Spread II (Including Coupon)

Table C4 shows the regressions on the yield spread, divided into two time panels. Model 18 includes observations before 01-01-2015. Model 19 includes observations after 01-01-2015. We choose this particular date because it's the exact middle of our time frame. Where the variable Coupon is excluded from all regressions in the main models, the variable Coupon is included in all regressions given in the appendixes.

	(18) Yield Spread	(19) Yield Spread
Bid-Ask Spread	0.2193***	0.4038***
	(0.0045)	(0.0073)
TMT (Years)	-0.0002***	0.0005***
	(0.0000)	(0.0000)
VIX	0.0003***	0.0001***
	(0.0000)	(0.0000)
CPI	0.0007***	-0.0000***
	(0.0000)	(0.0000)
Credit Rating	0.0011***	0.0010***
	(0.0000)	(0.0000)
Age	-0.0002***	0.0003***
	(0.0000)	(0.0000)
Log Volume	0.0002***	0.0004***
	(0.0000)	(0.0000)
Coupon	0.0078***	0.0083***
	(0.0000)	(0.0000)
Amount Outstanding	0.0000***	-0.0000***
	(0.0000)	(0.0000)
Constant	-0.2049***	-0.0330***
	(0.0007)	(0.0005)
Observations	356860	427848
R^2	0.8644	0.8723
Adjusted R^2	0.8644	0.8723

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$