Liquidity and Liquidity Premium for Infrequent OTC Bonds: a Pair Analysis Approach

Hua, Yan

Awarding institution:
King's College London

Liquidity and Liquidity Premium for Infrequent OTC Bonds: a Pair Analysis Approach

Yan Hua

Supervised by: Alexandru Preda, Zhong Chen

King's College London

Nov 2022

Abstract

The liquidity of financial instruments has always taken a prominent place in financial research. Compared to the studies of equities and exchange-traded bonds, estimating an OTC-traded bond's liquidity premium has not been fully understood. OTC bonds are usually less frequently traded, which results in insufficient data, therefore, most current methodologies cannot be efficiently used. The question is: how is the liquidity premium of OTC bonds decided? Can we find an easier way to assess or estimate the liquidity premium of a bond lacking transaction or quotation data?

Based on two main research methods, the existing multi-factor regression and the instrument pair analysis, our research applies a "bond pair" method to study the relationship between bond liquidity factors and liquidity prices. This thesis studies the comparative "difference" of the liquidity factors and premium of a pair of bonds, avoiding the need for a large amount of data, and proposes a solution for the similar data lacking question. As a conclusion, our result shows that at a specific cross-sectional time, the bond issuing amount size and the life length have significant impacts on comparative bond liquidity premium, but the cash flow structure and credit rating (limited to investment-grade bonds) have a relatively small impact. It is also found that the issuing amount dominantly determines the bid-ask spread premium when a bond is close to maturity. We suggest the following research direction in the last, and a feasible estimation method is proposed for quotation considering liquidity premium. The potential advantage is that there is no need to consider and collect a large amount of data, avoiding potential data problems.

Keywords: OTC bond, Liquidity, Pair analysis

List of Contents

List of Tables

List of Figures

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Liquidity and Liquidity Premium for Infrequent OTC Bonds: a Pair Analysis Approach

Chapter 1 Introduction

1.1 Background

Bonds are favoured by investors due to their “fixed income” returns. As future cash flows are pre-determined, bond investors are able to design the relative future business activities correspondingly. Therefore, bonds have become one of the most important types of financial instrument in the market. The majority of bonds are traded bilaterally in OTC (over-the-counter) markets due to the non-standard structure type and the sizeable transactional amount. Typically, the OTC markets do not provide an apparent price reference in the same way as exchange markets, therefore, the OTC bond pricing matters.

In practice, traders prefer to bid yield instead of price. The two most important factors that traditionally determine the bond yield are credit rating and market interest rate. The corresponding parts of the yield are credit risk premium and market risk premium. However, in secondary market trading practice, it was found that these two factors cannot explain all yield premiums. Some experienced traders believe bonds with larger issuing amounts to be more liquid and more expensive. The on-the-run/off-the-run difference in the US Treasury markets is also an indication that age influences bond valuation. In addition, investors must often pay additionally when buying or selling non-frequent traded bonds. Such additional costs are generally expressed due to the spread of buying and selling, which is called the bid-ask spread. These are called liquidity impacts and other returns or costs are called liquidity premium.

Previous studies have explored bond liquidity and liquidity premium, including liquidity definition, liquidity proxies, liquidity premium, the correlation between liquidity proxies and returns, and the interaction between liquidity determinants. However, relatively few studies have attempted to estimate the individual OTC bond liquidity premium and these estimations have generally been under certain limitations. The estimation of the liquidity premium of an OTC-traded bond has yet to be fully understood.

1.2 Brief on Current Literature and the Research Question

In financial markets, there are generally two types of liquidity: asset liquidity (or market liquidity) and funding liquidity (or accounting liquidity). (Brunnermeier & Pedersen, 2008). In this thesis, asset liquidity is used.

Bond liquidity measures how easily a bond can be traded or converted to cash without having any impact on its market price. Early studies defined liquidity from many trading convenience dimensions, including execution price, transaction speed, transaction cost and market price elasticity (Lippman and McCall, 1986; Amihud & Mendelson, 1986; O'Hara, 1995; Glen, 1994; Massimb and Phelps, 1994). Amihud and Mendelson (1986) provided a typical representative summary of these thoughts, considering liquidity to be execution cost, a measurable definition. Harris (1991) established a 4-dimension (immediacy, width, depth and resiliency) liquidity characteristic model that is widely accepted by researchers who study liquidity.

Regardless of the definition of liquidity, it cannot be watched or measured directly. Therefore, finding a liquidity proxy is necessary for further studies. There are two main ways of doing this, the first being to look for original determinants. For example, Amihud and Mendelson (1986) studied product features, trading venue, the number of participants and their quotes, trade size of trades and the total available notional. Sometimes piecewise regression is run using dummy variables with these original determinants. The other way is to look for or create a quantifiable liquidity proxy formed by watchable elements. The bid-ask spread and the Amihud ILLIQ (Amihud, 2002) are the two proxies that are most popular in the liquidity studying areas. The zero-returns-days related proxies, such as LOT and zero-day percentage (Lesmond, Ogden, & Trzcinka, 1999), are more applicable to OTC bonds, while other proxies, such as serial covariance (Roll, 1984) and turnovers, are also valid proxies in different areas.

Bond liquidity premium is the additional required return (or cost from the buyer side) for illiquid bonds in comparison to liquid ones. From the perspective of an investor, additional costs/returns are partly from extra yield and partly from extra bid-ask spread. Bid-ask spread plays the dual roles of liquidity proxy and liquidity premium. Amihud and Pederson (1986, 2002) initially linked price differences with bid-ask spread and ILLIQ (Amihud, 2002) as a means of illustrating the liquidity difference in the security market. However, they still paid more attention to the stocks. Chen, Lesmond and Wei (2007) examined corporate bond liquidity by testing bid-ask spread and two zero-returns-day related measures, concluding that more illiquid bonds earn a higher yield spread. Friewald et al. (2012) studied the time-series changes in US corporate bond yields and discovered that market liquidity changes, with the bid-ask spread as a proxy, provide an explanation for a third of the changes in the yields. Researchers have also noticed and verified the positive feedback between bond credit risks and liquidity risks, finding the role to be more obvious for high-yield bonds (Huang & Huang, 2012; He and Milbradt, 2014; Min & Luo, 2016). A report that was published by the Bank for International Settlements (CBIS – CGFS, 2016) performed a direct comparison of the different liquidity dimensions (breadth and depth) in different markets, finding there to be significant differences among "volume-based" liquidity proxies but not on "price-based" ones, e.g., the bid-ask spread in different markets with different market makers (or brokers).

Researchers have estimated the liquidity premium through the use of different approaches. They are as follows: liquidity proxy-based regressions, credit premium-based methods, the optimal solution via the supply-demand pricing (microstructural models) and the comparing solution via similar/pair instruments (pair analysis methods). All of these methods have advantages and limitations, and they will be examined in more detail in the literature review section.

Several studies have illustrated how liquidity determinants impact the liquidity premium for bonds as a type of instrument and some have explained how individual liquidity is affected by market liquidity conditions. However, it is notable that relatively few studies have attempted to estimate individual OTC bond liquidity premium, and these estimations are generally under certain limitations. It can be interpreted that the above research approaches are more applicable to bonds with higher trading frequencies and larger transparency in data, including exchange-traded bonds. A key limitation of OTC bonds is the data and a lack of credible data on spreads or quotes is a major impediment in analysing liquidity and its impact on yield spreads (Goodhart & O'Hara, 1997; Chen et al., 2007). Although some approaches and proxies are more applicable for OTC bonds, including market liquidity portfolio matching regression, CDS pair analysis and zero proxies, there are still many practical limitations. Due to the contradiction between spread data limitation and the number of factors, the use of quantitative analysis for directly estimating the liquidity premium of an individual bond is not easy.

Therefore, the aim of this study is to examine the liquidity and pricing of an OTC-traded bond (particularly for an infrequently traded bond). It is hoped that the following question can be answered: Can a method be found for estimating the liquidity premium of an infrequent OTC-traded bond more easily, reducing the requirements on the number of factors and market data of traditional methods?

1.3 Data and Preliminary Results

Unstructured bonds issued by US-owned companies in US markets between January 2013 and June 2018 are examined in this study. Bonds are chosen with at least ten quotations during this period as a minimum threshold. Basic information and daily quotes of the bonds are collected manually from Bloomberg Terminal, with additional data relating to the historical credit change and the number of transaction days from the TRACE database.

The focus is solely on plain vanilla bonds, with a total of 4,900 bonds and 2,267,398 quoting behaviours following the screening process. The liquidity premium is then examined based on these quotations. It is noticed that the bond liquidity premium exhibits distinct characteristics at different life cycle stages, not only the on-the-run/off-the-run differences but also significant differences during the issuing, persisting and expiring periods.

The "pair analysis" approach is adopted to avoid traditional information data requirements and it is inspired by the cash flow matching approach in financial engineering. The liquidity determinants are classified according to their characteristics and impacting method, “bond pairs” are constructed according to specific rules, similar impacts are eliminated by comparing similar bonds and analysis is conducted based on the results of different bond pair types. Many determinants of crucial importance are offset between pairs during this process, simplifying the model and reducing the data requirements.

By conducting tests of different liquidity determinants, the relationship between key factors and the spreads is verified. It is found that bond issuing amount, bond life and bond age are most influential and significant when the liquidity premium difference between two bonds is determined. Bond cash flow structure impact is relatively minimal and rating impact is not as significant as anticipated. In addition, yield difference and bid-ask spread difference perform differently as bond age increases. This is interpreted as the yield and bid-ask spread standing for different liquidities. Yield liquidity premium is the additional premium investors are willing to pay for an illiquid bond due to its liquidity feature. Bid-ask spread liquidity premium stands not only for the bond liquidity feature but also for the bond liquidity status in the market, which is driven by both individual liquidity and market supply. The former has a closer relationship to the features of the bond itself, whereas the latter is more relevant to the market. As bond age increases, the liquidity of the two has a tendency to be consistent.

1.4 Contributions

The liquidity determinants of OTC corporate bonds, including liquidity measurements and impacting factors, are tested and verified. Unlike previous studies, special attention is paid to the liquidity premium difference between two bonds using a pair analysis approach. Bond issuing amount and bond age are the two main factors that determine the liquidity premium difference between two bonds, while cash flow structure and credit rating are not significant. Different bond yield liquidity premium and bid-ask liquidity premium performance is noticed when far away from or close to maturity and the bid-ask spread can be dominated by the issuing amount if the bond is about to expire. The potential difference and the internal relationship between the two premiums are interpreted.

The pair analysis results reveal that some determining factors can explain the liquidity premium difference between two bonds. A feasible method is proposed for the estimation of the liquidity premium of an individual infrequently traded OTC bond using pair analysis. This method does not require the collection of too much data relating to specific bond issuer information, financial information, market environment information and historical quotations or transactions, which makes the estimation simpler and more efficient.

Chapter 2 Literature Review

It was considered that bond valuation factors are mainly market interest rates and credit status. Researchers designed indicators for identifying and measuring factors and established structured models for the accurate calculation of bond valuation and price. However, since the 1990s, it has been noticed that these two factors fail to fully explain the bond price. Taking trading zero-coupon US Treasuries (without credit risk) as an example, investors found that those with the same maturity dates did not have the same market price. Differences were marginal, but they were noticeable. In particular, significant price differences were observed between newly-issued and old Treasuries, which is known as the on-the-run/off-the-run difference. Researchers interpreted this as liquidity impact. They defined liquidity, found liquidity determinants, established liquidity proxies and explained the premium.

In this thesis, the literature review relates only to bond liquidity, liquidity risk and liquidity pricing.

2.1 Concepts of Liquidity, Liquidity Premium and Liquidity Risk

It is essential to first clarify the fundamental concept of liquidity and define the differences between the concepts of liquidity, liquidity premium and liquidity risk.

In financial terms, there are two different types of liquidity: asset liquidity (or market liquidity) and funding liquidity (or accounting liquidity). Asset liquidity is the ease with which an asset is traded and converted to cash without having any effect on its market price, whereas funding liquidity is the ease of obtaining funding to meet the financial obligation (Brunnermeier & Pedersen, 2008). Both types of liquidity interact with each other. For example, selling is difficult when other investors are encountering funding problems, and obtaining funds when the collateral is difficult to sell is also problematic (Brunnermeier & Pedersen, 2008). The focus of this thesis is on asset (bonds) liquidity and the related concepts of liquidity, liquidity risk and liquidity premium are all related to the asset (bonds).

Bond liquidity measures the ease of trading a bond and converting to cash without influencing its market price. A typical bond liquidity proxy is the bid-ask spread. Bond liquidity premium is the additional cost required (or cost from the buyer side) for an illiquid bond in comparison to a liquid bond. This additional return or cost can be in any form.

The concept of bond liquidity risk is often unclear. In a European Central Bank working paper, Nikolaou (2009) concluded liquidity risk to also be the cost, in the same way as liquidity premium. In some other cases, it accounts for trading cost uncertainty: the uncertainty of the additional bid-ask spread of transactions to buy or sell in a short period of time. This uncertainty is not examined in this paper.

2.2 Bond Liquidity: Definition and Proxies

2.2.1 Liquidity Definition

Market liquidity research began before there was a clear distinction between markets and before the establishment of funding liquidity. Early studies believed liquidity to be a price allowing immediate execution (Lippman & McCall, 1986; Schwartz, 1988). The influence of liquidity meant the speed of completing a transaction, which is also known as immediacy (Amihud & Mendelson, 1986; O'hara, 1995). Amihud and Mendelson (1986) provided a typical representative summary of these thoughts, where liquidity was considered as the ***execution cost***: the cost of transactions to buy or sell in a short period or the expected time required for completing a transaction at a reasonable price. They took the cost as an input for testing the sensitivity of the spread. Treating liquidity as cost was convenient for the calculation of liquidity, and this became widely accepted by the financial industry. A report on fixed income liquidity that was published by the Bank for International Settlements accepted the concept of cost (BIS-CGFS, 2016), but clarifying and measuring the “short period” was difficult. It was easy to comprehend the cost concept, but applying it was difficult.

Some studies that did not treat liquidity as cost still considered liquidity to be a rapid execution ability that would not have any significant effect on market price (Glen, 1994; Massimb & Phelps, 1994) (known as resiliency). The resiliency reflected the relationship between normal price and a slight price shock, but it could still not solve the problems when a trend changed.

Harris (1991) established a liquidity description model and examined liquidity from four dimensions: immediacy, width, depth and resiliency. **Immediacy** refers to whether a transaction can be executed immediately if an investor wishes to trade. Not only does it mean immediacy at a reasonable price, but it also referred to unobstructed trading facilities in some studies, including a facility to reduce the costs of searching, information asymmetry and currency transmission. **Width** (or breadth) refers to the different available prices in the market caused by the variety of participants and standing for the cost of instant execution. It is generally measured by the current spread on the market. **Depth** shows the quantity that is available for sale or purchase away from the current market price and measures the trading capacity of a considerable amount. **Resiliency** is the ability/speed of a price return to equilibrium due to some shock volumes. The entire industry has accepted the liquidity concept. For example, the Bank for International Settlements adopted it as a definition of liquidity in its most recent liquidity research report. (BIS – CGFS, 2016)

2.2.2 Liquidity Proxies

Appropriate liquidity proxies and accurate quantitative measures must be employed when studying liquidity impacts. However, many proxies or measures are only available or perform well in specific environments. A perfect proxy or measure that fits all situations is yet to be found.

(1) Bid-Ask Spread

Bid-ask spread (the difference between the best bid and the best ask) is one of the most traditional and popular liquidity proxies. Amihud and Mendelson (1986) found that illiquidity could be measured by the cost of immediate execution using the bid-ask spread: "an investor willing to transact faces a trade-off: he may either wait to transact at a favourable price or insist on immediate execution at the current bid or ask price" (Amihud & Mendelson, 1986).

As market microstructure theory developed, several **other spreads** developed. The effective spread (the difference between transaction price and middle price) measures the execution cost of orders by assuming the previous middle price estimated the security value unbiasedly (Glosten & Harris, 1988). The realized spread (the difference between transaction price and middle price in the next timestamp) slightly alters the effective spread by supposing the post-transactional middle price to be a more accurate estimation of true value (Stoll, 1989; Huang & Stoll, 1996). The positioning spread (the difference between the middle price and the middle price in the next timestamp) is a combination of the two previous spreads, and it reflects the "true value" change (Naik & Yadav, 1999). However, in many environments, these spreads require further transactional information, while the bid-ask spread only requires quotation information and is used more regularly in OTC markets.

It should be noted that the bid-ask spread is not just a liquidity proxy. It is also a combination of three different types of costs according to market microstructure theory: order processing cost (Demsetz, 1968), inventory holding cost (Demsetz, 1968; Garman 1976, Amihud & Mendelson, 1980) and information asymmetry /adverse selection (Bagehot, 1971; Copeland & Galai, 1983; Glosten & Milgrom, 1985; Glosten & Harris, 1988; Brennan & Subrahmanyam, 1996). It also helps interpret what liquidity is.

However, the bid-ask spread and the other spread methods all have limitations. In the majority of OTC markets, bid-ask spreads cannot always be directly achieved for all times or all bonds, particularly with less frequently traded bonds. This means that the bid-ask spread is only applicable for bonds that are above a certain trading frequency level and not for illiquid bonds. For example, many high-yield bonds are illiquid and do not have existing public bid-ask spreads. This limits the use of the bid-ask spread in the measurement of the liquidity of high-yield bonds. In addition, the bid-ask spread does not contain trading amount information. Investors must further quote for the exact price with the transactional amount, resulting in the bid-ask spread not considering market depth and being unsuitable for the prediction of individual transaction prices. Without any market depth consideration, the bid-ask spread will be a biased liquidity estimation (Naik & Yadav, 1999). Naik and Yadav examined the change of spreads for the study of liquidity rather than using them directly, taking advantage of the elimination of biased estimation in an individual transaction process.

(2) Volume-Price Impact Indicators

Amihud and Mendelson (2002) created a new liquidity proxy through summing the daily return per volume (ILLIQ). The visual explanation of ILLIQ is clear: liquidity can be measured based on the return (price change) per volume. The higher the price change per volume is, the lower the liquidity will be. ILLIQ was tested and found to be one of the best liquidity proxies for the measurement of entire stock market liquidities in different countries and developing countries were found to generally have a higher liquidity premium (Kim & Lee, 2013; Min & Luo, 2016; Amihud, Hameed, Kang, & Zhang, 2015). The proxy has also been used in bond markets for testing liquidity (Wang, 2013; Zhu, 2013; Wang & Wen, 2016). The Amihud model was developed as a means of illustrating the systematic liquidity risk of the entire market (Amihud & Mendelson, 2015).

The Amihud model has the advantage of clearly showing the relationship between price and volume. However, it requires a series of transactional data for examining the relationship between price and volume during different short periods. When the liquidity of an individual bond is estimated, this model may be limited as transactional data is only available for a few bonds in the OTC markets.

(3) Volume Indicators

Volume indicators (order volume, transaction volume and turnover) have also been used for proxies in the research of buying/selling activity liquidity. Glen (1994) took the trading volume of the current price as a liquidity proxy and overcame the weakness of data collection in some illiquid instruments.

In comparison to different markets, the exchange-traded bond turnover indicators do not represent liquidity as well as those of the stock markets (Johnson, 2008; Min & Luo, 2016). In addition, studies have shown that the volume presents different correlations (either positive or negative) with spreads in different markets (Datar, Naik, & Radcliffe, 1998; Degryse, 1999). At the same time, a number of studies have treated volume indicators as proxies for volatility than liquidity and trading volatility is considered different to liquidity in current research (Kamara, 1994). However, high volatility does not necessarily always mean high liquidity. A recent study supported that volatility varies more than liquidity, using volume indicators as proxies (Oprică & Weistroffer, 2019).

Volume indicators are not generally treated as liquidity proxies alone and are often used in combination with others, such as the relationship with price change (price impact). When researching individual bond purchasing/selling activities (market microstructure), volume indicators are widely considered together with the spread and other factors.

(4) Zero Indicators

LOT, zero-returns-days and zero-trading-days methods were specifically useful for securities or markets where there are less frequent transactions. Lesmond et al. (1999) introduced the zero-return-days method for estimating liquidity (LOT model). The LOT model provides a comprehensive estimate of liquidity based on the spread and other costs and contains a measurement of zero-returns-days occurrence. The economic interpretation is that transactions only occur when the value of private information is high enough for transaction costs to be fully covered and when liquidity is low, the high costs prevent the transaction. Bekaert, Campbell, and Lundblad (2003) proved that the percentage of zero-returns-days is a reasonable liquidity proxy when studying transactions in emerging markets. Lee (2011) also used zero-returns-days as a means of measuring liquidity in emerging markets. Zero-trading-days is quite similar to zero-returning days, but the most significant difference is counting the “no transaction days”. Chen et al. (2007) argued that the zero-returns-days indicator is a noisy liquidity measure, although it estimates liquidity more accurately. The researchers compared LOT, zero-returns-days and bid-ask spread in an analysis of liquidity premium regression on liquidity determinants. Dick-Nielsen et al. (2009, 2012) collected zero-trading days for corporate bonds with adjustment, studying liquidity change both before and after the subprime crisis. Kang & Zhang (2014) adopted zero-volume days for measuring emerging market liquidity.

These methods are principally transaction frequency methods. In comparison to other proxies, they only require the collection of the time series of daily returns, so they can overcome data collection difficulties in certain markets. They are applicable for the measurement of OTC corporate bond liquidity, but explanatory power is less than the price-impact indicator or the bid-ask spread proxy in an environment of sufficient data

(Chen et al., 2007; Kim & Lee, 2013).

(5) Other Proxies

Correlation between past and subsequent prices. Roll (1984) used the covariance between past and subsequent prices as a means of measuring liquidity in the US market, which is known as serial covariance. Stoll (1989) raised a similar indicator, and Bao, Pan, and Wang (2011) adopted a Roll extraction method for examining the liquidity of investment-grade and high-yield grade bonds in the US market, testing them with a drawback of being reliant on the hypothesis of an information-effective market.

Issuing/outstanding amount. In practical bond markets, the issuing amount (outstanding volume) is one of the most important factors in OTC bonds that is used by traders. There is no rigorous verification due to a lack of OTC bond liquidity studies, but it remains widely accepted in practice. Many traders have found price differences between two bonds where everything is equal with the exception of the issuing amount. Longstaff et al. (2005) performed verification using regressions and found firm evidence of the importance of the outstanding principal amount in the explanation of bond liquidity.

(6) Proxies Comparison

Previous studies have tested the effectiveness of liquidity proxies by performing a comparison of the correlation between liquidity proxies and liquidity premiums (this will be introduced later) in different market environments. Edwards, Harris, and Piwowar (2007) considered the bid-ask spread to be the most simple and intuitive that does not perform any worse than others through qualitative analysis. Chen et al. (2007) discovered that the bid-ask spread and LOT perform better and exhibit more cross-sectional variation in the yield spread for investment-grade bonds, and LOT performs better for speculative bonds. Wang (2016) tested the Amihud ILLIQ indicator, bid-ask spread and zero-trading-days indicators on the Chinese bond markets, finding evidence of liquidity premium, but only Amihud ILLIQ and the bid-ask spread remained significant in all cases. In recent research on US corporate bond secondary market liquidity, Boyarchenko, Giannone, and Shachar (2018) tested the bid-ask spread and found it to contain better information for the prediction of future liquidity than other measures, including market-wide volatility and market-maker constraints.

2.3 Bond Liquidity Premium

Bond liquidity premium is the additional required return (or cost from the buyer side) for illiquid bonds in comparison to liquid bonds. By definition, liquid bonds can easily be traded and converted to cash without having any effect on market price. Some studies have referred to the calculation of the liquidity premium of a bond as liquidity pricing.

An individual bond has liquidity premium and the entire bond market also has it. However, individual bond liquidity premium is entirely different to the entire market liquidity premium. Individual investors consider the liquidity of individual bonds or specific types of bonds more important for the purpose of conducting transactions. Individual investors treat entire market liquidity as an external liquidity factor, not one that is to be estimated. Conversely, policymakers are more focused on the estimation and prediction of the liquidity of the entire market. These two views are not conflicting and focus on different dimensions from the four dimensions of liquidity definition (Harris, 1991) that were previously mentioned.

This thesis focuses on individual bond liquidity from the perspective of individual investors. Liquidity premium is analysed by separating the liquidity premium components into the yield part and the bid-ask spread part. Further details are given in Chapter 5.

2.3.1 Liquidity Premium of Individual Bonds from the Perspective of Investors

From the perspective of individual investors, people consider whether a positive liquidity premium exists, how it has been affected and how it can be estimated to be an important factor. Researchers have found a positive correlation between returns and liquidity determinates or proxies through an investigation of the relationship between bond liquidity determinants, proxies and returns. Chen et al. (2007) examined corporate bond liquidity and reached the conclusion that more illiquid bonds earn higher yield spreads and liquidity improvement results in yield spread reduction. The liquidity component in the yield spread was found to be associated with each of the three tested liquidity measures, including bid-ask spread, LOT measures and zero-returning days.

A similar conclusion was also reached for yield spread changes, regardless of whether they control the credit rating or macro-economic environment. Lin, Wang, and Wu (2011) expanded the traditional Fama-French (1993) model for bond yield pricing, considering market capital factors, market ratio factors, default factors, maturity factors and liquidity factors. Evidence was found that liquidity risk is critical in expected corporate bond yields. Friewald et al. (2012) studied the time-series changes in US market liquidity and US corporate bond yields using bid-ask spread as the proxy. It was discovered that market liquidity change explains approximately a third of yield changes, while the changes explained 50% in the economic crisis. However, the volume had no explanatory power. It was also found that financial institution bond liquidity is far higher than general corporate bond liquidity. Huang (2012) and He, & Milbradt. (2014) verified the positive feedback between credit risks and liquidity risks of OTC bonds, the role being more obvious for high-yield bonds. Ming and Luo (2016) discovered that the ILLIQ proxy represents liquidity premium and applies not only to the US market but also to the exchanged traded bond market. They also tested the positive feedback between credit and liquidity in the market. In addition, some particular illiquid securities have been found to have higher liquidity sensitivity (Amihud & Mendelson, 2015).

A research group from the Bank for International Settlements (CBIS – CGFS, 2016) examined liquidity changes from a different perspective, whether or not the market maker participates in the market. The report identified significant differences between volume-based liquidity proxies in different markets with different market makers (or brokers). They discovered this liquidity differentiation to only be significant with transaction volume and "*price-based metrics of these costs, such as the bid-ask spreads and liquidity premia, provide little evidence of any significant changes thus far.*" This was the first time the different dimensions of liquidity (breadth and depth) had been compared directly in different markets. In addition, this conclusion supports less consideration of the differences between market makers when the bid-ask spread is used as a proxy for liquidity analysis.

2.3.2 Liquidity Premium of the Entire Market from the Perspective of Policymakers

The perspective of policymakers is more focused on time-series data, answering the question of how the illiquidity of the entire market can be estimated and predicted. They are also interested in the use of liquidity for illustrating systematic risk. A model based on a similar methodology was established as a means of investigating the entire market. Pastor and Stambaugh (2003) tested the liquidity coefficient of the US market, finding it to be auto-correlated. Amihud and Mendelson (2005) found a significant and extremely high coefficient when using an autoregressive model, which demonstrated that a rise or fall in illiquidity will likely persist for a considerable period of time and it is higher in illiquid stocks. These findings helped establish the liquidity-adjusted CAPM model (Acharya & Pedersen, 2005) where an estimated 1.1% annualised premium of systematic liquidity risk existed in the US stock market, 0.6% systematic liquidity risk premium existed in US investment-grade bonds and 1.5% systematic liquidity risk premium existed in US and speculative-grade bonds (Bongaerts, De, & Driessen, 2012). Tang & Yan (2007, 2013) studied the liquidity characteristics of some OTC bonds, compared the liquidity features of the same bonds and reached the conclusion that OTC markets have higher liquidity premiums. However, they did not examine the liquidity pricing of individual bonds. Choudhry (2010) established a composite aggregate liquidity score model with related features for measuring the liquidity of the entire OTC bond market. He introduced 13 features, with issuing amount, bid-ask spread and turnover being the only three features related to transactional activities, while the others were all external environmental or macro-economic features. Adrian, Fleming, Shachar, & Vogt (2017) investigated the bid-ask spread as a liquidity proxy, finding liquidity to change with the market situation and introduced regulations by using a quantitative regression. Some individual bond liquidity researchers have recently focused on how individual liquidity is affected by market liquidity conditions, known as liquidity commonality (Roll & Subrahmanyam, 2000). Although this is related to individual bonds, it is still the view of policymakers.

2.4 Estimating Bond Liquidity Premium: Different Approaches

As was previously mentioned, liquidity premium is the additional return/cost for compensating bond liquidity. As bond liquidity cannot be directly measured, researchers have developed different approaches for the validation and estimation of liquidity premium. These studies and their results have been sorted by the following approaches:

2.4.1 Liquidity Proxy-Based Regression Models

Quantitative finance analysts ran regression predictions with liquidity proxies on liquidity premiums. Through this regression process, they learned how much liquidity proxies can account for liquidity premiums.

There are two common specific approaches when establishing regression models. The first is studying the relationship between different determinant factors and the returns in order to predict market price changes. Researchers have studied bond liquidity through the addition of different types of determinant factors, including liquidity proxies (such as credit rating factors). However, as was mentioned in the liquidity determinants section, as a result of a large number of factors, the regression model can be complicated. Some studies have only established regressions of liquidity premium on a few selected liquidity proxies and controlled other factors. The main focus of these studies has been on an examination of the chosen proxies. Chen et al. (2007) used regressions as a means of testing liquidity proxies in the US corporate bond market, including bid-ask spread, LOT and zero–returns-days indicators. Regarding the condition of credit, sensitivity was found to significantly increase in riskier bonds (Chen et al., 2007; Acharya et al., 2013), thereby verifying the finding in old research that liquidity premium was tiny in T-bills (Elton & Green, 1998; Strebulaev, 2001). Elton et al. and Strebulaev found that the premium, in particular the T- bills premium, was actually from tax rather than liquidity, as was also recognised in the Chinese Treasury notes market (Chang, 2015).

The second uses traditional panel data regression or portfolio matching techniques for more directly estimating liquidity premium. This approach is reliant on the liquidity measurement of a market portfolio and the aforementioned Amihud ILLIQ is one of the most accepted forms of market liquidity measurement. The approach has also been frequently used in combination with the above first approach by adding liquidity factors to the regression model. Lesmond et al. (1999) established a two-factor model (risk-free rate and equity return) for estimating liquidity premium. Martinez, Rubio, and Tapia (2005) established a market portfolio regression estimation, finding bond liquidity pricing to have a negative correlation with the book-market ratio of issuers. Chen et al. (2007) developed Lesmond et al.'s model through the addition of two duration factors, testing three liquidity proxies by regressing liquidity premiums on liquidity determinants that included these proxies. Covitz and Downing (2007) followed a similar approach and found it to have a negative correlation with bond term. Lin et al. (2011) expanded the traditional Fama-French (1993) model for bond yield pricing by regression yield on market capital, market ratio, default ratio, maturity and liquidity factors. Min and Luo (2016) tested the liquidity proxies using a Fama-French model based on bond pricing regression by the Fama-Macbeth process and considered time series influence, obtaining much more rigorous results. At the same time, they observed positive interactions between liquidity and default risk premium.

The regression models can be used to directly examine liquidity characteristics and their significance in different products/markets. However, it has been noticed that regression methods have almost all been used in the study of liquidity of bond market or bond groups. For individual securities, the regression methods (particularly models that contain Amihud ILLIQ proxy) appear more on stocks or exchange-traded bonds, and only limited proxies (e.g. zero-returns-days) are applicable to individual bonds. This can be interpreted as the regression method containing continuous average calculation in a certain period of time. Therefore, adequate price and volume data are required on at least a daily basis. The regression cannot solve bonds that contain too less transactional data, resulting in limitations to the methodology when it is applied to OTC bonds, especially to infrequent individual OTC bonds.

2.4.2 Credit Premium-Based Methods

Some studies have observed the relationship between credit and liquidity. Researchers have estimated or directly collected traditional credit spread quotations as a means of estimating the pure credit part in the spread and the remaining liquidity part. Estimation of the credit part generally requires the use of structural models, including credit intensity models.

At the start of this century, researchers discovered that the traditional default risk model underestimated the credit spread and credit risk only explained part of the results. The reason for the differences was explained as liquidity impacts or liquidity premiums (Huang & Huang, 2003; Eom, Helweg, & Huang 2004; Covitz & Dowing, 2007). A positive correlation was found between liquidity premium and credit premium, and they were both found to have a negative correlation with maturity, whether the Merton (1974) option hypothesis model or the traditional credit expected loss model was used (Ericsson & Renault, 2006). This result was also confirmed and enhanced when time intervals of liquidity and credit events were dependent (Li & Cheng, 2013). In addition, a positive interaction impact between liquidity and credit factors was found, liquidity premium being interpreted by "pure liquidity premium" and "default drove liquidity premium" and vice versa (He & Milbradt, 2014). This resulted in similar conclusions in some papers that used the aforementioned regression model methodology. Chen, Cui, He, & Milbradt (2014) developed a structural credit model for examining how interactions between default and liquidity affect corporate bond pricing, finding that 25% to 30% of the credit spread level changes could be interpreted by default-liquidity interactions for bonds with a credit rating above Aa and 35%-40% for bonds with a credit rating between Aa and Ba.

However, credit premium-based models emphasise default probability and bond price decision. They estimate bond value rather than trading price in a dynamic trading market. These models still require trading activity data and the credit premium-based model cannot be used without taking trading activity data into consideration.

2.4.3 Optimal Solution via Supply-Demand Pricing (Microstructural Models)

These methods are often used in price decision-specific conditions or dynamic trading strategy optimisation. Cetin, Jarrow, and Protter (2004) introduced a random supply curve to the APT pricing model, establishing and testing the function and finding "liquidity required return" to be dependent on asset price and volume. Bao, Pan, and Wang (2011) investigated the spread component on the market maker side, finding liquidity risk to increase the required return.

Vayanos and Weill (2007) investigated the OTC bond market from a different perspective. Based on the hypothesis of the issuing amount being the best liquidity proxy, they believed bond short sales to increase volume and liquidity. They established a bond buy/short-selling model and discovered the short-sale ratio was a better liquidity proxy. They defined the "specialness" of a bond in their study and found specialness premium to be a dominant factor rather than liquidity. They also reached the conclusion that liquidity premium will disappear when no short selling is allowed through their model. They interpreted "specialness" as meaning that a bond may not be widely accepted by traders and become illiquid.

2.4.4 Comparing Solution via Similar/Pair Instruments (Pair Analysis Methods)

Regarding the phenomenon of price differences between similar instruments or bonds in the market, some researchers have studied similar instruments and considered the price difference to be liquidity premium. This is known as the pair analysis approach.

The first typical pair analysis to be performed was on-the-run and off-the-run difference analysis. Amihud and Mendelson (1991) compared short-term bills/medium-term notes with the same maturity, finding practical evidence of newly issued short-term Treasuries being traded relatively cheaply and quickly in the US Treasury market. They proposed the hypothesis that investors value liquidity, which is why the notes-bill liquidity premium exists. Warga (1992) measured and documented the impact age and maturity have on the average premium. Duffie & Kan (1996) and Krishnamurthy (2002) explained the difference by demanding in the repo market, which is the collateral of reverse repo that supports the short selling of bonds. Studies have successfully detected and illustrated the on-the-run/off-the-run difference in various markets and the major part of the differences has been interpreted as liquidity premium, although debate continues regarding the causes of liquidity premium. (Boudouck & Whitelaw, 1993; Fleming, 2003; Guo & Yang, 2006).

When studying bond liquidity premium, the pair analysis approach has not only been applied to on-the-run/off-the-run pairs, but also to other pairs. Ejsing, Grothe, and Grothe (2012) compared Government bonds and Government secured bonds, giving everything else equal. They explained the liquidity difference and found that liquidity changes affect single bond prices. Han and Zhou (2008) compared CDS spreads with the credit spreads of bonds as a means of investigating bond liquidity in the US market and explained liquidity differences. However, Ejsing et al. (2012) argued that the CDS spread is too volatile in comparison to a regular credit status for large countries such as France and Germany, firmly believing that the CDS (particularly sovereign CDS) premium to not be reliable for large countries. Fontana and Scheicher (2016) analysed Euro area sovereign CDS and its relationship with government bonds, successfully linking price spread difference with market liquidity among different markets.

However, there are limitations to pair analysis, the main one of which is that the estimation only works when there is a corresponding instrument of the bond. Taking CDS pairs as an example, there is not a corresponding CDS to every bond. This limits application scenarios for the methods. The more precision required by the pair matching, the harder it is to find such pairs. Another limitation is that data information may be lost when screening pairs, thereby reducing the applicability of this approach.

2.5 OTC Bonds Liquidity Pricing Summaries

Previous studies have explored bond liquidity and liquidity premium, including liquidity definition, liquidity proxies, verification of liquidity premium, correlations between liquidity proxies and returns and interactions between liquidity determinants. Several studies have demonstrated the impact liquidity determinants have on the liquidity premium of bonds as a type of instrument. Some have explained how individual liquidity is affected by market liquidity conditions. However, relatively few studies have attempted to estimate individual OTC bond liquidity premium and there are generally limitations to such estimations. The main limitation is data problems.

The lack of credible data on spreads or quotes has long been a major impediment when analysing liquidity and its impact (Goodhart & O'Hara, 1997; Chen et al., 2007). There are many different types of OTC-traded bonds, they lack fixed participants and are generally without continuous quotation. Although some frequently-traded OTC bonds may have market makers that provide continuous bid-ask quotations, the majority of OTC bonds are less frequently traded. They can only conduct with repeated request-for-quotations from both parties, meaning that spreads or quotes data cannot always be guaranteed for a number of individual bonds that are traded in the OTC market.

There is also a lack of availability of basic information relating to bond issuers. Certain bond issuers, including listed companies or "big-name" companies, may release relevant data and make it easier to estimate the liquidity premium of their bonds. However, this data is generally released at a particular frequency level. For example, financial reports are published quarterly or semi-annually, which means that an information time lag exists, weakening the capabilities of liquidity premium estimation through time. Furthermore, listed companies are only one of many issuers and the majority of other bond issuers do not publish data periodically, thereby limiting the range of bond selection. Liang, Lien, Hao, & Zhang (2017) described it as the parameter uncertainty of corporate bond information, and concluded that parameter uncertainty give rise to the high level of illiquidity in the corporate bond market through testing monthly corporate bond data from 2005-2010.

As has been mentioned in the previous sections, some simplified liquidity proxies, including Roll, LOT and zero-returns-days indicators, have been created and were able to overcome the above problem to a certain extent. However, these liquidity proxies have been found to have less information than traditional proxies, e.g., the bid-ask spread. Therefore, they are not as robust, particularly for investment-grade bonds (Edwards et al., 2007; Chen et al., 2007). Regarding studies that are related to OTC corporate bonds, Chen et al. (2007) examined corporate bond liquidity and liquidity premium, finding significant evidence of three tested liquidity measures. However, no evidence was found regarding the importance of the outstanding amount in explaining bond liquidity. This does not coincide with the daily feel of a bond trader. In addition, the regression on the zero-returning-days indicators requires statistics of past transactions, and this leads to unsuitable for individual bonds that trade extremely infrequently.

In a small number of studies, researchers have explored individual bond liquidity premium. Two approaches have contributed significantly to individual bond liquidity pricing. The first is the portfolio matching approach that was previously mentioned in the regression model section. However, this approach is dependent on the collection of a variety of information on liquidity, particularly basic information about the issuer, including financial indicators and stock prices. This makes the approach more applicable to bonds that are issued by the public and transparent companies (generally listed companies). The second approach is pair analysis. However, this approach only works in particular circumstances: the bond that is to be studied must have a minimum of one correspondent pair instrument, such as a CDS product on the market, with trading information about this instrument also being available. There are limitations to these approaches and the estimation of an individual OTC bond liquidity premium is not yet fully understood.

Chapter 3 Research Question

As was mentioned in the previous section, several studies have explored bond liquidity and liquidity premium and in the area of individual bonds, many have not covered OTC corporate bonds. Previous studies are limited in their estimation of the liquidity premium for an individual OTC-traded bond. Although some approaches and proxies are more applicable for OTC bonds, including market liquidity portfolio matching regression, CDS pair analysis and zero proxies, many limitations remain in terms of practical application. Due to the contradiction that exists between the spread data limitation and the number of factors, it is difficult to use quantitative analysis for directly estimating the liquidity premium of an individual bond.

Therefore, the aim is to study the liquidity and pricing of an OTC-traded bond (particularly an infrequently traded bond). It is hoped that the following question can be answered:

Can a method be found for more easily estimating the liquidity premium for an infrequent OTC-traded bond that reduces the requirements on the number of factors and market data of traditional methods?

Chapter 4 Methodologies and Dataset

4.1 General Methodologies

The aim of this thesis is to study liquidity and pricing for individual OTC corporate bonds by determining an approach that avoids traditional data requirements. Evitable factors that impact bond liquidity on a minimum level are sought, the way they impact the yield and the bid-ask spread is examined and a potential approach is identified for estimating the liquidity premium of an individual bond.

Data scarcity remains the biggest challenge for infrequently traded/quoted OTC bonds. To solve it, this thesis follows the aforementioned comparative analysis process, adopting the bond-pair analysis approach. This is inspired by the cash flow matching approach that is used in financial engineering. Liquidity determinants are classified according to their characteristics and impacting method, bond pairs are constructed according to specific rules, similar impacts are eliminated by comparing similar bonds, and analysis is conducted on the basis of the results of different types of bond pairs. During this process, some determinants of crucial importance are offset between pairs, including factors relating to market rate, financial status and the entire market liquidity environment. In addition, by controlling the changes of other factors, the known liquidity premium of an existing bond with more market data to another with fewer market data can be inferred. Therefore, the data insufficiency obstacle of liquidity impact analysing and pricing is avoided.

In terms of detailed analysis, regression is used for examining the relationship between liquidity and liquidity premium. Liquidity determinants are put into an estimation model to regress the liquidity premiums on the basis of single bonds and regression is run on the liquidity factor differences between two bonds using bond-pair analysis. The factors are analysed individually in the pair analysis process and the difference between bid-ask spread and yield is calculated as a means of interpreting the regression results. In addition, a model construction approach is used for estimating the bond liquidity premium based on the difference between two bonds, particularly for infrequent bonds lacking data.

4.1.1 General Steps

Firstly, bonds with more known transaction quotations, transaction data and liquidity premiums are chosen and the relationship between their liquidity factors and premiums is analysed. During this process, there is reliance on the traditional panel data approach for examining the liquidity determinants of individual bonds. Bond pairs are then introduced. By constructing and analysing bond pairs, the significant factors that determine the difference in liquidity premium among bond pairs are explored. The steps are as follows:

(1) Analysing the liquidity factors and liquidity premium

The process starts with an analysis of the factors, including issuing amount, maturity and cash flow structures, such as the size of coupons and frequency of interest payments. Market environments and trading venues are also considered and analysed. The liquidity premium of bonds, including their yield and bid-ask spread, are also investigated as these two spreads are parts of the liquidity premium.

(2) Analysing the liquidity impact for individual bonds

The liquidity premium is regressed on liquidity factors as a means of examining liquidity factor impacts. The traditional panel data approach that is adopted by the credit groups of issuers is followed for exploring the liquidity determinants of the dataset. The question "what determines bond liquidity premium?" is tested and verified.

(3) Analysing the liquidity impact for bond pairs

Bond pairs are constructed using specific approaches that reduce the number of variables (liquidity factors) to just one. By running a regression of the liquidity premium on the one remaining factor, the question "what is the difference in liquidity premium between two bonds and how is it determined?" is tested and estimated.

In the last, the regression results are explained to provide economic meaning and the inner meaning the findings stand for is interpreted. At the same time, a feasible method of model construction is adopted to help estimate the liquidity premium of infrequent traded OTC bonds without the collection of too much data.

4.1.2 Assumptions

The pair analysis approach in this thesis requires the following assumptions:

(1) Each bond has a unique liquidity characteristic. This is an endogenous property of bonds. The static elements of the bond itself (including credit rating, cash flow structure and maturity) determine the liquidity characteristics. As long as the external factors, including overall market liquidity, do not change, the endogenous liquidity characteristics of the bonds will remain unchanged.

(2) Market segmentation assumptions. Bonds from the same category (such as rating approximation/industry/country differences) are similar. Therefore, they are concentrated on the same market participants (including traders, market makers and intermediaries) for a period of time. This assumption supports that bonds are comparable over time in certain circumstances and supports the method of deducting the same external influences in the form of bond pairs.

(3) The change to the bid-ask spread of the bond happens independently of platforms, market makers and participants. The change is the same when markets and participants change. The preliminary results of this study and those of some previous studies (BIS's report, 2016) support this assumption on a preliminary level. However, further study is required.

(4) The liquidity impacting factors exhibit the similar performance between data-sufficient bonds and data-deficient bonds. Based on this assumption, conclusions from the simple regression of bonds can be extended to other bonds directly. The assumption is also tested in the model.

4.2 Dataset and Preliminary Descriptions

4.2.1 Data Range

US dollar common corporate bonds that were available for public trading (as per Bloomberg) between January 2013 and June 2018 are examined. The bond market had a relatively stable performance during this period. The economic crisis in 2012 caused massive volatility in the bond market, which led to a serious "flight to liquidity" in the market. The majority of bonds with lower ratings had almost no quotations. Boyarchenko et al. (2018) adopted a similar segmented analysis method for their bond market research. Furthermore, for bonds that defaulted during the period, the bond transaction data for 30 days prior to the default is excluded, as was also done by Boyarchenko et al. (2018) for the same reason.

Focus is also on regular bullet bonds (or "plain bonds") at this initial stage of this thesis. These include all government and corporate bills, notes, commercial papers[1] and bonds, with no assets embedded, not callable, puttable or extendable and no change in principles. The types of bond coupon are not limited to zero or specific interest payments and there are no particular restrictions on bond issuers. As plain bonds represent the basic format of all bonds and contain all the common factors that are found in many other bonds, analysing them will help solve the liquidity issues of all other bonds.

4.2.2 Data Collection

The daily quotations of bonds (bid prices, ask prices and yield) are collected manually from Bloomberg Terminal. The daily last "Quoted Bid the Yield to Maturity" (Bloomberg field name, which represents the last bid price of market makers presented in yield) is taken as the bid price and the "Quoted Ask the Yield to Maturity" is taken as the ask price. The difference between the two is the bid-ask spread. Their mean value (mid-price) is calculated as the quoted yield spread. The quoted yield spread is used rather than the traditional valuation yield spread. The advantage of this is that the quoted one more fairly reflects market value and this method can also ensure bid-ask spread data and yield data are comparable as they come from the same source.

The basic information data of bonds, including bond type, coupon type, coupon rate,

[1] Due to the availability of research data, the research objects did not include commercial bank certificates of deposit (CD). Commercial bank certificates of deposit have similar characteristics to zero-coupon bonds, but generally have a longer maturity and are less frequently traded in the US secondary market. However, it should be noted that in some other markets, like in Chinese bond market, the secondary market transactions for CD is incredibly frequent, but this is not the study target of this thesis.

issuing date, maturing date, issuing/outstanding amount, credit rating, issuer type and trading venue, are also collected from Bloomberg Terminal. Data relating to the historical credit change and zero-trading days are obtained from the TRACE database in situations where Bloomberg data is unavailable.

Bonds not recorded in TRACE are excluded. A TRACE tag in Bloomberg helped with the removal of these bonds. Bonds that are missing crucial information (ISIN, issuing amount and maturity date) are removed as this is basic information that is required by the regression. Bonds that were quoted for less than ten days during the five-year period are also excluded in order to establish effective regression models. In addition, in the pair construction process, some bond pairs are also excluded due to specific reasons, for example, bond pairs including the same bonds under 144A and REGS rules.

Table 1: Bond data collection and screening steps

Bonds Data Collection and Screening Steps	**Number of Bonds**	**Number of Issuers**
1. US domestic corporate bonds	*285,930*	
2. US domestic vanilla bonds (2008-2018, excluding bonds with option embed or principle variable, etc.)	*113,813*	
3. US domestic vanilla bonds recorded in TRACE(2008-2018)	*84,469*	*2,806*
*4. Above bonds from 2013-2018 **	*38,415*	*1,926*
5. Above bonds excluding ISIN not applicable ones(3066) & ISIN/BBID duplicated ones(692)	*34,657*	*1,798*
6. Bonds in pairs(Bonds have at least one other correspondent bond issued by the same issuer with at least one similar factor, regardless of the number of trading days)	*13,653*	*467*
*7. Bonds containing at least ten days of quoting data ***	*4,900*	*304*

* This thesis starts by selecting data from 2008 to 2018. Five years of data are ultimately chosen from 2013 due to relatively low data quality and achievability for earlier bonds.

** Most bonds have few quotations.

Period: 2013-2018.6 Source: Bloomberg

4.2.3 Bond Distributions

In terms of industry, government bonds account for 45%, financial institution bonds account for 31%, and the remaining bonds are distributed mostly in industry, consumer retail, public utilities and transportation. The distribution ratio is similar to that of the overall US dollar bond market.

Row Labels	Count of BBID
Government	2217
Financials	1513
Industrials	399
Consumer Discre	268
Utilities	160
Communications	132
Consumer Staple	64
Energy	56
Health Care	48
Technology	35
Materials	8
Grand Total	**4900**

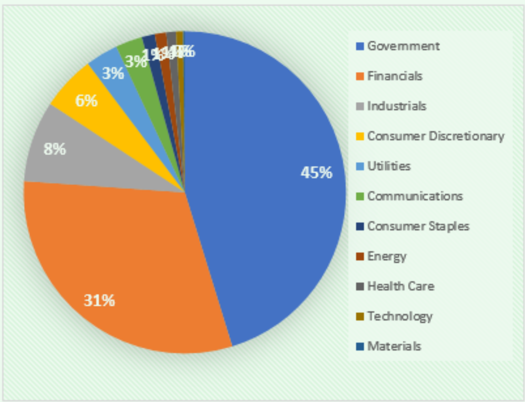


Figure 1: Sector distribution

In terms of bond maturity, they are divided into two years or less, five years or less, ten years or less and ten years or more. It should be noted that the relative number of bonds under two years is limited due to the markets under two years generally being flooded with CDs and government short-term Treasury bills and there are relatively few bond-based fixed-income instruments.

Life	# of bonds
<2Y	397
<5Y	1257
<10Y	1490
>10Y	1756
Grand Total	**4900**

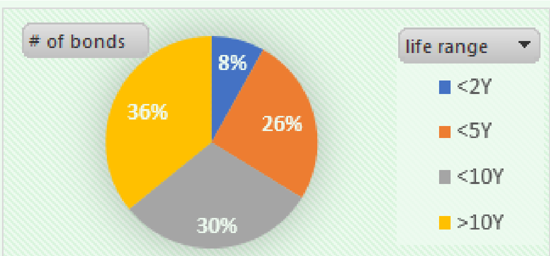


Figure 2: Bond life distribution

Regarding the rating of issuers, 4,065 bonds are investment grade, 88 bonds are high-yield bonds and 747 bonds have no public rating (Table 2). The lowest ratings of Moody’s, S&P and Fitch are selected for the bond and the earlier one between June 2018 and the bond maturity date is used as the rating date. Rating date selection here is only for statistical purposes in this situation. Later, the actual rating of the quoting days is used in detailed pair regressions.

The number of high-yield bonds is relatively small due to the screening criteria used in this thesis: ten or more quotes over the five-year period. High-yield bonds have a tendency to be incredibly risky and no one wants to bid on them. Therefore, quotation data on high-yield bonds are scarce. However, bond-pair regressions (discussed in more detail later) are theoretically less dependent on the credit rating of a bond than traditional regressions on single bonds. Ratings affect the liquidity premium of a single bond, but the difference between two bonds will most likely have little impact as the credit-issuing of both bonds can be offset. Therefore, the initial study is focused on the relationship between bond pairs where the bonds have sufficient data. If the liquidity premium of the bonds is not found to be under significant credit impact, the related conclusions will likely be extended to bonds under other ratings, including high-yield bonds.

Table 2: Bond rating distribution

Category	Rating	Number of Bonds	Category	Rating	Number of Bonds
Investment Grade (4065)	AAA	72	**Non-Investment Grade (88)**	BB+	19
	AA+	1700		BB+	7
	AA	13		BB-	21
	AA-	91		B+	9
	A+	193		B	7
	A	571		B-	6
	A-	289		CCC+	17
	BBB+	651		CCC	2
	BBB+	396	No Rating	NR	747
	BBB-	89			
Total: 4900					

Credit rating value: Lowest rating among Moody's, S&P and Fitch;
Credit rating value date: The earlier date between 15th Jun 2018 and maturity date.

Other characteristics

Regarding bond issuing amounts, 66% of the bonds (3,300) are under US $100 million, but the maximum single issuing amount is enough to be US $6 billion. Regarding bond interest payment frequency, 80% of the bonds (4,013) are the regular semi-annual interest payment, while the others are zero-coupon, monthly and seasonal. Regarding the bond issuing market, 77% of the bonds (3,781) are for the US domestic market, whereby 2,141 are issued separately, and 1,640 are MTN.

4.2.4 Spreads Performance over Time between Groups

The data diagram implies the existence of a distinct correlation between bonds and time. The consistency of the bid-ask spread among different bonds is observed, while the consistency of the yield is of even greater significance (Figure 3). By first visual observation, it is found that bonds have formed several groups based on the different extent of bid-ask spreads and bid-ask spreads are consistent within these groups. Yield consistency is more significant and the simultaneous change implies that bid-ask spreads and yield are affected by the same external factor. It reveals that a potentially strong correlation exists both in yields and bid-ask spreads between bonds. This is generally considered to be external market liquidity.

An uncommon high between-group R^2 shown both in the yield and the bid-ask spread regression reveals the potential consistency of spreads between bonds. It supports the feasibility of pair analysis. It will be discussed later in the single bonds regression section.

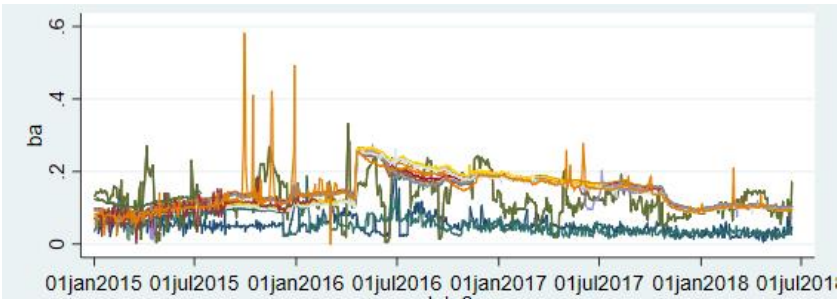
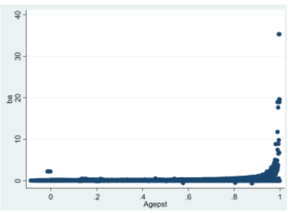
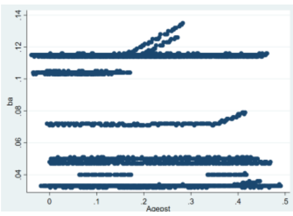


Figure 3: Bid-ask spreads and yield changing tendency for randomly selected bonds

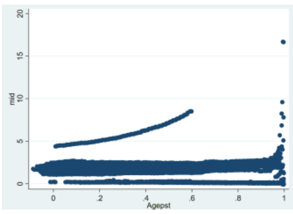
In addition, the following scatter charts (Figure 4) show that in the normal life cycle of a bond, the bid-ask spread and the yield perform differently during different stages. They exhibit relative stability for several long periods and change dramatically near the maturity date. The charts for bonds that are close to maturity (Figure 5) show that both bid-ask spread and yield increase rapidly when the bonds are near maturity for both long-term and short-term bonds. This demonstrates that liquidity will quickly deteriorate when the bonds are close to expiry. At the same time, the yield corresponding to the bid-ask spread shows significant decentralisation. Due to bond maturity, the mean value regression of yield occurs simultaneously.



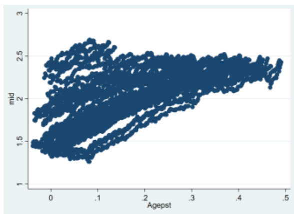
BA vs Lifecycle (all bonds within 2y)



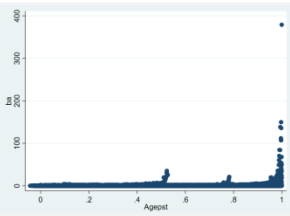
BA vs Lifecycle (BBB above bonds within 2y)



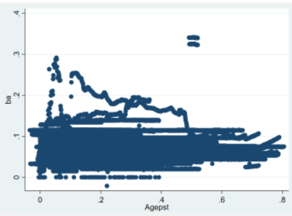
Yield vs Lifecycle (all bonds within 2y)



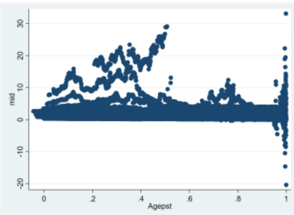
Yield vs Lifecycle (BBB above bonds within 2y)



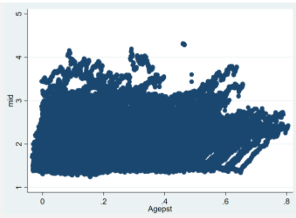
BA vs Lifecycle (all bonds 2y- 5y)



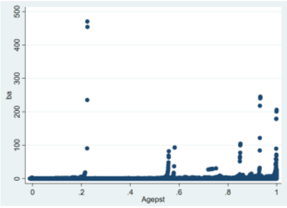
BA vs Lifecycle (BBB above bonds 2y- 5y)



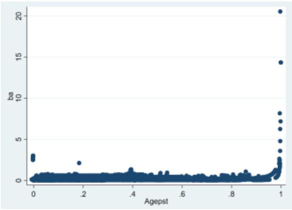
Yield vs Lifecycle (all bonds 2y- 5y)



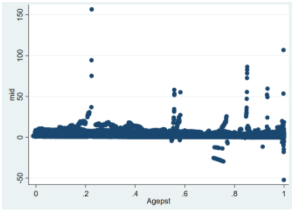
Yield vs Lifecycle (BBB above bonds 2y- 5y)



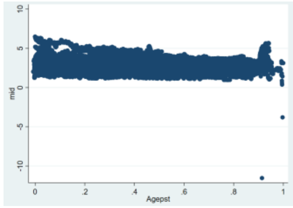
BA vs Lifecycle (all bonds 5y-10y)



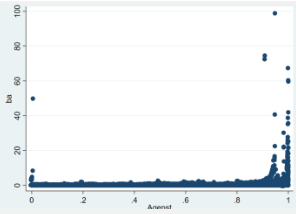
BA vs Lifecycle (BBB above bonds 5y-10y)



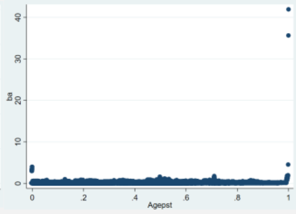
Yield vs Lifecycle (all bonds 5y-10y)



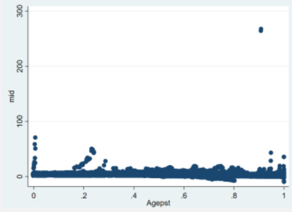
Yield vs Lifecycle (BBB above bonds 5y-10y)



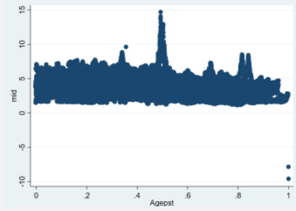
BA vs Lifecycle (all bonds above 10y)



BA vs Lifecycle (BBB above bonds above 10y)

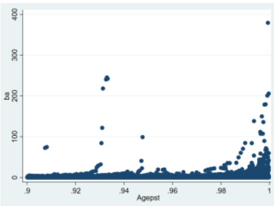


Yield vs Lifecycle (all bonds above 10y)

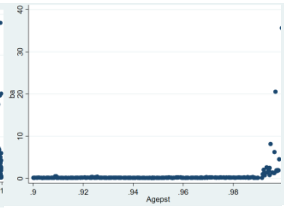


Yield vs Lifecycle (BBB above bonds above 10y)

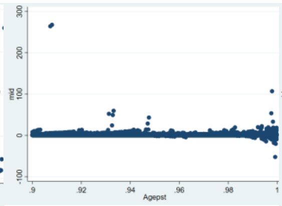
Figure 4: Changes of bid-ask spreads and yield over the entire life cycle of short-, medium- and long-term bonds



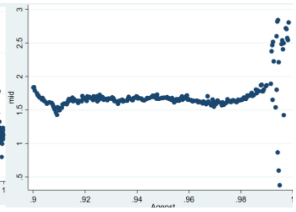
BA change when bonds are due



BA change when bonds are due (BBB above)

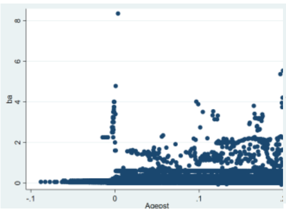


Yield change when bonds are due

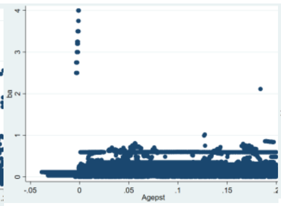


Yield change when bonds are due (BBB above)

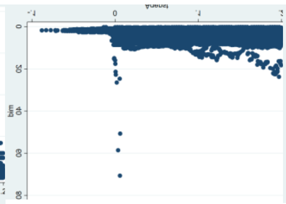
Figure 5: Bid-ask spreads and yield trends before and after all bond maturities



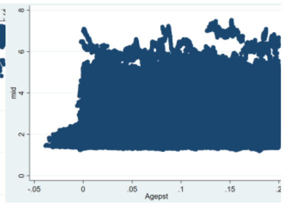
BA change when bonds are issued



BA change when bonds are issued (BBB above)



Yield change when bonds are issued



Yield change when bonds are issued (BBB above)

Figure 6: Bid-ask spreads and yield trends before and after the bond issuing date

In the above charts, BA refers to the bid-ask spread and AGEPCT refers to the percentage of elapsed time of the bond to the bond length. For example, with a 12-year bond, AGEPCT= 0.5 in the 6th year.

4.3 Regression Model Selection

This thesis adopts cross-sectional regression and panel data regression. Cross-sectional regression is used for testing the impacts of the liquidity factors and panel data regression is used when estimating the liquidity premium difference with pair models. Time series regression is not considered as the research object is not the prediction of future liquidity premium level from the past liquidity premium level of a bond. The expected liquidity premium level for a bond lacking data will be measured, so the analysis of traditional time series is of relatively little significance.

The bid-ask spread and the yield are not always available, as OTC bonds are not frequently quoted or traded. It means the panel data are highly unbalanced. However, it is not an extremely serious problem as yield and bid-ask spread are dependent variables in the pair analysis model.

In addition, many variables in the model used in this thesis are given and unchanged. Many individual liquidity determinants, including bond issuing amount and the cash flow structure of bonds, are time-invariant. A few factors, such as ratings, issuer financial indicators and market factors, are time-varying and change regularly or irregularly over time. In a traditional panel regression, these support a fixed-effects model. However, traditional regression forecasts the liquidity premium for bonds with available information and historical data, which are not easy to collect for infrequent traded OTC bonds. Therefore, it inhibits reliance on a traditional fixed-effect panel model for research. Instead, it is a greater likelihood of a random effect model focusing on the difference between groups than time.

Chapter 5 Model Construction

5.1 Liquidity Determinants

Many factors can be liquidity determinants. This thesis analyses liquidity determinants using the same approach as Chen et al. (2007). The determinants for bid-ask spread selected from those of Garbade and Silber (1979), Sarig and Warga (1989), Chakravarty and Sarkar (1999), Stoll (2000), Schultz (2001) and Brandt and Kavajecz (2004), while the determinants for the yield are selected from those of Elton, Gruber, Agrawal, and Mann (2001), Campbell and Taksler (2003) and Lin et al. (2011). In addition, two empirical factors are introduced – bond age and trading venue – as a means of improving the regression.

In order to simplify the analysis, these factors are divided into three categories: the inherent characteristics, market environment and trading venue environment of the individual bonds.

(1) Inherent Characteristics of Individual Bonds

This category includes factors that cause price fluctuations due to individual characteristics of bonds and includes the following:

- **Individual Structure Specialness.** This refers to factors that determine the cash flow structure of a bond, including bond life, maturity, frequency and coupon rate.

- **Issuer Specialness.** This includes unique issuer information, such as sectors and the credit rating of the issuer or bond. It also includes financial indicators that determine the liquidity premium, such as operating income-to-sales, debt-to-asset and debt-to-capital, as mentioned by Elton et al. (2001).

- **Size Specialness.** This refers to the issuing amount or outstanding amount. The issuing/outstanding amount is one of the essential liquidity determinants from the practical view of traders. For bonds without options embedded, the issuing amount generally equals the outstanding amount.

(2) Market Environments

This category includes proxies relating to external market liquidity that may impact the price of individual bonds, such as overall market liquidity proxies, size of the entire market and risk-free market rates. The entire market liquidity proxies can be LOT measure, Amihud ILLIQ, or average or weighted market bid-ask spread. This category also includes any risk-free market rates that can affect individual liquidity.

(3) Trading Venue Environments

This category refers to the particularity of the external trading environment of a bond, including trading platform, number of participants and market makers.

The main factors and their characteristics over time can be seen in the following table:

Table 3: Bond features analysis

In this table, the related factors are classified according to the type of feature. Whether they are time-varying or time-invariant is also analysed. Some factors constantly or volatilely change over time, while others remain stable. One column shows whether they can be eliminated by constructing pairs. This determines whether they should be put in the pair model.

Categories		**Factors**	**Whether time-varying**	**Whether sudden change**	**Whether pairs can eliminate it**	**Proxies in this thesis**
Inherent Characteristics	Individual Features	Cash Flow Structure (*Frequency/ Coupon, etc.)*	*No*		*No*	The portion of each Cash flow structure category
		Age (Bond age percentage)	*Stable change*	*No*	*No*	Using a comparing US with the same maturity, measured by the percentage
		Bond life (bond length)	*No*		*Yes*	Bond length
		Maturity	*No*		*Yes*	Remaining days to maturity
		Principle	*No*			Principle
	Issuer Features	Credit	*No*	*Yes*	*Yes*	Rating when issuing, compared by groups
		Financial Indicators	*volatile*	*Yes*	*Yes*	Debt/Operating Income/Asset/Capital/
	Size Feature	Issuing/	*No*		*Yes*	Issuing amount

		Outstanding Amount			*(conditional)*	(Outstanding amount if applicable)
Trading Venue Environment		Venue	*No*	*Yes*	*Yes (conditional)*	Number of public trading venues
		Issuing Market	*No*	*Yes*	*Yes (conditional)*	Issuing market type
Market Liquidity Environment		Market Liquidity	*volatile*		*Yes*	Market average bid-ask spread
		Market Volatility	*volatile*		*Yes*	The standard deviation of market return
		Risk-free rates	*volatile*		*Yes*	Treasury rates, etc.

5.2 Liquidity Premiums

As was mentioned in the literature review section, bond liquidity premium is an additional required return (or cost from the buyer side) for illiquid bonds in comparison to liquid ones. The required return (or cost) is partly reflected in the yield and partly in the bid-ask spread. In this thesis, analysis starts by separating the components of the liquidity premium according to the liquidity source.

5.2.1 Liquidity Premium in Yields

The first part of liquidity premium is the additional yield between a bond and a perfect fully liquid bond. If it is assumed that the bond and the perfect fully liquid bond are from the same market, the additional yield will be only caused by the characteristics of the bond. This can also be called "individual liquidity premium". For example, if a plain bond is calculated into one index, people may suppose it will be more popular in future, thereby creating more liquidity. Therefore, the bond price will go higher (and the yield will be lower), demonstrating a liquidity premium compared to other, similar, non-index bonds.

Individual liquidity premiums are expressed as the actual yield of a bond minus the market risk and credit risk components. In this study, individual liquidity premium is also calculated based on the above form, namely:

$$\boldsymbol{Y_{mid}} = \boldsymbol{Y_{liquidity}} + \boldsymbol{Y_{valuation}}$$

$$= \boldsymbol{Y_{liquidity}} + (\boldsymbol{Y_{interest}} + \boldsymbol{Y_{credit}})$$

Y_{mid} is the actual yield of this bond. The mid-yield that is provided by market makers is used rather than the traditional valuation yield to make it additive to the bid-ask spread. This can be directly observed. $Y_{interest}$ is the market risk premium, which is a known factor that can be calculated from a risk-free rate. Y_{credit} is the credit risk premium for the credit risk factor of the issuer, which can be calculated for the specific bond by default probability. This thesis considers it to be a known element. $Y_{liquidity}$ is the above individual liquidity premium part of the yield, one of the study objects of this thesis. This is also a function of the individual features of the bond (including cash flow structure, issuer and size.) and some external environments (including trading venue and market liquidity)

The following regression with the yield spread as the dependent variable and the various yield spread determinants as independent variables can also be made. This yield regression model is the same as that used by Chen et al. (2007). Details are provided later in the regression sections.

$$\begin{aligned} \boldsymbol{Y_{mid}} = \boldsymbol{\alpha} + \boldsymbol{\beta_1 Liqduity} + \boldsymbol{\beta_2 Maturity} + \boldsymbol{\beta_3 Issuing\ Amount} + \boldsymbol{\beta_4 Coupon} \\ + \boldsymbol{\beta_5 Bond\ rating} + \boldsymbol{\beta_6}\frac{\boldsymbol{Operating\ Income}}{\boldsymbol{Sales}} + \boldsymbol{\beta_7}\frac{\boldsymbol{Debt}}{\boldsymbol{Assets}} \\ + \boldsymbol{\beta_8}\frac{\boldsymbol{Debt}}{\boldsymbol{Cap}} + \boldsymbol{\gamma_1 Volatility} + \boldsymbol{\gamma_2 TreasuryRate} \\ + \boldsymbol{\gamma_3 EuroDollar} + \boldsymbol{\gamma_4 TreasuryRate_{10Yr-2Yr}} + \boldsymbol{\delta_1 Venue} + \boldsymbol{\varepsilon} \end{aligned}$$

It is assumed that individual liquidity premium can be explained by the buyer/seller expecting a bond to have specific unique properties, which makes it easier or harder to buy and sell in the future. From a different perspective, individual liquidity premium does not change when the liquidity characteristics of a bond remain unchanged.

5.2.2 Liquidity Premium in Bid-Ask Spreads

The bid-ask spread is a direct initial manifestation of the liquidity premium of a corporate bond when it is traded in the market. It is also the traditional liquidity proxy. The bid-ask is only adopted from quotations, which are provided actively by market makers and can be directly traded, known as “firm quotes”. Although some market participants provide non-tradable bid and ask prices, this thesis does not use such data. In addition, the counterparty credit spreads that are embedded in the bond quotations are ignored as the bond transactions are generally T+1 settled.

From the perspective of investors, the difference between buying and selling prices is the immediate cost of a bond to complete a small amount of buying and selling transactions. The bid-ask spread can be considered to be the trading difficulty of the bond in the current market status. The transaction problem is related to the characteristics of the market and bonds, but it is closely related to the market.

From the perspective of market makers, the difference is decided by their costs. Microstructure theory states that these costs are collections of the order processing costs that are incurred by market liquidity providers, inventory holding costs and adverse selection costs. For two plain bonds that are issued by the same issuer and quoted by the same market maker, internal processing cost and reverse selection cost will not change significantly in a short period of time. Both costs can be treated as constant cross-sectional factors that are only related to the market maker. Inventory holding cost is related to the market maker and bond feature. If focus is placed on the difference between the bid-ask spreads of two bonds, the constant matters of market makers can be eliminated.

$$\mathbf{0.5} * \boldsymbol{BA}_{\boldsymbol{market\ maker\ j, bond\ i}} = \boldsymbol{Cost}_{\boldsymbol{processing}} + \boldsymbol{Cost}_{\boldsymbol{adverse\ selection}} + \boldsymbol{Cost}_{\boldsymbol{inventory}}$$

$$= \boldsymbol{C}_{\boldsymbol{processing}} + \boldsymbol{C}_{\boldsymbol{adverse\ selection}} + \boldsymbol{f}\ (\boldsymbol{Bond\ i\ features})$$

$$= \boldsymbol{C}_{\boldsymbol{market\ maker\ j}} + \boldsymbol{f}\ (\boldsymbol{Bond\ i\ features})$$

c is a constant

Transactional Liquidity Premium

The difference between final transaction price and mid-price at the same time of the transaction contributes in part to the liquidity premium. It illustrates the phenomenon that a single transaction can affect market price, which causes price movement. This part of the liquidity premium is only shown when the transaction occurs.

Click-down is the dominant transaction method in the exchange-traded market and this allows the direct confirmation of a transaction when transaction volume is small. This difference is zero. When the transaction volume must generate an outstanding order, the difference is the change in price at the time of the transaction. This is also the essence of Amihud's "price shock model".

Price movement is more important for transactions with a large notional amount. If it is assumed that transactions conclude independently to an external environment or market change, the static factors of a bond and the external environment do not change. Therefore, the liquidity characteristic and individual liquidity premium do not change and the mid-price remains stable. The final transaction price change is the trading spread. Therefore, liquidity pricing can be studied in this thesis without considering the volume impact.

$$\boldsymbol{Y}_{\boldsymbol{liquidity}(\boldsymbol{post-tade})} = \boldsymbol{Y}_{\boldsymbol{liquidity}(\boldsymbol{pre-tade})}$$

$$\boldsymbol{BA}_{\boldsymbol{liquidity}(\boldsymbol{post-trade})} = \boldsymbol{BA}_{\boldsymbol{liquidity}(\boldsymbol{pre-trade})} + \Delta\boldsymbol{BA}$$

$$\Delta\boldsymbol{BA} = \boldsymbol{f}_{(\boldsymbol{trade\ volume})}$$

5.3 Regression for Single Bonds

Following the qualitative analysis of liquidity determinants and premium, quantitative and regression analysis of the two factors are performed. In order to understand the performance of bond determinants, the traditional panel data regression method is used. The bid-ask spread and yield spread, which are the two main parts of the liquidity premium of a single bond, are analysed. The method that was used by Chen et al. (2007) is adopted in combination with the techniques and experience of Lin et al. (2011) and the regression model is adjusted slightly.

5.3.1 Regression for Yield

For the data selection to be comparable, outstanding plain US bonds (2013-2018) are adopted as stated above. There were 34,657 outstanding bonds in the period 2013-2018 from 1,330 issuers. However, not all issuers were listed companies. As financial data relating to the issuers (including operating income, sales, assets and debt) and the corresponding equity volatility will be used in the regression model, listed companies are analysed to ensure relatively accurate data is obtained from the financial reports. This results in two problems: the data scale is small and the chosen bonds are those with high ratings as most listed companies are high-rated issuers. 615 of the 1,330 listed companies are chosen after comparison. At the same time, to maintain continuity with the pair analysis, regression is performed on the bonds that need pair analysis. 1,634 bonds and 115 issuers are ultimately involved.

As the model requires data from financial reports, quarterly data are used for regression. Although the data frequency is different to that of pair analysis, it should not prevent any understanding of the possible liquidity characteristics and determinants of a bond.

$$\begin{aligned} Y_{mid} = \alpha &+ \beta_1 Liquidity + \beta_2 Age + \beta_3 Maturity + \beta_4 Issuing\ Amount \\ &+ \beta_5 Coupon + \beta_6 Bond\ rating + \beta_7 \frac{Operating\ Income}{Sales} \\ &+ \beta_8 \frac{Debt}{Assets} + \beta_9 \frac{Debt}{Cap} + \gamma_1 Volatility + \gamma_2 TreasuryRate \\ &+ \gamma_3 TreasuryRate_{10Yr-2Yr} + \delta_4 Venue + \varepsilon \end{aligned}$$

The model follows the approach that was adopted by Chen et al. (2007) with a few adjustments made. The liquidity determinants are mainly based on those of Elton et al. (2001), Campbell and Taksler (2003) and Lin et al. (2011). There are three most important adjustments: the first is adding the variable bond age to better show the potential on-the-run/off-the-run difference. The second is adding the trading venue to examine whether there are impacts on bond trading platforms. The third is removing the euro-dollar factor, as only bonds in the US markets are considered. Regarding the liquidity factor, a quarterly average bid-ask spread is used. It should be noted that the research of Chen et al. (2007) implied that in bonds of investment grade, the R2 of the bid-ask spread is slightly higher than that of zero percentage and LOTs. The above data statistics imply that the bonds in this study are mainly investment grade.

The regression tests and verifies the yield determinants again, including maturity, outstanding amount, financial indicators and market situation. The conclusions are consistent with those of previous studies. The fact that the above data runs significantly in bonds with a high rating is cause for concern, but the significance of the data decreases following the decrease in the credit ratings of the bonds. The explanatory power, within-group R^2, is relatively high (0.22) for such an estimation. Also, an uncommon high between-group R^2 (0.65) reveals the potential consistency of yield between bonds. It supports the feasibility of pair analysis discussed later.

Regarding several specific determinants, Table 4 shows the detailed performance of each factor. The first row shows the significance of the bid-ask spread as the liquidity proxy. The following two rows reveal the bond length and age to have significant impacts on the yield. With bond length, it is noticed that it has a positive correlation with yield. Conversely, bond age (the ratio of the elapsed days to the total bond length) has a negative correlation. This verifies that on-the-run/off-the-run difference exists in these corporate bonds. The significance exists in almost every bond grade. Regarding financial indicators, as shown in Row 6-8, their significance is generally only observed in investment-grade bonds, while there is no evidence in speculative bonds. Regarding equity volatility, this has a positive correlation with bond yield and is the only factor that does not significantly weaken as bond credit rating decreases. Equity volatility should be of great importance for the interpretation of bond yield, but it is only obtainable from listed companies.

The amount factor of investment grade bonds is significant in the regression model used in this study, which is different to the conclusion reached by Chen et al. (2007). This is because different yield collection methods were used in the two studies. The market maker quotation implied yield from Bloomberg is adopted in this study rather than valuation yield or actual return. This potentially indicates that implied yield represents more liquidity information of trading in the secondary market.

Table 4: Yield spread determinants and liquidity tests

As was previously mentioned, yield is determined by the inherent characteristics of a bond, including coupon, age, length, outstanding amount, credit rating and the financial data of issuers. All these factors are taken into consideration. BA is the quarterly-average bid-ask spread, stands for the liquidity proxy. Maturity, age, amount and coupon are all bond features. The three financial indicators are collected from quarterly financial reports. US Treasury rates are directly from Bloomberg. Standard deviations of equity return are adopted as the volatility indicator. For the venue, the number of quoting market makers collected from Bloomberg is used. In addition, regressions are made by different credit groups and for simplicity, there are seven different rating groups.

	Investment grade bonds			High-yield bonds			Unrated	**All**
	AAA&AA	A	BBB	BB-	B	C&below	bonds	**bonds**
ba_avg	0.222***	-0.251***	0.170***	0.496	0.813***	0.920	-2.498***	**0.125*** **
	(7.96)	(-6.70)	(4.28)	(1.36)	(5.22)	(0.65)	(-6.16)	**(6.29)**
length	0.000238***	0.000185***	0.000237***	0.000197	0.000261***	-0.00113	0.0136**	**0.000181*** **
	(21.54)	(14.22)	(33.57)	(1.82)	(3.68)	(-0.38)	(3.21)	**(25.08)**
age	-2.165***	-0.769***	-1.493***	-1.308*	-4.552***	3.296	-0.986	**-1.592*** **
	(-27.10)	(-7.50)	(-25.41)	(-2.04)	(-7.82)	(1.11)	(-1.50)	**(-30.84)**
amount	-1.92e-10***	8.41e-11	-1.20e-10***	1.58e-11	1.06e-10	-4.67e-10	-1.34e-09***	**-1.22e-10*** **
	(-5.78)	(1.56)	(-5.92)	(0.02)	(0.17)	(-0.59)	(-3.74)	**(-4.89)**
couponrate	0.0585**	0.163***	0.0723***	1.032**	0.217	0	-0.178***	**0.157*** **
	(2.79)	(5.30)	(4.52)	(2.89)	(1.93)	(.)	(-3.29)	**(9.89)**
incomesales	-0.0150***	0.00234	0.00227***	0.0121	-0.00835	0	-0.0408***	**-0.00172** **
	(-3.75)	(1.25)	(4.09)	(0.81)	(-1.89)	(.)	(-4.37)	**(-2.74)**
debtasset	-0.00352	0.0154***	-0.00111	-0.0343*	0.00121	-0.106	0.0610***	**0.000908**

	(-1.20)	(7.07)	(-1.59)	(-2.38)	(0.30)	(-1.12)	(6.00)	**(1.32)**
debtcap	-0.00233*	-0.00182	0.00284**	0.0535	-0.00128	-0.000553	-0.151**	**-0.000479****
	(-2.18)	(-1.62)	(3.00)	(1.24)	(-1.95)	(-1.23)	(-2.86)	**(-3.03)**
vol_equity	0.0864***	0.355***	0.360***	0.329***	0.677***	0.626***	0.110	**0.316*****
	(4.49)	(14.18)	(35.85)	(3.51)	(10.33)	(4.88)	(1.66)	**(34.42)**
us1y	0.760***	0.616***	0.705***	0.207	0.0407	-1.395*	0.901***	**0.690*****
	(36.02)	(25.38)	(61.20)	(1.68)	(0.23)	(-2.08)	(16.42)	**(62.41)**
us10y2y	0.629***	0.787***	0.744***	0.406**	-0.288	-1.106*	1.189***	**0.731*****
	(20.96)	(22.88)	(47.33)	(2.99)	(-1.44)	(-2.26)	(18.12)	**(47.94)**
venue	0.0288	0.315***	0.241***	0.820	1.183**	-0.0688	0	**0.200*****
	(0.90)	(5.52)	(10.86)	(1.20)	(2.69)	(-0.32)	(.)	**(8.67)**
rating	0.579***	0.188***	0.225***	-0.129	0.793*	1.380*	-9.196**	**0.0339*****
	(3.92)	(3.71)	(7.16)	(-0.11)	(2.08)	(2.08)	(-3.18)	**(4.69)**
length	0	0	0	0	0	0	0	**0**
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	**(.)**
_cons	-0.160	-2.787***	-1.618***	-0.453	-7.777	0	0	**-0.0423**
	(-0.54)	(-8.61)	(-6.07)	(-0.03)	(-1.48)	(.)	(.)	**(-0.43)**
N	5543	4241	7177	248	421	69	285	**17984**
r2_w	0.190	0.186	0.478	0.169	0.295	0.570	0.744	**0.224**
r2_b	0.814	0.666	0.826	0.728	0.737	0.982	0.976	**0.653**

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.3.2 Regression for Bid-Ask Spread

The regression for bid-ask spread is similar and the equation is as follows:

$$\begin{aligned} \boldsymbol{BA\ spread} &= \boldsymbol{\gamma + \mu_1 Liquidity + \mu_2 Age + \mu_3 Maturity} \\ &\quad \boldsymbol{+ \mu_4 Issuing\ Amount + \mu_1 Bond\ Volatility + \mu_2 Bond\ Rating} \\ &\quad \boldsymbol{+ \varepsilon} \end{aligned}$$

For the liquidity factor, as the bid-ask spread is the dependent variable, quarterly average bid-ask spread is no longer suitable to be the liquidity proxy. Quarterly zero-quoting days is used for analysis. However, liquidity proxies are not studied in this regression, but the impacting factors are verified as a means of determining the bid-ask spread using single bonds regression, so the choice of liquidity factor is not an issue. Bond volatility uses the yield standard deviation.

The regression results also verify the determinants of the bid-ask spread, including maturity, outstanding amount and age. Similar to what was observed in the regression for yield, these factors are significant for high-rating bonds and the significance also decreases following the decrease in the credit ratings of bonds.

Regarding several specific determinants, bond length no longer affects bid-ask spread size, but bond age is still correlated: the older the bond, the bigger the bid-ask spread. Bond volatility also has a direct correlation with bid-ask spread, which is consistent with common sense. However, the impact of the amount factor is not found to be significant for all rating groups. Although this accords with the findings of Chen et al. (2007), it is still quite different to the experience of bond traders. Therefore, the investigation will continue in the following pair analysis.

Table 5: Bid-ask spread determinants and liquidity tests

	Investment grade bonds			High-yield bonds			Unrated	**All**
	AAA&AA	A	BBB	BB-	B	C&below	bonds	**bonds**
zerodays	0.0952	0.197***	0.117***	-0.171	0.183	0.00971	0.0502	**0.0904****
	(1.14)	(9.65)	(3.78)	(-0.81)	(0.35)	(0.25)	(1.79)	**(2.71)**
length	-0.0000005	-0.00000300	-0.00000494*	0.000000764	-0.0000300	0.0000950	-0.0000340	**-0.0000012***
	(-1.29)	(-1.26)	(-2.32)	(0.05)	(-0.67)	(1.23)	(-0.11)	**(-2.32)**
age	0.314***	0.270***	0.248***	0.729***	1.217**	-0.132	0.404***	**0.348*****
	(3.76)	(13.21)	(10.02)	(3.85)	(2.67)	(-1.92)	(7.62)	**(10.45)**
amount	1.92e-11	-6.56e-11***	-1.43e-11	3.55e-10*	5.94e-10	9.48e-11*	-1.95e-10***	**-5.87e-12**
	(0.29)	(-4.79)	(-1.66)	(2.16)	(1.08)	(2.26)	(-6.49)	**(-0.28)**
vol_yield	0.660***	1.326***	-6.54e-11	0.425	1.616***	0.0996**	0.0324	**-3.83e-11**
	(13.33)	(90.18)	(-0.14)	(1.64)	(5.92)	(3.08)	(1.12)	**(-0.05)**
rating	0.124	-0.00898	-0.0413***	-0.164	0.149	-0.0274	0.0115	**0.000520**
	(0.81)	(-0.84)	(-3.81)	(-0.78)	(0.46)	(-1.62)	(0.06)	**(0.10)**
_cons	-0.287	-0.00290	0.401***	1.509	-2.685	0	0	**0.0996***
	(-0.61)	(-0.05)	(4.79)	(0.69)	(-0.61)	(.)	(.)	**(2.37)**
N	5543	4281	7287	261	421	85	285	**18163**
r2_w	0.0281	0.679	0.00782	0.0346	0.0759	0.267	0.236	**0.00361**
r2_b	0.179	0.419	0.216	0.246	0.382	0.558	0.0376	**0.0556**

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.3.3 Summary of Single Bonds Regression

Following the panel data regression for bonds, the relationship between liquidity determinants and premiums is preliminarily explored. Based on an existing yield pricing model and a bid-ask spread model, the impacts of various factors on the liquidity premium are verified with the bond dataset. Individual cash flow, amount, financial status of the bond issuer, volatility of equities/stocks and bond market environment are all found to be decisive for the yield. There is also evidence that the bid-ask spread, as the liquidity proxy, has significant impacts on the yield. The amount also plays a decisive role. Maturity and bond age are found to be significant for the bid-ask spread. The significance of these factors in high-rating bonds is stronger than those in low-rating bonds for both yield and bid-ask spread, which potentially indicates that low-rating bonds may be affected by more unknown factors and that liquidity premium is unstable.

The single bonds regression process inevitably has some restrictions, making the estimation of the liquidity of individual bonds more challenging. Only bonds that were issued by listed companies can be involved due to the need for data regarding the financial status of issuers and their equity volatility. In addition, yield estimation can only be conducted on a quarterly basis as this is the regularity with which financial reports are published. For observable bond individuals, many factors and information are still required for the regression. Therefore, the pair analysis approach is adopted in the following to avoid the need for factors that are difficult to obtain, which will simplify the model.

5.4 Regression for Bond Pairs

As has been mentioned in previous sections, to navigate the data scarcity obstacle, the pair analysis method is adopted. By analysing whether or not each determinant factor can be offset in pairs (Table 3), bond pairs are constructed according to specific rules. This approach eliminates similar impacting factors from different bonds and helps conduct the following bond-pair regression.

5.4.1 Model Construction Methodology

The aforementioned yield pricing model is used and to clarify the analysis, the equation is simplified by grouping the determining factors.

$$\begin{aligned} Y_{mid} = a + b_1 Liquidity + b_2 CSFS + b_3 Amount + b_4 Age + b_5 Maturity \\ + b_6 Rating + b_7 Financial\ indicators \\ + b_8\ Equity\ Volatility + b_9\ Market\ Rates + b_{10} Venue + \varepsilon \end{aligned}$$

$$\begin{aligned} \boldsymbol{BA\ spread} \quad &= \boldsymbol{c + d_1 Liquidity + d_2 CSFS + d_3 Amount + d_4 Age} \\ &\quad \boldsymbol{+ d_5 Maturity + d_6 Rating + d_7\ Bond\ Volatility + \varepsilon} \end{aligned}$$

Liquidity is still the aforementioned liquidity proxy. It can be any liquidity proxy, including bid-ask spreads, zero and ILLIQ. CSFS stands for the cash flow structure of a bond, including coupon and frequency. Age and maturity are the age and the remaining life of a bond, respectively. Volatility is the equity volatility previously mentioned. Financial indicators and market rates represent a series of relevant indicators in the above equation. Equity volatility and bond volatility are the standard deviation statistics of equities and bonds, respectively.

When there are two bonds (bond a and bond b), the yield difference between the two bonds can be calculated by simple subtraction:

$$\begin{aligned} &\boldsymbol{Yield_a - Yield_b} \\ &\quad = \boldsymbol{A + b_1(Liquidity_a - Liquidity_b) + b_2(CSFS_a - CSFS_a)} \\ &\quad \boldsymbol{+ b_3(Amount_a - Amount_b) + b_4(Maturity_a - Maturity_b)} \\ &\quad \boldsymbol{+ b_5(Age_a - Age_b) + b_6(Rating_a - Rating_b)} \\ &\quad \boldsymbol{+ b_7(Financial_a - Financial_b) + b_8(Vol_a - Vol_b)} \\ &\quad \boldsymbol{+ b_9(Market_a - Market_b) + b_{10}(Venue_a - Venue_b) + \varepsilon} \end{aligned}$$

$$\begin{aligned} \mathbf{BA}_a - \mathbf{BA}_b &= \boldsymbol{C + d_1(Liquidity_a - Liquidity_b) + d_2(CSFS_a - CSFS_a)} \\ &\quad \boldsymbol{+ d_3(Amount_a - Amount_b) + d_4(Age_a - Age_b)} \\ &\quad \boldsymbol{+ d_5(Maturity_a - Maturity_b) + d_6(Rating_a - Rating_b)} \\ &\quad \boldsymbol{+ d_7(Vol_a - Vol_b) + \varepsilon} \end{aligned}$$

Market risk-free ratings are offset with each other. If both bonds are issued by the same issuer, it is possible to offset the financial, rating and volatility factors. In this case, the yield liquidity premium difference is simply the difference of yield, while the difference of bid-ask spread liquidity premium is simply the difference of bid-ask spread.

$$\begin{aligned} &\mathbf{Yield}_a - \mathbf{Yield}_b \\ &\quad = \boldsymbol{b'_1 * (Liquidity_a - Liquidity_b)} \\ &\quad \boldsymbol{+ b'_3(Amount_a - Amount_b) + b'_4(Age_a - Age_b)} \\ &\quad \boldsymbol{+ b'_5(Maturity_a - Maturity_b) + b'_{10}(Venue_a - Venue_b) + \varepsilon} \end{aligned}$$

$$\begin{aligned} \mathbf{BA}_a - \mathbf{BA}_b &= \boldsymbol{d'_1(Liquidity_a - Liquidity_b) + d'_2(CSFS_a - CSFS_a)} \\ &\quad \boldsymbol{+ d'_3(Amount_a - Amount_b) + d'_4(Age_a - Age_b)} \\ &\quad \boldsymbol{+ d_{\prime 5}(Maturity_a - Maturity_b) + d_7(Vol_a - Vol_b) + \varepsilon} \end{aligned}$$

During the actual process, as there may be a nonlinear relationship between factors and the dependent variable, the regression model adopts subtraction (difference) and division (ratio) as a means of exploring the coefficient relationship. It is also observed whether any significant difference exists between the subtraction produced and the division produced liquidity and liquidity premium difference. The specific equation is not repeated.

At the same time, the principle of pair construction is also analysed using the Fama-French portfolio matching model (Fama-French, 1993; Lin et al. 2011) and highly consistent results are obtained. The specific process can be seen in Appendix.

5.4.2 Pair Construction Approach

Bond pairs in different cases are constructed. For each bond pair, one single liquidity impacting factor is maintained following the elimination of others, which is then tested and the difference in the factors on the bid-ask spread and the yield is measured. The key factors that are observed in single bond regression, including amount, bond age and maturity, are all studied individually. Bond pair constructions begin with cases of bonds issued by identical issuers. Further studies should focus on dealing with bonds issued by different issuers with similar ratings or by completely different issuers.

Specifically, in the process of bond pair construction, some details and difficulties should be dealt with as follows:

(a) More than Three Bonds Meeting the Pair Construction Condition.

If more than three bonds satisfy the screening conditions of bond pair construction, they are sorted according to the factor to be tested and paired in turns. A single bond can exist in more than one pair. For example, when studying the impact of the outstanding amount of bonds, if five bonds are found to have identical issuers, maturities and cash flow structures and only the outstanding amounts are different, they are sorted based on their outstanding amounts and bond pairs are formed in turns, with a total of four pairs.

(b) Two Bonds with Everything Identical except ISIN

In the process of constructing bond pairs, some pairs are generated with two bonds that have everything identical except ISIN. A bond can be endowed with two different ISINs due to policy restrictions, such as under 144A and REGS policies. In the 144A/REGS case, the bid-ask quotations may be slightly different in the starting dates and keep the exactly same after stabilisation. To avoid any potential error, this thesis removed such pairs only except when studying issuing market impacts.

(c) Define the Condition of "Same life" and "Same maturity"

It is possible for two bonds to have similar (but not identical) lengths, such as just several days difference. To involve data as much as possible, they may be treated as having the "same life". This condition is defined in pair constructions based on the ratio of bond length difference. If the difference between two maturity dates in a bond pair does not exceed 0.1% of the length of the shortest bond, the bond pair will be treated as having the same length in the model in this study. However, this does not apply to the same maturity. To avoid any potential market risk impact, this thesis adopts two bonds with precisely the same expiry date as the same maturity bonds.

(d) Define the Condition of "Same Amount"

Similarly, the issuing or outstanding amounts of two bonds cannot be exactly the same, but they can be incredibly close. The threshold of "same amount" is defined as 0.1%. For example, the model treats 1 million par value bonds and 1.001 million par value bonds as being the same amount.

(e) Define the Condition of "Close to Maturity"

It has been found that liquidity premium becomes dramatically high when a bond is close to expiry. Therefore, during the pair analysis process, liquidity impact is investigated by separating the bond into two stages: When a bond is close to or far from maturity, 5 % is chosen as a key point for examining the different stages: a bond is treated as being close to maturity when it has less than 5% of its life remaining. A maximum cap of 3 months is added to the remaining life in addition to avoid any potential situation that the 5% of life may still be too long, for instance, for a 30-year bond.

5.4.3 Pair Cases: Without Considering Issuer Differences

In order to explore the impact factors other than individual issuers have on liquidity to the greatest possible extent, this thesis starts with bond pairs from the same issuer. Factors relating to issuer specialness, including rating, financial status and equity volatility, are ignored. Cases are also started without considering the difference in maturity because maturity is the main factor for calculating the bond yield in traditional discounting approaches. Cases with different types of pairs using the aforementioned factors, including bond life, issuing amount, issuing market and cash flow structures, are taken. Bond ratings and market makers are also analysed by different groups. It must be explained once more that the analysis is based on the difference/change between liquidity factors and premiums, including bid-ask spread and yield of the same bond, rather than the original value. However, in the regression process, the same factors for a pair of bonds are also included for measuring the magnitude of the impact.

(1) Issuing/Outstanding Amount

Single bond regression of yield provides evidence of the significance of the amount, while regression of bid-ask spread provides little. However, based on the practical experience of traders, bond issuing amount is one of the most important considerations for traders in terms of the liquidity of bond transactions. Due to this disparity, pair analysis is used for examining the amount again.

Bonds with ten consecutive quotations are accepted for data sufficiency consideration. Bond pairs are constructed by selecting bonds with factors (bond life, remaining days and cash flow structure) that are the same, with the exception of issuing/outstanding amount. Regarding the amount, if the historical outstanding amount is different to the issuing amount, the outstanding amount replaces the issuing amount information. However, Bloomberg data shows the outstanding amount to be basically the same as the issuing amount for the majority of bonds during their life. This makes sense as the investigated bonds are plain with no option embedded and the outstanding amount is generally equal to the issuing amount.

In addition, the notional value of the issuing amount (using the value of the smaller of the bonds) is maintained and its potential impact is examined in the regression model. The absolute values of bond ages, remaining life (days to maturities) and cash flow factors(coupon rate in this case) are also involved so the potential impacts can be observed. The detailed regression equations are as follows:

$$\mathbf{Yield}_a - \mathbf{Yield}_b = \boldsymbol{b'_1(Amount_a - Amount_b) + b'_2(Amount_{min}) + b'_4(Age_{ab}) + b'_5(Maturity_{ab}) + b'_6(CSFS_{ab}) + \varepsilon}$$

$$\mathbf{BA}_a - \mathbf{BA}_b = \boldsymbol{d'_1(Amount_a - Amount_b) + d'_2(Amount_{min}) + d'_4(Age_{ab}) + d'_5(Maturity_{ab}) + d'_6(CSFS_{ab}) + \varepsilon}$$

The statistics for observations and variables are as follows:

Table 6: Statistics for observations and variables (Case1)

Case 1 is to study the impact of the amount by setting the same issuer, same maturity and same bond length. Bonds are arranged in descending order of issuer, maturity, life and issuing amount when constructing pair. Differences by division (~div) and subtraction (~diff) are calculated between bonds. Two bonds with a length difference of less than 0.1% are treated as the same life. Bonds that cannot find a partner in pairs will be excluded. The initial pair-constructing process formed 3,446 pairs in case1, but reduced to 1,204 pairs to satisfy the requirement that bid/ask are available on the same day and that no less than ten quotes over five years. The data reduction is significant. It will be discussed in the limitation and drawback section in the last chapter of the thesis.

Observations	**Initial**	**above 10 quotes**		**Case1: Amount**	
Pairs	3446	1204			
bonds	5938	2170			
Issuers	265	166			
Variables	**Obs**	**Mean**	**Std.**	**Min**	**Max**
yielddiff	625,873	0.108862	0.8001	-50.321	70.684
yielddiv	607,991	1.008671	0.1032	-15.4186	21.159
badiff	625,872	0.005786	0.3734	-38.627	138.29
badiv	605,751	1.086768	1.0278	-27.25	309
amount(M)	1,644,498	263.2368	497.36	0.001	6000
amountdiv	1,644,498	11.73741	355.28	0.0000436	12338
amountdiff	1,644,498	-8.67283	228.15	-2990	3489.5
life	1,644,498	11.93432	10.284	0.3342	100.07
quotesdays	1,644,498	548.3793	418.07	10	1364

The panel data regression output, displayed in Table 7, shows issuing amount to be of importance. The left part of the table displays the situation for where the bond is close to maturity and the right part displays where it is far from expiring. The number of observations of the latter is much larger than that of the former, representing a more general situation.

The first two rows of Table7 reveal the difference in the amount to be decisive on the bid-ask spread difference and the yield difference when a bond is both nearing and far from maturity. The significance of the impact is larger to the bid-ask spread when a bond is close to maturity. This proves that the issuing amount difference plays a crucial role, especially when the bond is close to expiry. The significance of the issuing amount decreases, demonstrating that the yield difference is not affected by it.

When a bond is close to maturity, the significance of bond life and rating exists. As a relatively smaller data scale, reasons of occasional issues cannot be rejected. This will be investigated in the later cases.

The bond age (AgePCT) is significant all the time. This implies that neither the yield difference nor the bid-ask spread difference is a static value that only relates to the original feature of the bond. It makes sense that the calculation of the liquidity premium should involve the decaying time factor, although the exact formula for this is not known.

Table 7: Pair regression for issuing/outstanding amount

BAdiff is the bid-ask spread difference of the bond pair measured by subtraction, while BAdiv is that which is calculated by division. Yielddiff is the quoted yield difference of the bond measured by subtraction. It is also the difference of the mid prices. Yielddiv is that which is calculated by division. Amount is the amount of Bond1 in the pair, which is measured in millions. Amountdiv is the ratio of amount between bond 1 and bond 2. Cpn is the coupon rate, which is measured in units of 100bp. Life is the length of the bonds in the pair, Remain-life is the remaining days till the maturity and Agepct is the ratio of days past. These three factors are the same for the two bonds in the Case 1. These factors are also used in the below regression and will not be repeated.

	Close to maturity (age >95% & remain life < 3m)				**Far from expiring (age <95% of life)**			
	badiff	badiv	yielddiff	yielddiv	badiff	badiv	yielddiff	yielddiv
Amountdiff	**0.0113*****	**-0.0654*****	**-0.00394***	**0.00199***	**-0.0000266**	**-0.000620*****	**-0.000259**	**-0.0000752*****
	(3.73)	(-4.75)	(-2.05)	(2.16)	(-1.70)	(-7.45)	(-1.64)	(-6.21)
Amount	**0.00240****	**-0.0182*****	-0.000669	0.000533*	**0.0000296***	**0.000197***	0.000000866	-0.00000199
	(3.08)	(-5.23)	(-1.76)	(2.30)	(2.00)	(2.42)	(0.01)	(-0.17)
Cpn	-5.102**	37.49***	1.667*	-1.213*	0.0118**	-0.0273	0.134**	-0.00529
	(-3.25)	(5.15)	(2.17)	(-2.49)	(2.88)	(-1.12)	(3.25)	(-1.54)
Remain.life	**4.271***	**-15.07****	-0.103	0.542	-0.00163***	-0.00515	0.000815	-0.00349***
	(2.49)	(-2.97)	(-0.12)	(1.60)	(-7.38)	(-1.74)	(0.95)	(-18.84)

Life	**-2.171***	**6.454***	**0.717***	**-0.189**	-0.000869	0.00775	-0.0120	0.00336***
	(-5.42)	(3.46)	(3.65)	(-1.52)	(-0.94)	(1.36)	(-1.32)	(4.50)
Rating	**1.123***	**-1.967***	**-0.300**	**0.0935**	-0.00140	0.00884	-0.0201	0.00228
	(5.70)	(-2.43)	(-3.11)	(1.73)	(-0.68)	(0.80)	(-0.98)	(1.42)
AgePct	**17.50**	**-265.3***	**3.191**	**9.357***	**-0.0236***	**-0.133***	**0.0479**	**-0.0495***
	(0.94)	(-4.75)	(0.35)	(2.50)	(-6.07)	(-2.56)	(3.17)	(-15.29)
_cons	27.34	0	-18.14	0	0.00566	1.090***	-0.131	1.019***
	(1.20)	(.)	(-1.62)	(.)	(0.53)	(17.67)	(-1.23)	(120.13)
N	450	393	450	393	118083	112226	118083	112230
r2_w	0.104	0.157	0.0277	0.0468	0.000492	0.0000705	0.000169	0.00326
r2_b	0.996	1	0.417	1	0.0745	0.291	0.0927	0.186

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(2) Bond Life and Bond Age

The result of the above single bond regression shows that bond life (length of a bond) and bond age (the number of days the bond lives) impact bond liquidity premium in different ways. Therefore, the impact of bond life and bond age must be examined using a pair analysis approach. The bond life and age are examined in different ways: life uses value (daily data converted to year), while age uses percentages (bond lives day as a percentage of total bond length) to split the value impacts from life. Age differences are calculated by only subtraction instead of division to keep data stable. When the absolute amount is too small, the difference by the division may fluctuate too much.

Bonds with ten consecutive quotations are adopted and those with the same factors apart from age are selected. Regressions are then run based on them being close to maturity or far from maturity. The absolute values of amount, life and age are involved and the regression equations are as follows:

$$\mathbf{Yield}_a - \mathbf{Yield}_b = b'_1(life_a - life_b) + b'_2(Amount_{min}) + b'_4(Age_{ab}) + b'_5(life_{ab}) + b'_6(CSFS_{ab}) + \varepsilon$$

$$\mathbf{BA}_a - \mathbf{BA}_b = d'_1(life_a - life_b) + d'_2(Amount_{min}) + d'_4(Age_{ab}) + d'_5(life_{ab}) + d'_6(CSFS_{ab}) + \varepsilon$$

Table 8: Statistics for observations and variables (Case2)

Case 2 is to study the impact of the bond age/life by setting the same issuer, same maturity and *same bond length. Bonds are arranged in descending order of issuer, maturity, issuing amount and life when constructing pair. Differences by division (~div) and subtraction (~diff) are calculated between bonds. Two bonds with an amount difference of less than 0.1% are treated as the same amount. The amount figure adopts the outstanding amount in Bloomberg, and if not achievable, turns to the issuing amount. Bonds that cannot find a partner in pairs will be excluded. The initial pair-constructing process formed 1,322 pairs in case2, but reduced to 612 pairs to satisfy the requirement that bid/ask are available on the same day and that no less than ten quotes over five years.*

Observations	**Initial**	**above 10 quotes**		**Case2: Bond life & age**	
Pairs	1322	612			
bonds	2417	1117			
Issuers	210	160			
Variables	**Obs**	**Mean**	**Std.**	**Min**	**Max**
yielddiff	268,153	0.318304	1.2372	-31.669	70.684
yielddiv	238,980	1.002609	0.1012	-8.277778	21.159
badiff	268,153	0.054894	1.4619	-160.947	245.25
badiv	238,859	1.036044	0.7121	-22.75	114
amount(M)	834,768	566.1858	656.32	0.1	4072.2
life	834,768	11.65059	11.843	0.5753425	100.07
lifediv	834,768	0.955104	0.1265	0.1148168	1.1171
agepctdiff	515,719	-0.01943	0.0679	-0.885183	0.1049

Table 9 presents the panel data regression results. It is found that length and age matter and that they work differently in different stages. The left part shows the situation where the two bonds are both close to maturity, the mid part displays where they are both far from expiring, and the right part presents the rest.

The first row illustrates that the difference in bond length does not directly affect liquidity risk premium in the bid-ask spread when the bond is close to maturity. The significance impacts on yield reveal the existence of on-the-run/off-the-run premium. It also has a positive correlation with the difference in the bid-ask spread when far from expiry. The fourth row implies the significance of bid-ask be positively impacted by the days of the bond life, but without strong evidence.

The fifth row shows the age of the bond is always decisive. It stands that the difference in premium between long- and short-term bonds exists both on the yield and the bid-ask spread. It is also verified by the fixed effect regression on the age percentage factors (Table 10). Generally, it appears that a trader must consider the liquidity impacts from bond life when it is close to maturity. The trader must then pay more in liquidity premium for a bond.

Table 9: Pair regression for bond life/age

	when close to maturity (age >95% of life)				When far from expiring (age <95% of life)				Others (two bonds on different stages)			
	badiff	badiv20	yielddiff	yielddiv	badiff	badiv	yielddiff	yielddiv	badiv	badiv	yielddiff	yielddiv
lifediv	-0.577	-1.053	**5.271*****	**-0.509****	**-0.595****	**0.504*****	-0.240	0.0550*	1.614*	0.0454	0.144	0.0904**
	(-0.12)	(-0.47)	(4.09)	(-2.58)	(-2.68)	(3.58)	(-0.50)	(2.27)	(2.25)	(0.10)	(0.24)	(2.66)
amount	-0.00189	-0.00508**	0.000192	0.000266	0.0000359	-0.00012**	-0.0000207	-0.0000048	-0.000939	-0.00103*	-0.0000824	0.0000526*
	(-1.29)	(-2.58)	(0.38)	(1.85)	(0.71)	(-2.94)	(-0.15)	(-0.67)	(-1.95)	(-2.29)	(-0.20)	(2.31)
cpn	1.685***	1.454**	0.310**	-0.0519	0.0952***	-0.00698	0.223***	0.00145	0.529***	-0.0513	0.210**	-0.0132
	(6.14)	(3.13)	(3.22)	(-1.43)	(6.13)	(-0.47)	(5.18)	(0.54)	(5.81)	(-0.44)	(2.75)	(-1.78)
life	-0.681***	-0.379	-0.0998	0.00945	-0.014***	0.00179	-0.00629	0.00042	-0.172**	0.135	0.0276	-0.00685
	(-3.81)	(-1.86)	(-1.83)	(0.49)	(-4.31)	(0.62)	(-0.70)	(0.82)	(-2.77)	(1.30)	(0.54)	(-0.97)
agepctdif	**212.2**	**9.523**	**-124.9*****	**16.83*****	**0.804*****	**-0.331*****	**0.439*****	**-0.0118***	**-2.702*****	**-0.899*****	**-0.247**	**-0.145***
	(1.86)	(0.26)	(-6.41)	(3.93)	(3.65)	(-4.75)	(4.19)	(-1.99)	(-4.29)	(-11.63)	(-0.72)	(-2.04)
_cons	0.644	1.764	-5.130***	1.550***	0.446*	0.600***	-0.0770	0.943***	-1.491**	0.560**	-0.418	0.961***
	(0.15)	(0.86)	(-4.47)	(8.62)	(2.24)	(5.07)	(-0.19)	(47.05)	(-3.18)	(3.05)	(-1.08)	(58.55)
N	1791	570	1791	570	83533	62972	83533	62983	412	287	412	287
r2_w	0.00218	0.000048	0.0228	0.0275	0.000125	0.000353	0.000204	0.000065	0.0468	0.333	0.00141	0.00124
r2_b	0.433	0.416	0.415	0.208	0.192	0.0760	0.170	0.0528	0.651	0.525	0.550	0.423

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Pair regression for bond age, fixed effect

This is a fixed effect panel data regression for case 2. The only time-varying factors are the bond ages (measured by the percentage of the age to the whole life, AGEPCT). One of the three factors has a coefficient of zero due to the s perfect collinearity.

	badiff20	badiv20	yielddiff20	yielddiv20
Agepctdiffs	-3.180***	-0.812***	-0.0519	0.0770***
	(-9.84)	(-9.16)	(-0.44)	(9.91)
Agepct bond2	1.788***	0.254***	-0.0617	-0.0133***
	(18.69)	(8.33)	(-1.78)	(-4.96)
Agepct bond1	0	0	0	0
	(.)	(.)	(.)	(.)
_cons	-0.928***	0.836***	0.810***	1.010***
	(-15.41)	(47.56)	(37.19)	(655.56)
N	85736	63829	85736	63840
r2_w	0.00410	0.00152	0.0000792	0.00156
r2_b	0.0176	0.00594	0.0421	0.0249

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11 shows the correlation between the bid-ask spread difference and the yield difference, measured respectively by subtraction and division. As only the ages are different in this case, we can observe the relationship more easily. There is little evidence for the correlation between differences in yield and in bid-ask spread when the bond is far from maturity, and the relationship strengthens when close to maturity. This study observes no correlations between them in other cases, which implies the correlation is weak compared to other determinants.

Table 11: Correlation between yield and bid-ask spread

AgePCT	<0.1	0.1-0.3	0.3-0.5	0.5-0.7	0.7-0.9	>0.9
bid-ask diff vs yield diff	0.2612	0.1462	0.3922	0.3189	0.508	0.479
bid-ask div vs yield div	-0.0008	-0.1138	-0.0492	0.3336	0.0092	-0.175

(3) Issuing Markets and Venue

Issuing market shows whether a bond is mainly targeted at the US, the European public market or the private equity market. It should be noted that in this study, these markets refer to the markets of investors when the bonds were first issued. Once a bond is issued and is available in the secondary market, any OTC investors are able to trade it without restriction from the issuing market, in theory.

Table 12: Statistics for observations and variables (Case 3)

Case 3 is to study the impact of the issuing market and venue, with everything else included in case1 and case2 equal.

Observations	**Initial**	**above 10 quotes**		**Case3: venue**	
Pairs	799	454			
bonds	1531	864			
Issuers	196	138			
Variables	**Obs**	**Mean**	**Std.**	**Min**	**Max**

yielddiff	212,482	0.1443509	0.768353	-6.945	70.684
yielddiv	201,668	1.004252	0.107461	-8.277778	21.15929
badiff	212,482	0.0147939	0.566554	-35.735	138.292
badiv	201,547	1.055007	0.7516352	-1.165049	114
amount(M)	619,256	604.6703	600.0716	0.1	4072.197
life	619,256	12.67193	12.59885	1.49589	100.0685

By running different issuing markets groups, including the European market(Bloomberg Tag: EURO), US market(Bloomberg Tag: US DOMESTIC or DOMESTIC),Global market (Bloomberg Tag: GLOABL), private market(Bloomberg Tag: PRIVE PLACEMENT), etc., it is found that issuing amount difference and bond life difference still impact liquidity premium and the direction and extent are stable. This result is in accordance with the conclusion reached by BIS-CGFS (2016). The BIS-CGFS report argues that liquidity premium remains stable in the market with different market makers. It also accords with the result from the single bond regression process.

Table 13: Pair regression for issuing market

	(1) badiff	(2) badiff	(3) badiff	(4) badiff	(5) badiff
cpn	-.001	-.006**	0	-.002***	.563
	(.002)	(.003)	(.002)	(.001)	(10.119)
AmountMillion	0	0	0*	0***	-1.478
	(0)	(0)	(0)	(0)	(6.47)
freq	0	.001	0	.002**	.671
	(.004)	(.002)	(0)	(.001)	(4.864)
lifeyear	0	.003***	0	0***	-.27
	(0)	(.001)	(.001)	(0)	(3.891)
agepct	-.017***	-.007***	-.009***	.003**	-.002
	(.001)	(.002)	(.002)	(.002)	(.002)
_cons	.019		.001	.01***	
	(.012)		(.003)	(.003)	
Observations	180472	13952	22820	6087	2994
Pseudo R^2	.z	.z	.z	.z	.z

Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

(4) Cash Flow Structures (CSFS): Coupon Size and Frequency

Yield is calculated on the basis of the current market price of a bond, the frequency of interest payments on it, coupon rate and maturity dates. The impact of bond cash flow is also considered. Bonds are classified into two categories based on whether their coupons are paid semi-annually. They cannot be broken down further as the observations would be too small. After repeating the regression, results show that neither frequency nor coupon significantly impacts the final liquidity premium result.

5.4.4 Pair Cases: Considering Different Issuers with the Same Rating

Firstly, it can be concluded from the previous regression models that three factors are most important: issuing amount, age and whether the bond is nearing maturity. When focusing on bonds issued by different issuers but with the same credit rating level, only a few bond pairs have the same life, maturity and issuing amount. The remaining data cannot support traditional regressions. Therefore, the life and the issuing amount are controlled as a means of investigating the role of different issuers in the same ratings.

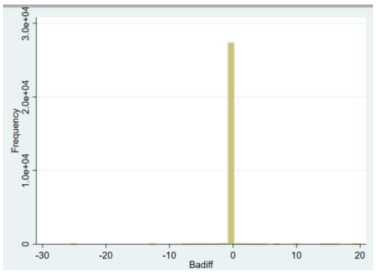
After running regressions, no significant evidence is found that the difference in credit rating has an impact on the difference of bid-ask spread premium or yield premium. It is interpreted that the difference in yield or bid-ask spread mainly reflects pure liquidity. Credit rating impact may work through the absolute value of other existing liquidity factors, including issuing amount. However, when there is an issuing amount difference, a lower rating bond has a more considerable sensitivity of issuing amount to the liquidity premium. No contradiction can be found between this conclusion and evidence relating to positive feedback on credit and liquidity risk from previous studies (Huang & Huang, 2012; He & Milbradt, 2014). The difference and the absolute value of credit status will potentially work differently.

(1) Controlling Issuing Amount

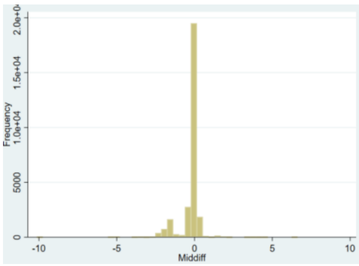
For bond pairs with the same issuing amount and same rating but with different issuers, almost all bid-ask spread and the yield difference are concentrated at zero (Figure 7, upper charts). This proves that the liquidity factors discussed above still apply to bonds of different issuers within the same credit rating range and liquidity is completely controlled.

(2) **Controlling Bond Life**

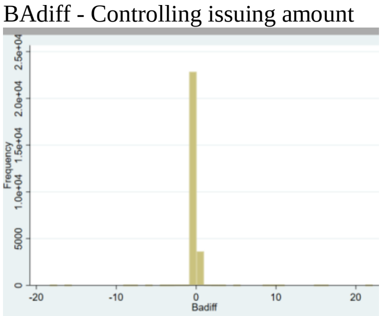
It is found that with bond pairs with the same bond life and same rating but different issuers, all the bid-ask spread and the yield difference are concentrated at zero (Figure 7, lower charts). This proves that the liquidity determinants discussed above still apply to bonds of different issuers within the same rating range and liquidity is almost completely controlled.



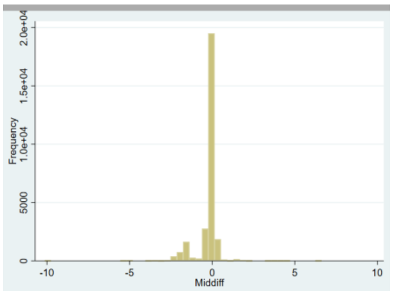
BAdiff - Controlling issuing amount



Middiff - Controlling issuing amount



BAdiff - Controlling bond life



Middiff - Controlling bond life

Figure 7: Histograms of paired bid-ask spread difference and yield difference

Table 14: Rating impacts analysis - controlling amount

Only four types of credit ratings are involved in this group of bonds: AA+, AA, A and A-. They are divided into two groups: higher rating (above AA) and lower rating (below A). It can be seen that the impact of credit rating is not obvious in such a high credit rating situation.

	(1)(all bond pairs) middiff	(2) (Higher rating pairs) middiff	(3) (Lower rating pairs) middiff	(4) (all bond pairs) badiff	(5) (Higher rating pairs) badiff	(6) (Lower rating pairs) badiff
rating	-.025			0		
	(.05)			(.023)		
lifediv	-.011	.036	.141	-.019	-.015	.019
	(.161)	(.16)	(.14)	(.076)	(.081)	(.013)
agepct	.585***	.481***	2.084***	-.122***	-.148***	.153***
	(.021)	(.02)	(.107)	(.016)	(.018)	(.011)
Amount	0	0	-.002***	0	0	0***
	(0)	(0)	(0)	(0)	(0)	(0)
_cons	-.079	-.13	.827***	.014	.016	-.003

	(.139)	(.125)	(.174)	(.066)	(.063)	(.016)
Observations	26647	23080	3567	26606	23080	3526
Pseudo R^2	.z	.z	.z	.z	.z	.z

Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

Table 15: Rating impacts analysis - controlling life

Only six types of credit ratings are involved in this group of bonds: AAA+AAA, AA+, AA, A, A-. and BBB. Ratings above AA are classified as higher and the other three as lower. It can be seen that the rating shows no significant impact on the liquidity premium, but when bond life is controlled, the bond issuing amount has a stronger impact on liquidity as the rating decreases.

	(1)(all bond pairs) middiff	(2) (Higher rating pairs) middiff	(3) (Lower rating pairs) middiff	(4) (all bond pairs) badiff	(5) (Higher rating pairs) badiff	(6) (Lower rating pairs) badiff
rating	.071***			-.015		
	(.023)			(.011)		
amountdiv	-.119***	-.147***	-2.197*	-.041**	-.043**	.244**
	(.045)	(.039)	(1.182)	(.021)	(.021)	(.102)
agepct	-.11***	-.151***	-.152	.01	-.015	.266***
	(.028)	(.015)	(.225)	(.016)	(.017)	(.067)
Amount	0	.001***	0	0*	0***	0***
	(0)	(0)	(.001)	(0)	(0)	(0)
_cons	.01	.154***	2.612**	.046*	.022	-.241**
	(.052)	(.024)	(1.144)	(.024)	(.014)	(.098)
Observations	26463	23451	3012	26422	23451	2971
Pseudo R^2	.z	.z	.z	.z	.z	.z

Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

5.5 Summary

After pair regressions were run, the following key findings were discovered:

(1) Issuing amount difference performs differently between yield difference and bid-ask difference at different stages. When it is far from maturity, there is a significant impact for the yield but not the bid-ask spread. When close to maturity, the impact on the bid-ask spread difference increases. This increase is amplified with the absolute value of the issuing amount.

(2) Bond life difference impacts the yield difference when the bond is nearing maturity. When it is far from maturity, life difference impacts the bid-ask spread instead of the yield, but this turns significantly when close to maturity.

(3) There is little but not much evidence for the correlation between yield difference and bid-ask spread difference when the bond is far from maturity, and the relationship strengthens when close to maturity. The direction and extent of the difference between bid-ask difference and yield difference have a tendency to be consistent.

(4) Cash flow structure (coupon, frequency) and market venue difference have little impact on either yield difference or bid-ask difference.

(5) Impact does not vary for different credit rating groups within the investment grades.

The impacts of each factor can be summarised as follows:

Main Factors	**Impacts on B-A spread**	**Impacts on Yield**	**Others**
Cash Flow Structure	N	N	significance in some groups, not universal
Ratio/Difference of Issuing amount	**Y** (close to maturity only) • Negative impacts ***, • Positive impacts on absolute amount *** • Impacts increase when closer to maturity	Y(far from maturity only) • Negative impacts ***	Correlation between BA spread and Yield increase when close to maturity
Ratio/Difference of Bond age	**Y** Positive impacts ***(close to maturity),	**Y** • Positive impacts ***(close to maturity), • Negative impacts ***(far from maturity)	Minor impacts when away from maturity
Ratio/Difference of Bond life	**Y** (far from maturity only) • Positive impacts *	**Y** (close to maturity only) • Positive impacts *	Minor negative correlation between life and spread
Groups of Market markers/Venue	N	N	The level of parameter keeps stable between different groups Similar conclusion with BIS-CGFS(2016)

Groups of Credit	N	N	The tendency appears that impacting the power of issuing amount differs in different credit groups

Chapter 6 Conclusions

6.1 General Conclusions

Firstly, this research analyses the way in which liquidity premium is affected by liquidity factors from the perspective of bond pairs. When a bond is compared with other known bonds, it is found that liquidity premium differences (the difference of the bid-ask spread and the difference of the yield spread) can be expressed by the difference of liquidity factors, specifically:

Bond issuing amount difference and bond age difference are dominant with respect to liquidity premium difference and the impacts are significant. Issuing amount is always dominant and the impact it has on bid-ask spread is more substantial if the bond is closer to maturity date. Age only impacts yield difference if the bond is close to maturity date. When a bond is close to maturity date, the impact issuing amount and bond age have on bid-ask spread increases, determining the majority of changes in the bid-ask spread. This leads to the possibility of prediction.

Little evidence is found regarding the significance of credit rating difference impacting the bid-ask spread difference or yield spread difference of an individual bond. However, only investment-grade bonds are involved due to a scarcity of data. The possibility that the difference still works for high-yield bonds cannot be rejected. However, the actual impacts on high-yield bonds require verification.

Little evidence is found regarding the significance of bond cash flow structure difference impacting the liquidity premium difference of an individual bond.

In addition, it is known that yield spread difference and bid-ask spread difference vary as bond age increases. There is no direct correlation between yield difference and bid-ask spread difference when bonds are far from the maturity date, but the correlation improves when bonds are closer to the expiry date.

6.2 Interpretation of Conclusions

The above conclusions are in accordance with the phenomenon observed in the market. A bond that has a larger issuing amount is generally quoted with a lower yield and a lower bid-ask spread. This means that the bond has a capacity for more funds and investors and therefore, it is more attractive.

But why does the yield spread difference exhibit stable performance while the bid-ask spread changes dramatically when the bond is close to maturity? This is because a bond that is issued earlier is generally quoted with higher yield spreads than a bond that is issued later when the two bonds mature on the same day, particularly when the former is nearing maturity. Ordinarily, investors wish to replace an older bond with a younger one as a means of maintaining their investment strategy and so they are willing to pay for the premium. At the same time, when a bond is close to maturity, due to the "maturity reinvestment risk" in an investment-grade bond market that is dominated by prudent investors, bonds with a small issuance volume are of little interest.

A further interpretation is that yield spread liquidity premium, driven by the individual bond liquidity feature, stands for the additional cost investors are willing to pay for an illiquid bond. This shows expectations among investors of how easy a bond can be turned to cash in future. Bid-ask spread liquidity premium stands not only for the bond liquidity feature but also the bond liquidity status on the market, which is driven by both individual liquidity and market supplies. It measures or how easily the bond can now be turned to cash now.

When a bond is far from maturity date, there is less of a correlation between the individual liquidity feature and market liquidity status. When it is close to maturity, there is no difference between trading and holding-to-maturity, so there tends to be consistency with the direction and extent between the bid-ask difference and yield difference. When a bond is close to maturity, the liquidity premium of both will be highly correlated with each other and can be more easily estimated.

6.3 Inferring Unknown Bond Liquidity Premium from Known Bonds

Based on the significant impacts issuing amount and bond age have on yield difference and bid-ask spread difference, particularly the relationship when the bond is close to maturity, the price of one bond (particularly that with sufficient consistent quotation data) can potentially be used for the prediction of the price of another. This works better when there is a need to estimate the spread of an illiquid bond with few quotations and relatively little other information.

Starting with known bonds with sufficient quotations and transaction price data, the relationship between risk premium difference and liquidity factors difference can be examined by constructing “bond pairs” based on specific rules. Similar impacts are controlled and eliminated from different bonds, comparative analysis is performed and the liquidity determinants are put into an estimation regression model on liquidity premiums. If the number of available bond partners is sufficient, the approach can be expanded from a simple pair approach to a pair portfolio, similar to the portfolio matching analysis. As the impacting factors of the pair regression only include the amount, age and individually specific characteristics of the unique study target cases, the accuracy and efficiency of the pair-analysis model are improved in comparison to those for traditional single bonds, which require the collection of more data relating specific bond issuer information, financial information, market environment information and historical quotations or transactions.

In addition, this study finds no evidence that credit rating difference impacts yield premium difference and bid-ask spread premium difference. It is noticed that the impacts are no different under different rating levels within the investment grade. Therefore, the possibility that the conclusion and the feasible method can work with other rating bonds such as high-yield bonds cannot be rejected. However, this is only a preliminary interpretation and the actual impacts on the high-yield bond require verification.

6.4 Limitations and Drawbacks

This thesis contributes to the study of OTC bond liquidity using bond pairs. However, there are limitations and drawbacks.

The first limitation relates to data. The pair bond analysis requires precise matching of bonds, resulting in a reduction to sample size when the data is screened and narrowing down the application range. There is a significant reduction in the amount of data available through the analysis process. However, the fastest reduction stage is screen data with at least ten quotations, which is not the limitation of pair analysis but the limitation of the OTC bonds. Therefore, all available data is used in the model to satisfy the minimum data requirement of the study, but this is perhaps not enough for result verification. In future studies, the data range should be expanded to enable the verification of conclusions.

The second limitation is that the estimation method used in this study only works when the same issuer has previously issued a bond that exists in the OTC market due to the high requirement for precise matching. Although the insignificance of the rating potentially implies that it also works for different issuers with the same rating, this has yet to be tested. This also can be treated as another type of data limitation.

The third limitation is that the thesis is only focused on plain bonds. However, plain bonds are the basic type of all bonds and complex bonds are basically formed on their structure.

Fourthly, potential sudden credit rating fluctuations during the life cycle are ignored in this study. It is assumed that bond pairs have the ability to eliminate mild credit issues, but this is not tested in wild credit rating change circumstances.

Finally, individual issuer credit differences under the same rating level are also not fully considered. This is limited by data scale and the credit rating issue is a direction that can be expanded upon in future studies.

6.5 Potential Directions for the Future Study

Future studies could be expanded to pair selections. Some “big name” issuers may issue many similar bonds that are available for the construction of “pairs”. of the method for selecting one or a series of the most suitable bonds could also be studied. The latter situation could improve the estimation through the introduction of a portfolio-matching regression on pairs.

Future studies could also be expanded through the involvement of more issues in the real market. A typical example is transaction volume. Considering “market depth” in the model could optimise the result of estimating the final premium in a separate transaction. Other examples include sudden credit changes and market noises.

The interpretation of interesting phenomena that were observed in this study could be further studied. For example, rapid changes to bid-ask spread difference when a bond is close to maturity date. The different roles of yield spread difference and bid-ask spread difference could be further interpreted and verified, including different types of liquidities, different features against more liquidity factors and the interaction between both roles.

There is also scope to expand this study to the high-yield bond area through the introduction of more data relating to speculative-grade bonds.

References

Abudy, M.M., & Raviv, A. (2016), How Much Can Illiquidity Affect Corporate Debt Yield Spread? Econometric Modelling: Capital Markets - Risk eJournal.

Acharya, V. V., Drechsler, I., & Schnabl, P. (2013), A Pyrrhic Victory? Bank Bailouts & Sovereign Credit Risk, mimeo, NYU-Stern.

Acharya, V.V. & Pedersen, L.H. (2005), Asset Pricing with Liquidity Risk. Journal of Financial Economics, 77, 375-410.

Adrian, T., Fleming, M., Shachar, O., & Vogt, E. (2017), Market Liquidity after the Financial Crisis. Federal Reserve Bank of New York Staff Reports,

Ai. C., Zhang. Y, & Cui. C. (2015), An Empirical Study on the Correlation Premium of Bond Liquidity & Default Risk, J. Manag. Sci. CHINA, vol. 18, no. 5

Amihud, Y. (2002), Illiquidity & stock returns: Cross-section & time-series effects. Journal of Financial Markets 5: 31–56 ,

Amihud, Y. & Mendelson, H. (1986), Asset pricing & the Bid-Ask Spread. Journal of Financial Economics 17: 223–279 .

Amihud, Y. & Mendelson, H. (1991), Liquidity, maturity & the Yields on U.S. government securities. Journal of Finance 46: 1411–1426,

Amihud, Y. & Mendelson, H. (2015), The Pricing of Illiquidity as a Characteristic & as Risk, Multinatl. Financ. J., vol. 19, no. 3, pp. 149–168,

Amihud, Y. & Mendelson, H. & Pedersen, (2013), L.H.Market Liquidity: Asset pricing, risk & crises. Cambridge University Press,

Amihud, Y. & Mendelson, H. (1989), The Effects of Beta, The Bid-Ask Spread & Residual Risk, & Size on Stock Returns. Journal of Finance 44: 479–486.

Amihud, Y., Hameed, A., Kang, W. & Zhang, H. (2015), The illiquidity premium:

International evidence, Journal of Financial Economics, 117, issue 2, p. 350-368.

Anderson, A. & Long, M. (2014), Explaining the On-the-Run Puzzle. SSRN Electronic Journal. 10.2139/ssrn.2388246.

Bagehot, W. (1971), The Only Game in Town. Financial Analysts Journal, 27, 12-14.

BAO, J., PAN, J., & WANG, J. (2011). The Illiquidity of Corporate Bonds. The Journal of Finance, 66(3), 911–946.

Bekaert, G., Campbell, H., & Lundblad, C. (2003), Liquidity & expected returns: Lessons from emerging markets, Working paper, Duke University.

Bongaerts, D., De Jong, F., & Driessen, J. (2011), Derivative Pricing with Liquidity Risk: Theory & Evidence from the Credit Default Swap Market. Journal of Finance, 66, 203-240.

Boudouck, J. & Whitelaw, R. (1993), Liquidity as a choice variable: a lesson from the Japanese government bond market. Review of Financial Studies 6, 265–292.

Boyarchenko, N., Giannone, D. & Shachar, O. (2018), Flighty Liquidity. FRB of New York Staff Report No. 870,

Br&t, M. & K. Kavajecz. (2004), Price discovery in the U.S. Treasury market: The impact on order flow & liquidity on the yield curve, Journal of Finance 59, 2623–2654

Brennan, M. & Subrahmanyam, A. (1996), Market microstructure & asset pricing: On the compensation for illiquidity in stock returns, Journal of Financial Economics, 41, issue 3, p. 441-464.

Bruche, M. & Segura, A. (2016), Debt Maturity & the Liquidity of Secondary Debt Markets (January 21, 2016). Bank of Italy Temi di Discussione (Working Paper) No. 1049

Brunnermeier, Markus K., & Pedersen, L.H. (2009), Market Liquidity & Funding Liquidity. Review of Financial Studies 22 : , 22, 2201-2238.

Campbell, J. & Taksler, G. (2003), Equity volatility & corporate bond yields, Journal of Finance 58, 2321–2349

Cetin,U. Jarrow, R., & Protter, P. (2004) ,Liquidity Risk & Arbitrage Pricing Theory, Finance & Stochastics, Vol. 8, No. 3, pp. 311-341.

Chakravarty, S. & Sarkar, A. (1999), Liquidity in fixed income markets: A comparison of the bid ask spread in corporate, government, & municipal bond markets, Working paper, Federal Reserve Bank of New York

Chang, J. (2015), Analysis of the Liquidity Effect in the Pricing of Inter-bank Treasury Bonds, J. Dongbei Univ. Financ. Econ., vol. 1, pp. 1–5,

Chen, H., Cui, R., He, Z., & Milbradt, K. (2014), Quantifying Liquidity & Default Risks of Corporate Bonds Over the Business Cycle (October 2014). NBER Working Paper No. w20638,

Chen, L., Lesmond, D.A., & Wei, J. (2007), Corporate Yield Spreads & Bond Liquidity. The Journal of Finance, 62, 119-149.

Chen, R.R., Fabozzi, F., & Sverdlove, R. (2010), Corporate credit default swap liquidity and its implications for corporate bond spreads. J. Fixed Income 20, 31–57

Chordia, T., Roll, R. & Subrahmanyam, A. (2000), Commonality in Liquidity. Journal of Financial Economics, 56, 3-28.

Choudhry, M. (2010), Measuring bond market liquidity: devising a composite aggregate liquidity score, Applied Financial Economics, 20:12, 955-973,

Cob&ag Guloglu, Z, & Ekinci, C. (2021), Liquidity measurement: A comparative review of the literature with a focus on high frequency. Journal of Economic Surveys. 36.

Committee on the Global Financial System, (2016), Fixed income market liquidity, BIS website, no. 55

Copel&, T.E. & Galai, D. (1983), Information Effects & the Bid-Ask Spread. Journal of Finance, 38, 1457-1469.

Covitz, D. & downing, C. (2007), Liquidity or Credit Risk? The Determinants of Very Short-Term Corporate Yield Spreads. Journal of Finance. 62. 2303-2328.

Daley,B. & Green,B.(2016), An Information-Based Theory of Time-Varying Liquidity, J. Finance, vol. 71, no. 2, pp. 809–870

Datar, V.T., Naik, N.Y., & Radcliffe, R. (1998), Liquidity & Stock Returns: An Alternative Test. Journal of Financial Markets, 1, 203-219.

De Jong, F. & Driessen, J. (2006), Liquidity Risk Premia in Corporate Bond Markets. European Finance.

Degryse, H. (1999), The total cost of trading Belgian shares: Brussel versus London. Journal of Banking & Finance 23, 133 1- 1355.

Demsetz, H. (1968), The Cost of Transaction. The Quarterly Journal of Economics, 82, 33-53.

Díaz, A. & Escribano, A.(2017),Liquidity measures throughout the lifetime of the U.S. Treasury bond, J. Financ. Mark, vol. 33, pp. 42–74,

Dick-Nielsen, J. (2009), Liquidity biases in TRACE[J], Journal of Fixed Income, Vol.19, PP43–55.

Dick-Nielsen, J., Feldhütter P., & L&o D. (2012), Corporate bond liquidity before & after the onset of the subprime crisis [J]. Journal of Financial Economics , 103(3): 471-492.

Duffie, D. & Kan, R. (1996), A Yield factor Model of Interest Rates. Mathematical Finance, 6, 379-406.

Edwards, A.K., Harris, L.E., & Piwowar, M.S. (2007), Corporate Bond Market Transaction Costs & Transparency. Journal of Finance, 62, 1421-1451.

Ejsing, J. & Grothe, M., & Grothe, O.(2012), Liquidity & Credit Risk Premia in Government Bond Yields (May 24, 2012). ECB Working Paper No. 1440,

Elton, E. J. & Green, C.(1998), Tax & Liquidity Effects in Pricing Government Bonds.NYU Working Paper No. FIN-98-029,

Elton, E., Gruber,M. Agrawal, D., & Mann,D. (2001), Explaining the rate spread on corporate bonds, Journal of Finance 56, 247–277.

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

Fama, F. & French, R. (1993), Common risk factors in the returns stocks & bonds *, vol. 33, pp. 3–56,

Fleming, M.J.(2001), Measuring Treasury Market Liquidity FRB of New York Staff Report No. 133,

Fons, J.S. (1994), Using Default Rates to Model the Term Structure of Credit Risk. Financial Analysts Journal, 50, 25-32.

Fontana, A., & Scheicher, M. (2010), An Analysis of Euro Area Sovereign CDS & Their Relation with Government Bonds. Derivatives eJournal.

Friewald, N., wagner, C. & zechner, J. (2014), The Cross-Section of Credit Risk Premia & Equity Returns. The Journal of Finance, 69: 2419-2469.

Friewald, N., Jankowitsch, R., & Subrahmanyam, M. G. (2012), Illiquidity or credit deterioration: A study of liquidity in the US corporate bond market during financial crises, Journal of Financial Economics, 105, issue 1, p. 18-36,

Garbade, D. & Silber, W. (1979), Structural organization of secondary markets: Clearing frequency, dealing activity, & liquidity risk, Journal of Finance 34, 577–593.

Garc,B. (2017), Conditional Conservatism & Management Earnings Forecasts, no. December

Garman, M. B. (1976), Market microstructure. Journal of Financial Economics, 3, 257-275.

Glen, J. (1994), An introduction to the microstructure of emerging markets. International Finance corporation Discussion Paper No. 2 Washington D.C. IFC.

Glosten, L. & Milgrom,P. (1985), Bid, ask, & transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 2, 24–38.

Glosten, L.R. (1989). Insider Trading, Liquidity, & the Role of the Monopolist Specialist. The Journal of Business, 62, 211-235.

Glosten, L.R. & Milgrom, P.R. (1985) Bid, Ask & Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. Journal of Finance, 14, 71-100.

Glosten, Lawrence R. & Lawrence E. Harris, 1988, Estimating the components of the bid-ask spread, Journal of Financial Economics 21, 123-142.

Goodhart, C. & O'Hara, M. (1997), High frequency data in financial markets: Issues & applications, Journal of Empirical Finance 4, 73–114.

Gregoriou, A. & Ioannidis, C. (2007), Generalized method of moments & present value tests of the consumption-capital asset pricing model under transactions costs: evidence from the UK stock market. Empirical Economics 32, 19–39 (2007).

Guo, H. & Yang, Z. (2006), On-the-Run & Off-the-Run Treasury Market Liquidity, working paper Shanghai Stock Exchange..

Han, S., & Zhou, H.(2008), Effects of Liquidity on the Non-default Component of Corporate Yield Spreads: Evidence from Intraday Transactions Data, Finance and Economics Discussion Series 2008-40. Board of Governors of the Federal Reserve System (U.S.).

Harris, L. (1991), Stock price clustering & discreteness. Review of Financial Studies 4(3), 389-415.

He, Z., & Milbradt, K. (2014), Endogenous Liquidity & Defaultable Bonds, Econometrica, vol. 82, no. 4, pp. 1443–1508

He, Z., & Xiong, W. (2012), Rollover risk and credit risk. J. Finance 67, 391–430.

Helwege, J., Huang, J., & Wang, Y. (2014), Liquidity effects in corporate bond spreads, Journal of Banking & Finance, 45, issue C, p. 105-116.

Huang, J. & Huang, M. (2002), How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk? (October 2002). NYU Working Paper No. S-CDM-02-05

Huang, R.D. & Stoll, H.R.(1996), Dealer versus auction markets:a paired comparison of execution costs on NASDAQ & the NYSE. Journal of FinancialEconomics 4 1, 3 13-358.

Huang, R.D & Stoll, H.R. (1997), The Components of the Bid-Ask Spread: A General Approach, Review of Financial Studies, 10, issue 4, p. 995-1034.

Johnson, T.C. (2008), Volume, liquidity, & liquidity risk, Journal of Financial Economics [J].

Kamara, A. (1994), Liquidity, Taxes, & Short-Term Treasury Yields. Journal of Financial & Quantitative Analysis, 29(3), 403-417.

Kang, W. & Zhang, H. (2013), Measuring Liquidity in Emerging Markets. S&P Global Market Intelligence Research Paper Series.

Kim, S.H. & Lee, K.H.(2014), Pricing of Liquidity Risks: Evidence from Multiple Liquidity Measures (June 19, 2013). Published at the Journal of Empirical Finance, Vol. 25, 112-133,

Krishnamurthy, A. (2001), The Bond/Old-Bond Spread. Journal of Financial Economics. 66. 463-506. 10.1016/S0304-405X(02)00207-6.

Lasser, D. J. (2010), The Effect of Short Selling on Market Reactions to Earnings Announcements, vol. 27, no. 2, pp. 609–638, Leal, D., Stanhouse, B.E., & Stock, D.R. (2020), Estimating the term structure of corporate bond liquidity premiums: An analysis of default free bank bonds. Journal of International Financial Markets, Institutions & Money, 67, 101217.

Lee, Kuan-Hui, (2011), The world price of liquidity risk, Journal of Financial Economics, 99, issue 1, p. 136-161,

Lesmond, D., Schill, M. & Zhou, C. (2004), The illusory nature of momentum profits, Journal of Financial Economics 71, 349–380.

Lesmond, D.,Ogden, J. & Trzcinka, C.(1999), A new estimate of transaction costs, Review of Financial Studies 12, 1113–1141.

Li, S. & Cheng,Y. (2014) Liquidity Analysis of Financial Derivative Products Based on Structural Model, J. TONGJI Univ. Sci., vol. 42, no. 11,

Liang, G., Lien, D., Hao, M., & Zhang, H. (2017) Uncertainty and liquidity in corporate bond market, Applied Economics, 49:47, 4760-4781,

Lin, H., Wang, J. & Wu, C. (2011), Liquidity risk & expected corporate bond returns, Journal of Financial Economics, 99, issue 3, p. 628-650.

Lippman, S.A. & McCall,J.(1986), An Operational Measure of Liquidity, American Economic Review, 76, issue 1, p. 43-55.

Liu, H. & Wang, Y.(2016), Market Making with Asymmetric Information & Inventory Risk . Journal of Economic Theory, Vol. 163,

Martinez, M.A., Nieto, B., Rubio, G., Tapia, M.(2005), Asset pricing & systematic liquidity risk: an empirical investigation of the Spanish stock market. International Review of Economics & Finance 14, 81-103

Massimb, M., Phelps, B., 1994. Electronic trading, market structure & liquidity. Financial Anal. J. 50 (1), 39–50.

Min, X. (2015), A empirical analysis on liquidity premium on Chinese bond market

Security Market Research,

Min, X. & Luo, H. (2016), Research on Corporate Bond Liquidity Premium Based on the Dual Effects of Level & Risk, Security Market Research, vol. 6,

Merton, C.R. (1974), On the pricing of corporate debt: the risk structure of interest

rates. J. Finance 29, 449–470.

Naik, N.Y. & Pradeep Yadav,K. (1999), The Effects of Market Reform on Trading Costs of Public Investors: Evidence from the London Stock Exchange, mimeo.

Nikolaou, K. (2009), Liquidity (Risk) Concepts: Definitions & Interactions. ECB Working Paper No. 1008

O'Hara, M. (1995), Market Microstructure Theory. Blackwell, Cambridge.

Oprică, S. & Weistroffer, C. (2019), Institutional Presence in Secondary Bank Bond Markets: How Does It Affect Liquidity & Volatility? ECB: Working Paper Series (Topic).

Orphanides, A. & Kim, D. (2007). The bond market term premium: What is it, & how can we measure it?. BIS Quarterly Review.

Pastor, L. & Stambaugh, R. F. (2003), Liquidity Risk & Expected Stock Returns. Journal of Political Economy, 111, 642-685.

Renault, O & Ericsson, J. (2000), Liquidity & Credit Risk. SSRN Electronic Journal. 10.2139/ssrn.424002.

Report.C (2016), Examination of Liquidity of the Secondary Corporate Bond Markets Consultation Report, Int. Organ. Secur. Comm.,

Roll, R. (1984). A Simple Implicit Measure of the Effective Bid‐Ask Spread in an Efficient Market. Journal of Finance, 39, 1127-1139.

Sadka, R. (2004), Momentum & Post-Earnings-Announcement Drift Anomalies: The

Role of Liquidity Risk. Journal of Financial Economics, Forthcoming, EFA Maastricht Meetings Paper No. 5290,

Sarig, O. & Warga, A. (1989), Bond price data & bond market liquidity, Journal of Financial & Quantitative Analysis 24, 367–378

Schultz, P. (2001), Corporate bond trading costs: A peek behind the curtain, Journal of Finance 56, 677–698.

Schwartz, R. A. (1988), Equity Markets: Structure Trading & Performance. New York, NY: Harper an Row ,

Sperna Weil&, R., Laeven, R. J., & De Jong, F. (2017), Feedback Between Credit & Liquidity Risk in the US Corporate Bond Market. 30th Australasian Finance & Banking Conference 2017

Stoll, H. (1989), Inferring the components of the bid-ask spread: theory & empirical tests, The Journal of Finance 44, 115-134.

Stoll, H. (2000), Friction, Journal of Finance 46, 1479–1514.

Strebulaev, I. (2002), Liqudity & Asset Pricing: Evidence from the U.S. Treasury Securities Market, working paper London Business School

Tang, D.Y. (2019), Comments on Measuring Corporate Bond Liquidity in Emerging Markets: Price- vs Quantity-Based Measures. BIS Paper No. 102f,

Tang, D.Y., & Yan, H. (2007), Liquidity & Credit Default Swap Spreads. AFA 2007 C hicago Meetings Paper, EFA 2008 Athens Meetings Paper

Tang, D.Y., & Yan, H. (2013),What Moves CDS Spreads, Capital Markets: Market Mi crostructure eJournal (June 21, 2013).

Tomura, H. (2016), Investment Horizon & Repo in the Over - the - Counter Market. Journal of Money, Credit & Banking, 48(1), 145-164.

Vayanos, D. & Weill, P.(2006), A Search-Based Theory of the On-the-Run Phenomenon. NBER Working Paper No. w12670,

Wang, X. (2013), Research on Corporate Bond Liquidity & Asset Pricing,

Wang, Y. T. & Wen, Z. Y. (2016), The influence of Chinese bond market's liquidity risk on bond pricing, Investment Research, 2016-6

Warga, A. (1992), Bond Returns, Liquidity, & Missing Data. Journal of Financial & Quantitative Analysis, 27(4), 605-617.

Zhang, S.X. (2018), Liquidity misallocation in an over-the-counter market. Journal of Economic Theory, 174. pp. 16-56. ISSN 0022-0531

Zhao, Y., & Yu, M. (2019), Measuring the liquidity impact on catastrophe bond spreads. Pacific-Basin Finance Journal.

Appendix

Theoretical Model Construction from Fama-French Approach

Lin, Wang, and Wu (2011) expanded the traditional Fama-French (1993) bond yield pricing model and according to this, the corporate bond yield can be expressed as:

$$\boldsymbol{r_{it}} - \mathbf{r}_{\boldsymbol{ft}} = \boldsymbol{a_i} + \boldsymbol{b_i^M Mt} + \boldsymbol{b_i^s St} + \boldsymbol{b_i^H Ht} + \boldsymbol{b_i^D Dt} + \boldsymbol{b_i^T T} + \boldsymbol{b_i^L Lt} + \boldsymbol{\varepsilon}$$

Where: $\boldsymbol{r_{it}}$: The return of bond I,

$\mathbf{r}_{\boldsymbol{ft}}$: The risk-free rate

$\boldsymbol{a_i}$: The cost of the transaction

M, S, H, D, T, L: Fama-French's Market portfolio factor, Market Capital factor, Market Ratio factor, Default factor, Date-to-Maturity factor, and Liquidity factor

It can be seen that in this model, the market risk factor can be embodied as a risk-free rate of return and the credit risk factor can be embodied as a default risk factor. M, S and H are all factors that correspond to stocks, which are mixed with market liquidity factors and individual liquidity factors. L is a pure market liquidity factor. If market liquidity and individual liquidity can be separated from the elements relating to the bond or equity, then:

$$\boldsymbol{r_{it}} - \boldsymbol{r_{ft}} = \boldsymbol{a_i} + \boldsymbol{b_i^E Equity} + \boldsymbol{b_i^D Dt} + \boldsymbol{b_i^T T} + \boldsymbol{b_i^L L(Market)t} + \boldsymbol{b_i^L L(Individual)t} + \boldsymbol{\varepsilon}$$

If there are two bonds (A and B), the yield difference between them can be calculated by simple subtraction: The risk-free yield of all bonds and market liquidity risk factors can offset each other – if the bond maturity date is the same, the maturity factor can be offset and if the bond rating is the same, the default risk can be offset. The market liquidity factors in M, S and H can also be offset and all that remains is the individual liquidity factor.

$$
\begin{aligned}
r_{At} - r_{Bt} &= a_A - a_B + (b_A^E - b_B^E)Equity + (b_A^D - b_B^D)Dt + (b_A^T - b_B^T)T + (b_A^L \\
&\quad - b_B^L)L(Market)t + (b_A^L - b_B^L)L(Individual)t + \varepsilon \\
&= c + (b_A^E - b_B^E)Equity + (b_A^D - b_B^D)Dt + (b_A^L \\
&\quad - b_B^L)L(Individual)t + \varepsilon
\end{aligned}
$$

According to the statement in the liquidity determinants section, individual liquidity factors are classified into the following categories:

CSFS: The Cash Flow Structure of a bond, including coupon, frequency, etc….
Size: Bond issuing/outstanding amount.
Age: Bond age.
Date-to-maturity: The remaining life of a bond. This will be offset when two bonds have the same maturity
Rating: Credit rating. This will be offset when two bonds have the same issuer
Market maker: Market maker of the two bonds

At the same time, bond yield is used to represent bond return and the above difference model can be simplified as:

$$
\begin{aligned}
&Yield_a - Yield_b \\
&\quad = c + b_1 * (CSFS_A - CSFS_B) + b_2 * (Amount_a - Amount_b) \\
&\quad + b_3 * (Age_a - Age_b) + b_4 * (Maker_a - Maker_b) + b_5 \\
&\quad * (Liquidity_a - Liquidity_b) + b_6 * (Venue_a - Venue_b) + \varepsilon
\end{aligned}
$$

At the same time, it is observed that the regression model only includes all the individual bond liquidity factors. If a liquidity proxy can be represented by these individual bond liquidity factors, then the right side is used to represent the liquidity proxy of the bond under different coefficients. For example, the bid-ask spread is:

$$
\begin{aligned}
BA_a - BA_b &= c' + b'_1 * (CSFS_A - CSFS_B) + b'_2 * (Amount_a - Amount_b) \\
&\quad + b'_3 * (Age_a - Age_b) + b'_4 * (Liquidity_a - Liquidity_b) + b'_5 \\
&\quad * (Venue_a - Venue_b) + \varepsilon
\end{aligned}
$$

The regression model can be simplified in different pairs, considering offsetting effects among the factors. For example, in bond pairs with the same issuer and maturity date, the regression is expressed as:

$$
\begin{aligned}
BA_a - BA_b &= C + b_1 * CSFS + b_2 * \%Size + b_3 * VENUE + \varepsilon \\
Y_a - Y_b &= C' + b'_1 * CSFS + b'_2 * \%Size + b'_3 * VENUE + \varepsilon
\end{aligned}
$$

In addition, considerate is considered that a nonlinear relationship may exist in the bond liquidity determination process. During the actual process, subtraction difference and division difference (ratio) are both used for exploring the coefficient relationship. The final specific regression equations are:

$$\frac{BA_a}{BA_b} = C + b_1 * \frac{CSFS_A}{CSFS_B} + b_2 * \frac{Amount_a}{Amount_b} + b_3 * \frac{Age_a}{Age_b} + b_4 * \frac{Venue_a}{Venue_b} + b_5 * \frac{Liquidity_a}{Liquidity_b} + \varepsilon$$

$$\frac{Yield_a}{Yield_b} = C' + b'_1 * \frac{CSFS_A}{CSFS_B} + b'_2 * \frac{Amount_a}{Amount_b} + b'_3 * \frac{Age_a}{Age_b} + b'_4 * \frac{Venue_a}{Venue_b} + b_5 * \frac{Liquidity_a}{Liquidity_b} + \varepsilon$$

Where:

$CSFS$: The Cash Flow Structures factors, including coupon type, frequency, etc.,

$Liquidity$: The market liquidity proxy. It can be the average bid-ask spread, the zero-trading days, or the Amihud ILLIQ. Finally, been mitigated in the pair data approach

$Amount$: Bond Issuing/Outstanding Amount at time t

Age : Bond Age, measured in percentage