ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

LIQUIDITY IN EMERGING COUNTRY BOND MARKETS: A COMPARISON OF BID-ASK SPREAD PROXIES AND DETERMINANTS OF BID-ASK SPREAD

Ph.D. THESIS

Emre SU

Department of Management Engineering

Management Engineering Programme

APRIL 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

LIQUIDITY IN EMERGING COUNTRY BOND MARKETS: A COMPARISON OF BID-ASK SPREAD PROXIES AND DETERMINANTS OF BID-ASK SPREAD

Ph.D. THESIS

Emre SU
(507142003)

Department of Management Engineering

Management Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Kaya TOKMAKÇIOĞLU

APRIL 2023

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

GELİŞMEKTE OLAN ÜLKE TAHVİL PİYASALARINDA LİKİDİTE: ALIŞ-SATIŞ YAYILMA ARALIĞI GÖSTERGELERİNİN KARŞILAŞTIRILMASI VE ALIŞ-SATIŞ YAYILMA ARALIĞININ BELİRLEYİCİLERİ

DOKTORA TEZİ

Emre SU
(507142003)

İşletme Mühendisliği Anabilim Dalı

İşletme Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Kaya TOKMAKÇIOĞLU

NİSAN 2023

Emre SU, a Ph.D. student of İTU Graduate School student ID 507142003 successfully defended the thesis entitled "LIQUIDITY IN EMERGING COUNTRY BOND MARKETS: A COMPARISON OF BID-ASK SPREAD PROXIES AND DETERMINANTS OF BID-ASK SPREAD", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Kaya TOKMAKÇIOĞLU** Istanbul Technical University	....
Jury Members :	**Prof. Dr. Oktay TAŞ** Istanbul Technical University	....
	Prof. Dr. Sudi APAK Esenyurt University	....
	Assoc. Prof. Dr. Cumhur Enis EKİNCİ Istanbul Technical University	....
	Assist. Prof. Dr. Umut UĞURLU Bahcesehir University	....

Date of Submission : 16 January 2023
Date of Defense : 12 April 2023

To my family,

FOREWORD

I am grateful to Assoc. Prof. Dr. Kaya TOKMAKÇIOĞLU for his valuable support, guidance and encouragement in this thesis. I would like to express my sincere gratitude for all of his contributions, advice, criticism and insight throughout the research. I am also grateful to my thesis jury members, Prof. Dr. Oktay TAŞ and Prof. Dr. Sudi APAK, for their expertise and recommendations.

I would like to thank my wife Özüm for her endless support and encouragement all through this thesis and my life. I sincerely would like to thank my mother and father for their support.

April 2023

Emre SU
(İşletme Mühendisi)

TABLE OF CONTENTS

Page

ABBREVIATIONS

AIC	: Akaike Information Criteria
BIC	: Bayesian Information Criteria
BIST	: Borsa Istanbul
CHL	: Close-High-Low
CPQS	: Closing Percent Quoted Spread
CS	: Corwin-Schultz
DSM	: Debt Securities Market
EHP	: Edwards-Harris-Piwowar
ETF	: Exchange Traded Fund
EU	: European Union
FHT	: Fong-Holden-Trzcinka
HL	: High-Low
IQR	: Interquartile Range
IRT	: Imputed Roundtrip Trades
LC	: Local Currency
LOT	: Lesmond-Ogden-Trzcinka
MAE	: Mean Absolute Error
MBE	: Mean Bias Error
OLS	: Ordinary Least Squares
RMSE	: Root Mean Squared Error
RS	: Relative Spread
RT	: Round Trip
OTC	: Over the Counter
TWS	: Time Weighted Spread
US	: United States of America

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIQUIDITY IN EMERGING COUNTRY BOND MARKETS: A COMPARISON OF BID-ASK SPREAD PROXIES AND DETERMINANTS OF BID-ASK SPREAD

SUMMARY

Many of the emerging market countries have been issued large amounts of local currency bonds in order to finance their budget deficits. Therefore, the attention to the emerging country bonds increased and, different aspects of these assets are investigated in order to analyse market microstructures. Liquidity is one of the most important characteristics of the financial assets and markets. As liquidity affects asset returns, investment decisions and market characteristics, all of the market participants are attracted to measurement and forecast of a financial security’s liquidity. So, liquidity measurement is essential to understand market dynamics. The bid-ask spread, which is the difference between the best bid and ask prices, represents the explicit trade costs, and is an important liquidity measure.

In this thesis, we look for an efficient way to estimate the bid-ask spread for sovereign bonds, and analyze the bond-specific and country-specific determinants of the bid-ask spread in the emerging market countries. The performances’ of the bid-ask spread measures are compared, and the bond-level characteristics' effects on the bid-ask are investigated.

We implement empirical analysis on two distinct data sets. For the first data set, we analyze Turkish sovereign bonds traded in Borsa Istanbul Debt Securities Market. We build a relative quoted bid-ask spread as a benchmark using intraday quote data, and compare the forecast performances of the bid-ask spread measures. According to the regression results, the spread measures that use low-frequency data are able to represent and estimate spread dynamics, and Closing Percent Quoted Spread outperforms the other measures. Moreover, a panel data analysis is applied to investigate the bond specific features’ effects on the bid-ask spread, and the regression analysis shows that bond liquidity is significantly affected by bond specific features. The shorter term to maturity or the greater trade volume, the narrower relative bid-ask spread. Further, bond type is significantly influential on the bid-ask spread.

For the second data set, we analyze sovereign bonds from six emerging market countries. A panel regression model is implemented in order to examine the country-level and bond-level features’ effects on the sovereign bond bid-ask spread and liquidity. According to our empirical analysis, issuer country credit risk, market return and bond yield are statistically significant determinants of the bid-ask spread for emerging country sovereign bonds.

GELİŞMEKTE OLAN ÜLKE TAHVİL PİYASALARINDA LİKİDİTE: ALIŞ-SATIŞ YAYILMA ARALIĞI GÖSTERGELERİNİN KARŞILAŞTIRILMASI VE ALIŞ-SATIŞ YAYILMA ARALIĞININ BELİRLEYİCİLERİ

ÖZET

Bir finansal varlığın likit olması mevcut piyasa fiyatını önemli ölçüde etkilemeksizin, düşük maliyetle, hızlı ve istenilen zamanda işlem yapılabilmesi olarak tanımlanmaktadır. Likidite, uygun maliyetle makul bir zaman içerisinde işlem gerçekleştirebilme yeteneği olarak da tanımlanabilir. Etkin ve verimli bir piyasa için likidite son derece önemlidir. Ayrıca, likidite varlıklardan elde edilen getiriyi de etkilemektedir. Bu sebeplerden dolayı yatırımcılar, ihraççılar, aracı kurumlar, işlem platformları, piyasa düzenleyicileri ve denetleyici kuruluşlar için likiditenin varlığı ve doğru bir şekilde ölçülebilmesi önem arz etmektedir.

Her ne kadar likidite kavramından bahsetmek ve tanımlamak kolay olsa da, likiditenin doğru bir şekilde ölçümü ve farklı boyutlarını kavrayabilme konusu oldukça zordur. Likiditenin farklı boyutlarından bahsedilmiş olmakla birlikte, alış-satış yayılma aralığı (bid-ask spread) işlem maliyetlerini direkt yansıtan bir özellik olduğu için oldukça önemlidir. Bu çalışmada likiditenin bu yönü üzerinde durulmuştur.

Likidite ve alış-satış yayılma aralığının ölçümünün önemi dolayısıyla literatürde çok sayıda model ve teori geliştirilmiştir. Bu modeller kullanılan veri sıklığına göre sınıflandırılabilir. Yüksek sıklıklı veri kullanan modeller güniçi emir ve işlem bilgilerini kullanırken, düşük sıklıklı veri kullanan modeller ise günlük, haftalık veya aylık sıklıktaki işlem ve emir bilgilerini kullanmaktadır.

Modellerin sayısının artmasıyla birlikte bu modellerin performanslarını karşılaştıran çok sayıda çalışmalar yapılmıştır. Geliştirilen modeller öncelikli olarak hisse senedi piyasaları için geliştirilmiş ve test edilmiştir. Ancak son yıllarda sabit getirili menkul kıymet, emtia, borsa yatırım fonu ve döviz piyasaları başta olmak üzere farklı piyasalar için de mevcut modeller kullanılmış ya da farklı piyasa dinamiklerine uygun olacak şekilde mevcut modeller modifiye edilmiştir. Ve yine bu piyasalarda ölçüm modellerinin performansları karşılaştıran çalışmalar da mevcuttur. Mevcut çalışmaların sonuçlarına göre tüm piyasalar ve varlık tipleri için her zaman en iyi sonucu veren bir model mevcut değildir, modellerin performansları farklı zamanlarda ve farklı piyasa yapılarında değişiklik göstermektedir. Ayrıca, bildiğimiz kadarıyla böyle bir performans karşılaştırması gelişmekte olan ülkelerin sabit getirili menkul kıymet piyasaları için yapılmamıştır.

Global finans krizi sonrasında gelişmekte olan ülkelere artan sermaye akışı sonrasında, bu ülke hazinelerinin sabit getirili menkul kıymet ihraçlarında önemli bir artış söz konusudur. Bu piyasalara ve sabit getirili menkul kıymetlere artan ilgi dolayısıyla bu varlıklar için likidite ölçümü de önem kazanmıştır. Bu çalışma kapsamında farklı alış-satış yayılma aralığı ölçüm modellerinin tahmin performansları karşılaştırılmış ve ek olarak alış-satış yayılma aralığını etkileyen değişkenler incelenmiştir. Alış-satış

yayılma aralığı tahmin modellerinin performansları ve alış-satış yayılma aralığını etkileyen değişkenlerin incelenmesi kapsamında iki farklı veri seti kullanılmıştır.

İlk veri setinde Borsa İstanbul Borçlanma Araçları Piyasası'nda işlem gören devlet tahvilleri incelenmiştir. En yüksek işlem hacmine sahip devlet tahvillerine ait gün içi en iyi alış ve satış fiyat farklarının zaman ağırlıklı ortalamaları kullanılarak, gerçek zamanlı bir alış-satış yayılma aralığı göstergesi hesaplanmıştır. Ayrıca her bir kıymet için Round Trip, Close-High-Low, Roll, Fong-Holden-Trzcinka, Corwin-Schultz, Interquartile Range ve Closing Percent Quoted Spread yöntemleri ile alış-satış yayılma aralığı tahmin edilmiştir. Aylık tahmin sonuçlarının performansları için zaman serisi ve kesitler arası regresyon modelleri uygulanmıştır. Korelasyon katsayılarının yanı sıra, ortalama hata (MBE), ortalama mutlak hata (MAE) ve ortalama hata kareleri karekökü (RMSE) ölçütleri kullanılarak modellerin performansları karşılaştırılmıştır. Ayrıca aynı veri seti ile tahvil özelliklerinin alış-satış yayılma aralığı üzerindeki etkileri incelenmiştir. Bu kapsamda işlem hacmi, vadeye kalan süre, dolaşımdaki tahvil tutarı, piyasa yapıcısının varlığı, valör günü ve tahvil tipi değişkenleri panel regresyona bağımsız değişkenler olarak dâhil edilmiştir.

İkinci veri setinde ise 6 farklı gelişmekte olan ülkeden 33 adet yerel para birimi üzerinden ihraç edilmiş devlet tahvili yer almaktadır. Bu tahviller için de gün içi en iyi alış ve satış fiyat farklarının ortalamaları kullanılarak, günlük alış-satış yayılma aralığı hesaplanmıştır. Ardından, varlık takas yayılma aralığı, durasyon, tarihsel oynaklık, baz puan değeri, getiri ve ihraççı ülke kredi temerrüt takası gibi bağımsız değişkenler kullanılmış, ve bu değişkenlerin alış-satış yayılma aralığı üzerindeki etkileri panel regresyon analizi ile incelenmiştir.

Sonuçlara göre düşük sıklıklı veri kullanan tahmin yöntemlerinin gerçek zamanlı alış-satış yayılma aralığı tahmininde etkili olduğu söylenebilir. Bu tip yöntemlerle tahminlerdeki tutarlılık belirli bir oranda azalsa da, daha az veriye ihtiyaç duyması ve daha az hesaplama yapılması, düşük sıklıklı veri kullanan tahmin yöntemlerinin kullanılmasını avantajlı hale getirmektedir. Diğer bir ifadeyle hesaplama ve veriye erişim kolaylıklarının getirdiği faydanın, tahmin doğruluğundaki azalmaya üstün geldiği söylenebilir. Ayrıca, günlük veri kullanımının alış-satış yayılma aralığı tahmininde yeterli olduğu da söylenebilir. Hem zaman serisi hem de kesitler arası regresyon sonuçlarına göre Closing Percent Quoted Spread yönteminin diğer modellerden daha başarılı tahmin sonuçları elde ettiği görülmüştür.

Ayrıca tahvil özelliklerinin de alış-satış yayılma aralığı üzerinde önemli etkileri vardır. İlk veri setine göre işlem hacmi, vadeye kalan süre ve tahvil tipi alış-satış yayılma aralığı üzerinde önemli derecede etkilidir. Daha uzun vadeye kalan süre veya daha düşük işlem hacmi, daha geniş alış-satış yayılma aralığına sebep olmaktadır. Ayrıca değişken faizli tahviller ortalama olarak daha dar alış-satış yayılma aralığına sahiptir. İkinci veri setine uygulanan panel regresyon sonuçlarına göre ihraççı ülke kredi temerrüt takas primi, ihraççı ülke piyasa getirisi ve tahvil getirisinin alış-satış yayılma aralığı üzerinde önemli etkileri mevcuttur. İhraççı ülke kredi temerrüt takasının yüksek olması, piyasa getirisinin yüksek olması veya tahvil getirisindeki azalış alış-satış yayılma aralığı üzerinde artırıcı etki yapmaktadır.

Düşük sıklıklı veri kullanan tahmin yöntemlerinin kullanımı ve tahvil özelliklerinin alış-satış yayılma aralığı üzerindeki etkilerine ilişkin elde edilen sonuçlar literatürde mevcut olan çalışmaları büyük ölçüde desteklemektedir. Literatürde farklı varlık tipleri ve pazarlar hakkında mevcut olan sonuçların, gelişmekte olan ülkelerin sabit getirili menkul kıymet piyasaları için de büyük ölçüde geçerli olduğu söylenebilir.

Bildiğimiz kadarıyla, literatürde finansal varlık ve ülke seviyesindeki değişkenlerin likidite ve alış satış yayılma aralığı ile ilişkilerini irdeleyen bir takım ampirik çalışmalar bulunmakta ve bunlar genellikle hisse senedi piyasalarını ya da gelişmiş ülkeleri kapsamaktadır. Bu çalışma ile gelişmekte olan ülkelerin sabit getirili menkul kıymet piyasalarında likiditeyi etkileyen olguları ampirik olarak irdelemekteyiz. Elde edilen sonuçlar tahvil ve ülke seviyesindeki değişkenlerin etkilerini göstermektedir. Piyasa düzenlemeleri, portföy yönetimi ya da işlem stratejilerine yönelik olarak sonuçlar belirli noktalara ışık tutabilecektir. Ayrıca, likidite ölçümü konusunda kolaylık sağlayacak yöntemlere yönelik olarak ampirik kanıt sunulmaktadır.

1. INTRODUCTION

Liquidity is an important aspect of the financial markets, and it is crucial for every kind of financial assets, especially during rough times. Even it is easy to define liquidity, liquidity is not obviously apparent for every asset every time. So, liqudity measurement is not always readily available. Additionally, liquidity plays an important role in the decision making processes of investors, traders, issuers, regulators, market makers, and exchanges. Hence, measuring liquidity accurately has great importance, and the accuracy of liquidity measurement is vital for the market participants.

Liquidity has many aspects, and most of these aspects are not observable every time. However, the bid-ask spread is an explicit measure of transaction costs. It is the price difference of the best bid (buy) and ask (sell) orders at a certain point in time for a given security, and represents execution cost of sequential buy and sell trades of equal amounts. Amihud and Mendelson (1986) say that the bid-ask spread is a natural measure of illiquidity. Additionally, Fleming (2001) asserts that the bid-ask spread is convenient for surveillance and appraisal of the US Treasury market liquidity.

Even though, the bid-ask spread of a financial security could be calculated with high-frequency data, availability of the required data could be restricted or access to such data could be costly. Thereupon, many kind of proxies that does not require high-frequency data are produced to reduce calculation burden and dependency on high-frequency data. There are plenty of liquidity measurement models in the literature. The models require different kinds of data, and are established on distinct underlying rationales. Therefore, the low-frequency methods are deeply investigated in order to discover how effectively they can measure liquidity and whether they are able to substitute the high-frequency models.

Most of the existing empirical researches in the literature, which predict the bid-ask spread, focus on stock markets. Due to the data availability and conveinence of stock markets, it is reasonable that the focus is on stock markets. On the other hand, the volume and market capitalization of the bond markets are greater than those of stock markets. Also, due to the increased data availability and importance of bond markets, the number of researches on the bond markets has increased. As the stock markets differ from the bond markets because of their characteristics and market structures, implementing of such measures in the bond markets may be questionable and effective estimation of the bid-ask spread may not be possible.

Firstly, the stocks are generally centrally traded and stock trades occur at organized exchanges, while ratio of bond trades in organized exchanges are quite lower, since OTC trades are more commen for bonds. Data availability of the OTC bond markets are mostly restricted, and even though some of the bond trades take place in an organized platform, the high-frequency data availability could be limited. Additionally, retail trades are more widespread in stock markets, whereas bond trades are mostly between institutional investors or dealers. Furthermore, an issuer typically have a great number of bonds, and each of these bonds has distinct characteristics, like coupon interest, coupon frequency, and maturity date. On the other hand, an issuer mostly have a single type of traded stock, and the stocks are homogeneous. Moreover, many bond investors buy and hold these securites instead of trading frequently. Buy and hold strategy of bonds reduces the number of available security that is tradeable. Due to market structures, it is reasonable that the total number of trades of a bond, on average, is less than the number of trades of a stock. For example, according to Darbha and Dufour (2013), only a few trades per security occur in a day, even for the most liquid bonds, in the largest electronic government fixed-income securities trading platform of Europe, Mercato dei Titoli di Stato, but the aggregate volume of fixed-income securities typically is greater than the aggregate volume of stocks. Because of lower data availability and ease of calculation, the low-frequency bid-ask spread proxies could be preferred for fixed-income securities. On the other hand, the models in the literature are mainly developed for stock markets, so using these models in centrally organized or OTC fixed-income securities markets may be doubtful.

On the other hand, many studies show that the low-frequency measures are able to measure liquidity properly for the bond markets (Diaz & Escribano, 2019; Langedijk et al., 2018; Schestag et al., 2016). Additionally, Diaz and Escribano (2017) assert that using the low-frequency measures in the corporate bond markets is economically and statistically significant. Moreover, Jahan-Parvar et al. (2019) report that the low-frequency measures are significantly biased upwardly for financial securities, which have lower volatility and higher transaction costs, such as bonds. Further, Chordia et al. (2004) assert that the correlations between bond and stock liquidities are significantly positive, and the patterns of these markets' liquidity movements are analogous. In other words, different asset types may have similar liquidity characteristics. According to the existing literature, the bid-ask spread measurement models which are mainly established for the stock markets may also be employed in the bond markets. Nevertheless, due to the differences in data and market structures, results may be biased. So, model selection must be performed diligently, and if there is any requirement for a modification in the models, the modifications must be implemented properly.

Different spread prediction methods are created in order to fulfill distinct needs, such as transaction cost analysis or liquidity measurement. Diaz and Escribano (2020) examine diffent liquidity measures and dimensions of liquidity, and report that there is no single measure that could be called as a solely liquidity indicator. Additionally, they assert that model selection must depend on the research question and asset characteristics, as there is no universal liquidity measure. Different market specifications, trade dynamics, and asset characteristics lead to discrepancy of the measures' performances, and there is no measure that is able to suit different asset classes.

There have been significant enhancements in the emerging country sovereign bond markets within the last decade. Many of the emerging country treasuries have issued significant amounts of fixed-income securities in order to finance their budget deficits. Particularly after the global financial crisis, capital flows have been steered into the emerging market countries. Since the market size increases and investors' focus head towards to those markets, identification of the liquidity determinants and measuring liquidity efficiently become more of an issue.

Additionally, there are plenty of researches in the literature which try to identify and model the relation of financial security features and its liquidity for different asset classes. However, as far as we know, none of the existing literature focuses on the emerging country sovereign bond markets. Therefore, we examine this relation for six emerging countries, so this study targets to contribute to the literature of bond liquidity in the emerging countries via finding efficient ways to measure liquidity and understand its determinants. The previous evidence about the liquidity on bond yield spreads and returns (Lesmond, 2005) indicates the significance of understanding the bond-level determinants of the bid-ask spread and liquidity.

The main motivation of this thesis is to compare the high-frequency and low-frequency bid-ask spread measures' performance, and to exhibit whether the low-frequency proxies are able to replace the high-frequency measures. Also, the bond characteristics’ effects on the bid-ask spread are analyzed for the emerging country fixed-income securities markets.

The findings of this thesis may be meaningful for fixed-income securities markets participants, such as investors, issuers, and regulators. Since the implicit and explicit transaction costs affect return on an investment, analysis of liquidity and its costs in the secondary markets become a major issue for the investors. Additionally, as the trade costs in the secondary markets have effects on the investors’ decisions, liquidity have impact on the cost of borrowing for the issuers. So, the issuers consider the secondary market structures and liquidity as significant attributes. Also, as the regulators’ objective is to make markets efficient and sustain fair competition, they investigate liquidity in the secondary markets to explore the relationship between market structure and efficiency.

The rest of this thesis is structured as follows: Section 2 includes the literature review on liquidity and introduces the existing bid-ask spread measures, Section 3 describes the data set and empirical methodology. Section 4 exhibits the results of models decribed in the previous section, and Section 5 concludes the study.

2. LITERATURE REVIEW

2.1 Concept of Liquidity

Foucault et al. (2013) describe liquidity as the degree of how an order could be executed at fair value within a reasonable time period. According to Stange and Kaserer (2009), liquidity is the ease of trade for a financial security. They assert that the liquidity term can be used in three distinct meanings. Firstly, it may be used to indicate the liquidity of a firm, which is also called solvency. The term "funding liquidity" can be used to refer to the liquidity of liability side of a company's balance sheet. Second use of liquidity is called "asset liquidity" or "market liquidity" which designates a characteristic of a tradable asset. From an investor's perspective, liquidity represents the ease of trading an asset. Thirdly, the liquidity term addresses the liquidity of whole economy.

Foucault et al. (2013) refer funding liquidity of a company as having enough cash or the ability to acquire credit in order to meet a company's obligations without large losses; and they denote monetary liquidity as the monetary base or broader monetary aggregates. They also mentioned about the relation between funding liquidity and market liquidity in several ways: Firstly, both of them are worthy for the same reason, people hold securities which can be immediately liquidated to meet consumption requirements; secondly, funding liquidity is a precondition for market liquidity, since the greater funding liquidity of a market maker, the greater market liquidity of a security; thirdly, more liquid markets enable traders to fund leveraged purchases or short sales cheaper which creates a feedback from market liquidity to funding liquidity. Additionally, they state that monetary liquidity has relationship with funding liquidity and market liquidity; expansionary (contractionary) monetary policy of a central bank increases funds available to banks which tends to increase (decrease) funding liquidity, thus market liquidity.

Kyle (1985) describes liquidity as a slippery and elusive concept. Similarly, Baker (1996) asserts that none of the liquidity definitions is universally accepted or theoretically correct, so there is no single definition of liquidity. Schwartz and Francioni (2004) say that a liquid asset is in cash or is readily convertible into cash.

Lybek and Sarr (2002) mention about four different types of liquidity: asset liquidity, an asset's market liquidity, a financial market's liquidity, and liquidity of a financial institution. According to them, asset liquidity refers to how easily an asset can be converted into cash under normal circumstances, and the emphasis of this definition is on immediacy and transactions costs; while an asset's market liquidity has a broader meaning, which is related to how easily large amount of a security can be disposed of rapidly at a reasonable price, in the absence of informative shock that changes the security's fundamental price. A financial market's liquidity refers to how liquid the each asset traded in a market and the interaction of these assets in terms of being substitutes, and the liquidity of a financial institution refers to how flexible an institution to cover its asset-liquidity mismatch by conducting financial transactions. In this thesis, the term of liquidity refers to asset liquidity.

O'Hara (2007) reports that the markets supply liquidity and price discovery, and liquidity problems occur due to the non-synchronicity of buyers and sellers. She also says that the market microstructure straightly has effect on asset returns via its impact on liquidity. Hence, a better market is able to supply liquidity less expensively, and let the buyers to pay less and the sellers to receive more for an asset.

Black (1971) says that a liquid market is a continuous and efficient market. In other words, in a liquid market, almost any amount of a security can be traded immediately, transactions with small amounts can be executed at prices very close to the market price, and large transactions can be executed over long periods of time at close prices to market prices on average.

Kyle (1985) mentions about three properties of the market liquidity: Tightness, the cost of executing a transaction in a short period of time; depth, the volume available to absorb quantities without creating a large effect on the market price; resiliency, how quickly the prices can recover from an unexpected shock. In addition to these three characteristics, Lybek and Sarr (2002) mentioned about two other characteristics of market liquidity; immediacy, the speed at which orders can be executed and represents the efficiency of a trading system; and breadth means that there are enough number and volume of orders which have very low effect on the prices.

Aquilina and Suntheim (2016) state that liquidity is a result of the relation between different features of a market; how the securites are traded, which kinds of participants involve in trading activity, and the characteristics of the securities. For example, they say that the bond market liqudity is expected to be lower than the equity market liquidity, since fixed income securities typically are less standardised than equities, and investors mostly hold such securities until maturity. Additionally, Benos et al. (2020) report that a centralized clearing leads to greater liquidity and narrower bid-ask spreads, as a result of increased dealer competition.

According to Amihud et al. (2006) the exogenous transaction costs are among the sources of illiquidity; when a security is traded, transaction costs, such as brokerage fees, order processing costs or taxes, occur for the buyer and the seller; additionally, the buyer awaits for an additional transaction cost in case of a possible future sale. They also denote that demand pressure and inventory risk are the sources of illiquidity. Demand pressure occurs since not all of the agents are available in the market every time, which means if a market participant have to liquidate a security in a short period of time, the usual buyers of that security may not be accessible to execute the sell order. Hence the seller may prefer to sell the security to a market maker, who is exposed to price change risk while he holds the security in his inventory. So, a market maker requires a compensation for such a risk, which increases the cost on the seller. They also say that trading may be costly due to a private information that the counterparty of the trade may have, such private information may be about the fundamentals of a security or the order flow.

Additonally, Amihud et al. (2006) say that the security prices should be affected by the illiquidity costs if investors demand extra return for bearing illiquidity, and risk averse investors may demand extra return due to the exposure to liquidity risk, since liquidity may change over time. They report that such influences of liquidity on asset prices are important, since investors require them to build investment strategies, and the cost of capital and the allocation of real resources are partially determined by the liquidity costs and risks.

O’Sullivan and Papavassiliou (2020) assert that a country’s sovereign bond market liquidity significantly affects the CDS spreads, volatility, asset returns and liquidity not only in its own region but also in its interconnected bond markets. Additionally, Bampinas et al. (2020) report that the bond market liquidity risk is associated with the risks in the European CDS market.

2.2 Measuring Liquidity

As mentioned earlier, there is no single definition of liquidity and liquidity is described as a complex and intangible notion. Amihud et al. (2012) mention about liquidity and its opposite, illiquidity, as elusive concepts and say “you know it when you see it, but it is hard to define”. In other words, liquidity cannot be observed directly. Instead, different aspects (or characteristics or dimensions) of liquidity are defined, and a vast number of measurement techniques are developed in order to measure liquidity.

According to Foucault et al. (2013), liquidity measurement is quite important for not only practitioners, i.e. investors and intermediaries, but also researchers and regulators. They say that practitioners aim to minimize the liquidity effects on their investments, as lack of liquidity impacts portfolio returns, while researchers and regulators analyze the relationship between market structure and market performance. For example, Ho and Ho (2020) assert that market regulations improve liquidity for the investment grade bonds, while there is no significant effect for the non-investment grade bonds.

Because of the differences between liquidity measurement techniques and the distinct characteristics of market structures and mechanism, market features must be investigated cautiously. Additionally, as liquidity measurement techniques are mostly based on historical data, they produce results that depends on the historical events only. So, one should be aware of the usage of historical data and its limitations.

One of the most widely used measures of asset liquidity is the bid-ask spread, which shows the cost of immediate execution. Wyart et al. (2008) assert that the liquidity of an asset is partly characterized by the bid-ask spread, which shows the cost of an instant round-trip trade of the asset. Nevertheless, the bid-ask spread only shows the cost for a marginal trade and does not gives information about how large the volume can be executed at this level of cost.

Bollen et al. (2004) say that understanding the determinants of a market maker’s bid-ask spread is crucial for several reasons. The bid-ask spread helps to determine the market design, i.e. number of market makers, and optimal price tick. Additionally, for a regulatory body, the bid-ask spread shows the fairness of the market makers.

Foucault et al. (2013) assert that illiquidity is often measured by the trading costs, which have implicit and explicit components. Transaction taxes, brokerage commissions, trading fees, clearing fees, and settlement fees are the explicit costs, which are clearly known and easily measurable by the final investors. However, the implicit costs generally show up due to market illiquidity, and they are usually gauged by the difference between the execution price and the benchmark price, which is used as a substitute for the price that could be available in a perfectly liquid market.

Most of the liquidity measures can capture only a single aspect of liquidity. While some measures may combine several dimensions, there is not a measurement technique that can capture all dimensions of liquidity. As Lybek and Sarr (2002) mention there is apparently no single measure that can measure depth, breadth, tightness, immediacy, and resiliency.

Amihud et al. (2006) assert that all of the liquidity measures have an error due to several reasons: Firstly, a single measure is not able to capture all aspects of liquidity; secondly, the measurement techniques, which are empirically derived, lead to a noisy estimation of the real parameter; and thirdly, the low frequency data used for estimation increases the measurement noise. They report that these errors bias the estimated coefficients downward (towards zero), which means that it is hard to identify the effects of liquidity when it exists.

Lybek and Sarr (2002) categorize liquidity measures into four classification:

- Transaction cost measures, which gauge the costs of trading financial securities and frictions in the secondary markets,
- Volume based measures, which primarily measure breadth and depth, differentiate liquid markets by comparing transaction volume with price variability,
- Equilibrium price based measures, which mainly gauge resiliency, attempt to capture price movements towards to equilibrium prices,
- Market impact measures, which measure resiliency and speed of price discovery, try to differentiate price movements due to the degree of liquidity.

Schestag et al. (2016) classify liquidity measures as high-frequency measures and low-frequency proxies; the high-frequency measures are based on intraday data, while the low-frequency proxies use daily price, quote and volume data.

2.3 The Bid-Ask Spread

Amihud and Mendelson (1986) assert that the cost of immediate execution is a measure of illiquidity, and the bid-ask spread is a natural measure of illiquidity. They say that a tradeoff occurs when an investor wants to transact a security: He has to wait until a favorable price appears in the market, or he executes at the available market prices. The quoted ask price contains a premium for immediate execution, and the quoted bid price includes a concession for immediate execution. The difference between the bid and ask prices, which is called as the bid-ask spread, is a natural measure of illiquidity, and it is negatively correlated with the certain liquidity characteristics.

Amihud and Mendelson (1986) study the effect of the bid-ask spread on the returns of stocks, and find out the following characteristics: Average returns are an increasing function of the bid-ask spread, asset returns increase with the bid-ask spread (net of trading costs), securities with higher bid-ask spread are held by the investors with longer holding periods, i.e. there is clientele effect exists, and the returns on higher-spread securities are less spread-sensitive, that produces a concave shaped return-spread diagram.

Amihud and Mendelson (1986) also say that the relationship between higher-yield and higher-spread encourages firms to increase the liquidity of their securities in order to reduce their cost of capital. Therefore, increasing liquidity of a firm’s traded securities may boost the firm value.

Amihud (2002) shows that stock returns are positively correlated with illiquidity, and market illiquidity increases ex ante excess stock return. This argument may also be valid for the fixed income securities. Chen et al. (2007) find that the liquidity of a bond is the key determinant of the bond yield spread, whether the bond is investment-grade or not.

2.4 Measures of the Bid-Ask Spread

2.4.1 Quoted spread

The quoted bid-ask spread is the difference between the best ask quote and the best bid quote available in the market. The quoted spread, *S*, is represented in equation 2.1.

$$S = a - b \tag{2.1}$$

where a is the best ask quote and b is the best bid quote.

According to Foucault et al. (2013), the quoted spread is an appropriate measure of trading costs for relatively small orders that can be filled entirely at the best price levels, it is the most widely reported liquidity measure, and it is what people really mean when referring to the bid-ask spread.

2.4.2 Relative quoted spread

The relative quoted spread is calculated via normalisation of the quoted spread, S, by the midprice, *m*, as in equation 2.2.

$$s = \frac{S}{m} = \frac{a - b}{(a + b)/2} \tag{2.2}$$

2.4.3 Weighted average spread

For large orders which cannot be entirely filled at the best bid or ask prices, the quoted spread may result in misinformation. Instead of the relative quoted spread, the weighted average of bid-ask spread would be more useful and appropriate. Moreover, implementation of a relative spread can help to grasp comparable results. The relative weighted average spread is calculated as represented in equation 2.3.

$$s(q) = \frac{\bar{a}(q) - \bar{b}(q)}{m} \tag{2.3}$$

where $\bar{a}(q)$ is the average execution price for a buy market order of size q, $\bar{b}(q)$ is the average execution price for a sell market order of size q, and m is the midprice.

In order to calculate the weighted average spread, order book data and quotes posted at various points in time are required, as such information is not always available, practitioners and researchers use measures based only on transaction prices and the quoted spread (Foucault et al., 2013).

2.4.4 Effective spread

The quoted spread or the other spread measures based on the quoted spread represent a snapshot of the liquidity available for a theoretical transaction, while the effective half-spread, which is the difference between the trade price of an executed market order and the midquote just before the trade, uses the actual obtained prices in the market (Foucault et al., 2013).

The effective half-spread is calculated as in equation 2.4, and the relative effective spread is defined as seen in equation 2.5.

$$S_e = d(p - m) \tag{2.4}$$

$$s_e = 2d\frac{p - m}{m} \tag{2.5}$$

where *p* is the execution price, *m* is the midquote just before the transaction and *d* is the order direction indicator that equals to “-1” for market sell orders and “1” for market buy orders (Blau et al., 2018).

2.4.5 Realized spread

According to Foucault et al. (2013), if the prices move in the direction of the trade after the trade, quoted spread and effective bid-ask spread are likely to overestimate liquidity demanders’ trading costs. In order to overcome that problem, they propose the realized half-spread, which is the difference between trade price and midprice at some time (Δ) after the trade. The time interval (Δ) should be long enough to make sure that the quotes in the market are adjusted to reflect the price impact of the trade. The realized half-spread is calculated as represented in equation 2.6 and equation 2.7.

$$S_r = d_t(p_t - m_{t+\Delta}) = d_t(p_t - m_t) - d_t(m_{t+\Delta} - m_t) \tag{2.6}$$

$$E(S_r) = E(S_e) - E(d_t(m_{t+\Delta} - m_t)) \tag{2.7}$$

where p_t is the transaction price, d_t is the direction of market order, and m_t is the midprice at time *t*.

According to equation 2.6 and equation 2.7, if the change in the midprice just after the the trade is positively correlated with trade direction, the average realized spread is narrower than the average effective spread (Foucault et al., 2013).

When the data required to calculate the midprice is unavailable, accessible data alternatives can be used. Daily close price, daily opening price or daily average prices are such alternatives. For example, volume weighted average price is calculated as in equation 2.8.

$$VWAP = \frac{Total\ LC\ value\ of\ trades}{Total\ traded\ quantity} = \sum_{t \in T} w_t p_t \tag{2.8}$$

where $w_t = \frac{|q_t|}{\sum |q_t|}$, p_t is the price of trade, q_t is the quantity of trade, and T is the chosen time interval.

2.5 Estimation of the Bid-Ask Spread

In the literature, there are plenty of models, which analyze and estimate the bid-ask spread. Based on the data frequency used, the bid-ask spread measures can be classified into two groups; low-frequency and high-frequency measures. Low-frequency measures require daily or less frequent data, such as daily high, low, close prices, quote, or volume data. The ability to forecast bid-ask spread accurately with lower data requirements and calculation burden makes the low-frequency measures favourable, since liquidity measurement is vital for the market participants. The low-frequency data is used in order to compute some metrics which are liquidity proxies, and these metrics approximate a financial security's liquidity. Whereas, the high-frequency measures depend on intraday data, such as order flow, intraday trade price, intraday quote, or trade-based volume information. These measures immediately gauge the real-time bid-ask spread of an asset available in the market.

Roll (1984) pioneered the low-frequency proxies by using serial covariances of asset returns in order to estimate the bid-ask spread. Since the model only relies on trade price data, it is advantageous and easy to use. However, it is criticised due to the drawbacks in long-term estimation performance.

According to Roll's model, the orders are not serially correlated, the probability of an order being buy or sell is equal. Choi et al. (1988) developed Roll's model by loosing the assumption of conditional probability of consecutive transactions at bid or ask being equal to 0.5.

Stoll (1989) said that Roll’s model estimates the realized spread but does not examine the relationship between quoted spread and realized spread. He modelled time series behavior of the spread, and examined the relation between realized spread and quoted spread. He added magnitude and probability of price reversal to the model in order to break down the bid-ask spread into three components; order-processing costs, inventory-holding costs, and adverse information costs.

George et al. (1991) extended the previous models by including expected returns into their model. They assumed that the probabilities of trades being at bid or ask prices are equal, and the spread does not depend on trade size. They showed that the previous estimators of the spread components are biased because of the positively auto-correlated time-varying expected returns. To introduce an unbiased and efficient model, they constructed a spread measure relying on the serial covariances of the differences between transaction returns and returns calculated using bid prices.

Hasbrouck (2004) applied a Bayesian analysis using a Markov chain Monte Carlo on Roll’s model in order to include latent variables for futures market by using transaction data. The observed variables are the trade prices and volumes, while the latent variables are trade direction indicators, efficient prices and implicit price tick size. Hasbrouck (2009) also used Gibbs estimator model on stock markets with daily data, and found out that Gibbs estimator that use low-frequency data is highly correlated with the results obtained with the high frequency measures.

Holden (2009) used spread proxies that correspond to three attributes of daily data; price clusters, serial price covariance and quoted spread available on no-trade days. He modified Huang and Stoll (1997) model by allowing the effective spread to alter from day to day, and eliminating the inventory holding risk component.

Lesmond et al. (1999) found that frequency of zero return days is directly related to the quoted bid-ask spread and Roll’s effective spread. Their model is based on the assumption that a zero return day happens if the transaction cost is higher than the expected gain, i.e. the threshold.

Based on Lesmond et al.’s (1999) Zeros measure, Fong et al. (2017) developed a spread proxy, which assumes that the unobserved true value return is normally distributed with zero mean and a constant finite variance, and the proxy is modelled as a function of the return distribution volatility and the proportion of zero returns.

Goyenko et al. (2009) and Holden (2009) assumed that the trade prices cluster at certain price steps, and analysed such clusters to proxy effective spread. Their model is called as the Effective Tick.

Corwin and Schultz (2012) used daily high and low prices to approximate the bid-ask spread. While Chung and Zhang (2014) developed CPQS model that uses the bid and ask prices at daily close. Schestag et al. (2016) used the quoted bid-ask spread data to estimate transaction costs.

The bid ask spread measures, which use high frequency data, differentiate depending on the approaches which they are based on and the data they use.

The regression based approaches compare orders or trades with a benchmark (calculated) value. These models use all data available, not leaving significant amount of data out. However, they may not be able to perform well in case of intraday price movements or missing data (Schestag et al., 2016). On the other hand, measures like RT (Feldhüter, 2012) require both buy and sell order or trade data within a certain period of time, and they ignore large amounts of data which are not eligible as model input. Another type of the measures uses trade price distributions, and such models are properly work even there is only one sided (buy or sell) information.

Glosten and Harris (1988) aimed to separate the adverse selection component by using order flow and decomposed the bid-ask spread into two parts. The first part is the transitory component, which allows market makers to earn profit to cover inventory-holding costs, clearing fees or monopoly profits. This part is called transitory, since its effect on the security price is unrelated to the value of the security. The second part is a result of asymmetric information, and it arises in case of a trade against an anonymous counterparty with superior information. To compensate such a potential loss, a market maker widens the spread, and this marginal widening is called as the adverse-selection component.

Madhavan et al. (1997) extended Glosten and Harris' (1988) model in order to understand the intraday patterns of the bid-ask spreads, price volatility, transaction costs, and autocorrelations between returns and quote revisions. Their model incorporates microstructure effects and public information shocks.

Huang and Stoll (1997) estimated the bid-ask spread by modelling the unobservable fundamental value of a financial assetand decomposing it into two parts; the private information unveiled by the last trade, and the public information. They found that the spread components significantly depend on the trade size.

Ball and Chordia (2001) developed a model by generalizing Huang and Stoll's (1997) model. Their model makes possible to separately estimate the sum of order-processing and inventory holding components, and the adverse selection component. Also, the model can parse such costs from the effect of price step (or tick size). More importantly, the model allows the spread to change from trade to trade, and to estimate a varying spread depending on market depth and prior trade size.

2.5.1 High-frequency measures

2.5.1.1 Schultz model

Schultz (2001) estimated the round-trip transaction costs in a corporate bond market by regressing the formula in equation 2.9.

$$\Delta_i = \alpha_0 + \alpha_1 D_i^{Buy} + \epsilon_i \tag{2.9}$$

where Δ_i is the difference between the estimated bid price and the price of trade *i*, and D_i^{Buy} is the dummy variable that equals zero if the trade *i* is a sell trade, and equals one if the if the trade *i* is a buy. α_1 is an estimate of price difference between buy and sell orders, i.e. it is an estimate of the average bid-ask spread.

Schestag et al. (2016) altered Schultz's model by using a relative scale and including interdealer trade information, which is represented in equation 2.10.

$$\Delta_i^{rel} = \alpha_o^{rel} + \alpha_1^{rel} D_i + \epsilon_i \tag{2.10}$$

where Δ_i^{rel} is the relative deviation of the mid quote, so α_1^{rel} estimates the half spread. Markit composite prices are used as the proxy of the mid quote. They adapted the dummy variable in order to capture interdealer trade information as in equation 2.11.

$$D_i = \begin{cases} 1 \; if \; the \; trade \; i \; is \; a \; buy, \\ 0 \; if \; the \; trade \; i \; is \; an \; interdealer \; trade, \\ -1 \; if \; the \; trade \; i \; is \; a \; sell. \end{cases} \tag{2.11}$$

2.5.1.2 Edwards-Harris-Piwowar (EHP) model

Harris and Piwowar (2006) assumed that the trade price is equal to the true unobserved value plus a price concession or premium depending on whether the trade is a buy or sell. They also estimated the price concessions separately for customer and interdealer trades, and assumed that the absolute customer transaction cost is measured as a fraction of the trade price depending on the trade size. Their estimation of half-spread is represented in 2.12.

$$c(S_t) = c_0 + c_1 \frac{1}{S_t} + c_2 log S_t + \kappa_t \tag{2.12}$$

where $c(S_t)$ is the half spread, S_t is the trade size, and κ_t is the part that is unexplained by the average transaction cost function.

Schestag et al. (2016) developed another model based on Harris and Piwowar (2006) and Edwards et al (2007). They used a composite price for the unobserved true value, and assumed that the measurement error of the unobserved true value return between sequential trades is proportional to the time between two trades. They estimated the half spread, c, via iteration of equation 2.13.

$$r_i^{Obs} - r_i^{Composite} = c(D_i - D_{i-1}) + \eta_i \tag{2.13}$$

where $r_i^{Obs} = log\left(\frac{P_i}{P_{i-1}}\right)$, and η_i is the error term with mean zero and variance σ_i^2, which calculated as in equation 2.14.

$$\sigma_i^2 = \sigma_{\epsilon_i}^2 + (2 - D_i^{Int})\sigma_c^2 + D_i^{Int}\sigma_\delta^2 \tag{2.14}$$

where D_i^{Int} is the dummy that equals 0, 1 or 2 when none, one, or two of consecutive trades are interdealer trades. $\sigma_{\epsilon_i}^2$ is the variance from the measurement of the unobserved true value returns, σ_c^2 is the variance from the measurement of the transaction costs, and σ_δ^2 is the variance from the measurement of the interdealer price concessions.

2.5.1.3 Average bid-ask spread

Hong and Warga (2000) used the weighted average of transaction prices in a day in order to estimate transaction costs. Their model can be formulated as seen in equation 2.15.

$$Average\ Bid\ Ask\ Spread = \frac{\overline{P_t^{buy}} - \overline{P_t^{sell}}}{\left(\overline{P_t^{buy}} + \overline{P_t^{sell}}\right)/2} \tag{2.15}$$

where P_t^{buy} is the weighted average of buy trades within a day, and P_t^{sell} is the weighted average of sell trades within a day. The weights are calculated with the trade volumes.

2.5.1.4 Roundtrip transaction (RT) cost

Feldhütter (2012) developed a roundtrip transcation cost measure which is based on IRT. According to him, for securities with low numbers of transactions, if consecutive trades occur within a very short period of time after a longer period of time without a trade, this situation most probably is a result of a pair transaction of a dealer, and the price difference is the fee collected by the dealer. He defined the IRT as the combination two or more trades with same volumes which occur within 15 minutes. The round-trip transaction cost is the difference of the greatest and the smallest trade prices of an IRT. In order to have a relative measure, the spread is normalized with the average of the greatest and the smallest trade prices of an IRT. The measure uses high-frequency data. Intraday trade volume, trade price and trade time data are required for the model. The measure is given in equation 2.16.

$$RT = \frac{2(P_{max} - P_{min})}{(P_{max} + P_{min})/2} \tag{2.16}$$

where P_{min} is the minimum trade price in an IRT, and P_{max} is the maximum trade price in an IRT.

2.5.1.5 Roll's intraday measure

Dick-Nielsen et al. (2012) uses an intraday version of the Roll Model (1984) in order to estimate the effective spread, as seen in equation 2.17.

$$Roll_t = \begin{cases} 2\sqrt{-cov(r_i, r_{i-1})}, & if\ cov(r_i, r_{i-1}) < 0 \\ 0, & otherwise \end{cases} \tag{2.17}$$

where, $r_i = \frac{P_i - P_{i-1}}{P_{i-1}}$ is the return of *i*th trade.

2.5.1.6 Interquartile range

Han and Zhou (2007), and Pu (2009) built a proxy of the bid-ask spread that uses the quartiles of trade prices within a day. According to them, the price volatility within a day is due to the trades occur at the bid or ask prices, in other words, the bid-ask spread leads to price volatility. They defined the interquartile range as the difference between 75th percentile and 25th percentile of trade prices within a day, and then, they normalized the interquartile range by the average of daily price to obtain a relative measure. The IQR measure is given in equation 2.18. The model ignores half of the trade prices as the outliers, and only considers half of trade data.

$$IQR_t^i = \begin{cases} \frac{p_t^{i\prime 75th} - p_t^{i\prime 25th}}{\bar{p}_t^i} \end{cases} \tag{2.18}$$

where $p_t^{i\prime 25th}$and $p_t^{i\prime 75th}$ are 25th and 75th percentiles of trade prices in day i, respectively. $\bar{p}_t^i$ is the average of the trade prices within day *i*.

2.5.1.7 Huang and Stoll model

Huang and Stoll (1997) developed a model to decompose the bid-ask spread. They defined the unobservable fundamental values as represented in equation 2.19.

$$V_t = V_{t-1} + \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t \tag{2.19}$$

where S is the spread, α is the percentage of the half-spread for the adverse-selection cost (due to asymmetric information).

This equation decomposes *V(t)* into two parts, which are the private information unveiled by the last trade and the public information part. They formulized the relation between the midpoint of spread and the fundamental value as in equation 2.20.

$$M_t = V_t + \beta \frac{S}{2} \sum_{i=1}^{t-1} Q_t \tag{2.20}$$

where β is the component of the half-spread for the inventory-holding cost and, the transaction prices follow equation 2.21.

$$P_t = M_t + \frac{S}{2} Q_t + \eta_t \tag{2.21}$$

By combining equation 2.20 and equation 2.21, they modelled price changes in order to estimate the spread as represented in equation 2.22.

$$\Delta P_t = \frac{S}{2}(Q_t - Q_{t-1}) + \lambda \frac{S}{2} Q_{t-1} + e_t \tag{2.22}$$

They extended the model by adding trade size categories. The spread is estimated depending on trade size, as well, which is given in equation 2.23.

$$\begin{aligned} \Delta P_t = {} & \frac{S^s}{2} D_t^s + (\lambda^s - 1)\frac{S^s}{2} D_{t-1}^s + \frac{S^m}{2} D_t^m + (\lambda^m - 1)\frac{S^m}{2} D_{t-1}^m \\ & + \frac{S^l}{2} D_t^l + (\lambda^l - 1)\frac{S^l}{2} D_{t-1}^l + e_t \end{aligned} \tag{2.23}$$

2.5.1.8 Ball and Chordia model

Ball and Chordia (2001) developed a model by generalizing Huang and Stoll's (1997) model. Their model makes it possible to separately estimate the sum of order processing and inventory holding components, and the adverse selection component. Additionally, the model can parse such costs from the effect of price step (or tick size). More importantly, the model allows to estimate a varying spread depending on market depth and prior trade size. They defined the observed price as in equation 2.24, and they assumed that the true prices follow the formula in equation 2.25.

$$P_t \equiv [p_t^{NR}]_{Round} = [m_t + (1-\lambda) s_t Q_t / 2]_{Round} \tag{2.24}$$

$$m_t = m_{t-1} + \lambda \frac{s_t Q_t}{2} + u_t \tag{2.25}$$

They modelled the spread estimate as in equation 2.26.

$$\ln(s_t) = \alpha + \beta \ln(s_{t-1}) + \sigma \ln \frac{V_{t-1}}{D_{t-1}} + \tau TIME_{t-1} + d_1 BEG_t + d_2 END_t + e_t \tag{2.26}$$

where *S(t)* is the true spread, *D(t-1)* is the depth, and *V(t-1)* is the volume of the previous trade. They also added dummy variables to the model. *TIME* is the time difference between the last trade and the second last trade, and *BEG* and *END* are included in order to indicate the first hour and the last hour of a trading day.

2.5.2 Low-frequency measures

2.5.2.1 Roll's model

Roll (1984) developed a model that directly infers the effective bid-ask spread from the market prices. The main idea of the model is that a financial security's price bounces between the best bid and ask prices, so a higher negative covariance of trade prices points to a greater bid-ask spread. The model uses only daily prices, and the model has two major assumptions: The market is informationally efficient, and the probability distribution of the observed price changes is stationary, for at least short intervals.

In addition, Roll model has the following additional assumptions:

- It is likely to be equal for a market order to be a buy or sell order,
- There is no serial correlation between buy and sell market orders,
- Market orders do not contain news,
- The fundamental value of the security follows a random walk.

Serial covariance of daily equity returns is used, and Roll showed that the covariance of daily returns is an estimate of the effective spread. Roll's spread estimator is given in equation 2.27.

$$Spread = 2\sqrt{-cov(\Delta P_t, \Delta P_{t-1})} \tag{2.27}$$

where ΔP_t is the daily price change on day *t*.

When the covariance of daily price changes is positive, the spread is undefined according to the Roll's original model. Goyenko et al. (2009) extended the model by assigning zero to undefined values. The extended version is given in equation 2.28.

$$Roll_t = \begin{cases} 2\sqrt{-cov(\Delta P_t, \Delta P_{t-1})}, & if\ cov(\Delta P_t, \Delta P_{t-1}) < 0 \\ 0, & otherwise \end{cases} \quad (2.28)$$

According to Dick-Nielsen et al. (2012) and Schestag et al. (2016), in order to have a relative measure, daily returns should be selected instead of daily price changes. The version with daily returns is given in equation 2.29.

$$Roll_t = \begin{cases} 2\sqrt{-cov(r_t, r_{t-1})}, & if\ cov(r_t, r_{t-1}) < 0 \\ 0, & otherwise \end{cases} \quad (2.29)$$

where r_t is the daily return on day *t*.

2.5.2.2 Choi, Salandro and Shastri model

Choi et al. (1988) extended Roll's (1984) model by allowing positive serial correlation for transaction types. According to them, Roll assumed that transaction types are serially independent, i.e. regardless of the previous transaction type, the probabilities of conducting the following transaction at bid or ask price are equal. However, they asserted that there would be serial dependence in transaction type. So, they extended Roll's model in order to allow the conditional probability of executing both of consecutive transactions at bid (or ask) to differ than 0.5.

After giving a probability distribution, they formulized the covariance between the transaction price changes as in equation 2.30, and the spread in defined as given in equation 2.31.

$$Cov(\Delta p_t, \Delta p_{t+1}) = -s^2(1-\delta)^2 \quad (2.30)$$

$$s = \left[\sqrt{-Cov(\Delta p_t, \Delta p_{t+1})}\right] / (1-\delta) \quad (2.31)$$

where δ is the conditional probability that the trade at time *t* and the trade at *t+1* are both at bid (or ask) price, and it is estimated with maximum likelihood as in equation 2.32.

$$L(\delta) = \prod_{i=1}^{n} [\delta I_i + (1-\delta)(1-I_i)] \tag{2.32}$$

where n is the number of observations (transactions), and I_i is a dummy variable that equals one if consecutive trades are the same type or equals zero otherwise.

2.5.2.3 Hasbrouck's Gibbs model

Hasbrouck (2009) extended Roll's model (1984) by using a Bayesian Gibbs sampler. The model is given in equation 2.33.

$$r_t = c\Delta D_t + \beta^m r_t^m + \epsilon_t \tag{2.33}$$

where m is a market factor with return r_t^m on day t, D_t is a sell side indicator, and c is the half-spread.

2.5.2.4 High-low (HL) spread estimator

Corwin and Schultz (2012) developed a bid-ask spread estimator based on daily high and low prices. They argued that the daily high prices correspond to buy trades, while the daily low prices are likely to be the results of sell orders. Therefore, the high-low ratio reflects an asset's variance and bid-ask spread. They also assert that the variance component of the high-low ratio should be proportional to time, while the bid-ask spread component should be constant. In order to split these two components, they use the high-low ratios on consecutive days.

The model only needs daily low and high prices, and the output is a daily estimate of the relative bid-ask spread. The model is given in equation 2.34.

$$HL = \frac{2(e^{\alpha}-1)}{1+e^{\alpha}} \tag{2.34}$$

where $= \frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}}$, $\beta = \sum_{j=0}^{1}\left(log\left(\frac{H_{t+j}}{L_{t+j}}\right)\right)^2$, and $\gamma = \left(log\left(\frac{H_{t\prime t+1}}{L_{t,t+1}}\right)\right)^2$

$L_t(H_t)$ is the lowest (highest) price on day t, and $L_{t,t+1}(H_{t,t+1})$ is the lowest (highest) price on two consecutive days t and $t+1$.

Additionally, in order to get rid of negative bid-ask spread estimates, Corwin and Schultz (2012) recommended an overnight return adjustment. The adjustment is as follows: If the close price on day t is less than the low price on $t+1$, decrease the low and high prices on day $t+1$ by that difference. If the close price on day t is greater than the high price on day t, increase the low and high prices on day $t+1$ by that difference.

2.5.2.5 Holden model

Holden (2009) used spread proxies that correspond to three attributes of daily data. First attribute is price clustering which means the higher likelihood for trade prices to be on rounder increments. Second attribute is serial price covariance. He extended Roll's (1984) model in order to estimate data for no-trade days. The last attribute is the quoted spread that is available on no-trade days.

Holden (2009) modified Huang and Stoll model (1997) via allowing the effective spread to change from day to day and eliminating the inventory holding risk component. In this model, the unobservable fundamental value of a security is calculated with equation 2.35.

$$V_t = V_{t-1} + \frac{1}{2}\alpha S_{t-1}Q_{t-1} + \varepsilon_t \tag{2.35}$$

where α is the percentage of the half spread attributable to the adverse selection, $Q(t-1)$ is buy/sell/no trade indicator, and ε is the public information shock. He defined the daily change in midpoint as in equation 2.36.

$$\Delta M_t = \frac{1}{2}\alpha S_{t-1}Q_{t-1} + \varepsilon_t + \Delta\omega_t \tag{2.36}$$

where $\omega(t)$ is the cumulative inventory effect of all previous trades. He also defined price change process as in equation 2.37.

$$\Delta P_t = \frac{1}{2}S_tQ_t - (1-\alpha)\frac{1}{2}S_{t-1}Q_{t-1} + e_t \tag{2.37}$$

2.5.2.6 Close-high-low (CHL) model

Abdi and Ranaldo (2017) proposed a method to estimate the bid-ask spread that uses daily close, high, and low prices. The rationale behind their model is that a financial security's price deviates from its efficient value is due to the transaction costs. Their model uses a wider set of data than Roll's (1984) model, and it is not dependent on the trade direction dynamics. In other words, they do not assume that the probability of consecutive trade directions being buy or sell are equal. They estimated the bid-ask spread as in equation 2.38.

$$CHL = 2\sqrt{E[(c_t - \eta_t)(c_t - \eta_{t+1})]} \tag{2.38}$$

where c is logarithm of the daily close price, and η is the average of logarithms of daily high price and low price.

They mentioned about three methods to deal with the negative estimates. First method is to calculate monthly average spread, and then set the negative estimates to zero. Second method is to estimate the squared spreads over two days, then to set the negative values to zero and calculate the average spread. Third method is to ignore the negative estimates, and calculate the average of remaining positive estimates.

The first correction method is called as monthly corrected estimate and represented in equation 2.39.

$$\hat{s}_{monthly\ corrected} = \sqrt{max\left\{4\frac{1}{N}\sum_{t=1}^{N}(c_t - \eta_t)(c_t - \eta_{t+1}), 0\right\}} \tag{2.39}$$

The second approach is called as two-day corrected estimate and represented in equation 2.40 and 2.41.

$$\hat{s}_{two-day\ corrected} = \frac{1}{N}\sum_{t=1}^{N}\hat{s}_t \tag{2.40}$$

$$\hat{s}_t = \sqrt{max\{4(c_t - \eta_t)(c_t - \eta_{t+1}), 0\}} \tag{2.41}$$

where N is the number of days the bond traded.

2.5.2.7 Closing percent quoted spread (CPQS)

Chung and Zhang (2014) proposed a bid-ask spread measure which relies on the end of day quote data. CPQS is calculated as the difference between the best bid and ask quotes, which are available at the end of the day, divided by the average of these quotes. The measure results in a relative bid-ask spread, and only requires the best buy and sell quotes available at the end of the day. The CPQS is given in equation 2.42.

$$CPQS_t = \frac{(Closing\ Ask_t) - (Closing\ Bid_t)}{((Closing\ Ask_t) + (Closing\ Bid_t))/2} \tag{2.42}$$

2.5.2.8 Fong-Holden-Trzcinka (FHT) model

Lesmond et al. (1999) asserted that the days with zero return or zero volume for a financial security are the signals of lower levels of liquidity, and they employed the ratio of zero return days as a proxy of liquidity. Based on their proposal, Schestag et al. (2016) defined the Zeros metric as the ratio of the number of days with zero volume in a month over number of trading days in a month, as seen in equation 2.43.

$$Zeros = \frac{number\ of\ days\ with\ zero\ volume\ in\ a\ month}{number\ of\ trading\ days\ in\ a\ month} \tag{2.43}$$

Fong et al. (2017) developed FHT model on the basis of the Zeros metric. According to them, the observed return is a function of symmetric transaction costs and unobserved true value returns, which is assumed to be normally distributed with zero mean and variance σ^2. The main idea of the model is that when the magnitude of true return is less than the transaction costs of buy and sell trades, a zero return day happens, i.e. any trade does not occur in that day. A greater value of the Zeros metric points to a greater bid-ask spread and higher transaction costs. The model is represented in equation 2.44.

$$FHT = 2\sigma N^{-1}\left(\frac{1 + Zeros}{2}\right) \tag{2.44}$$

where N^{-1} is the inverse of the cumulative standard normal distribution, and σ is the monthly standard deviation of the daily returns.

2.5.3 Performance comparison of bid-ask spread measures

Following Roll, many other researches, which focus on building powerful measures with less effort, took place in the literature. As a result of increased number of the measures in the literature, there are plenty of studies in the literature which compare the performances of the bid ask spread estimation models. Such researches focus on how precisely these measures can represent the actual bid-ask spread in a market. The vast majority of these mesasures are primarily designed for stocks, and their performances are tested in stock markets. Furthermore, they are implemented over different asset classes and countries. Such comparative studies are available for different asset classes and market structures: Goyenko et al. (2009) and Fong et al. (2017) covered exchange traded stocks, Marshall et al. (2011) covered commodity futures, Karnaukh et al. (2015) worked on foreign exchange rates, and Schestag et al. (2016) analysed OTC bond markets. However, the results differ across different market structures and asset classes. According to the previous comparative studies in the literature, there is no single measure to outperform others for different market structures every time.

Lesmond (2005) discovers that the low-frequency measure, the limited dependent varible technique, has a correlation coefficient with the high frequency measures greater than 0.8 for twenty three of thirty one markets. Also, Lesmond et al. (1999) and Lesmond (2005) run horseraces between the liquidity measures and find that the LOT measure dominates the other models, such as Roll, Illiquidity, Liquidity, and Turnover models.

Goyenko et al. (2009) report that the low-frequency proxies are able to approximate the high-frequency measures both at monthly and annual frequency. So, they look for the outperforming low-frequency proxy that dominates others. They implement nine low-frequency proxy in the US stock markets and compare these proxies with each other and against the high-frequency measures. According tho their results, Holden measure (Holden, 2009) has the best performance. Additionally, they find that Effective Tick measure (Goyenko et al., 2009; Holden, 2009) and LOT Y-split (Lesmond et al., 1999) dominate others, and they should be selected if someone desires to avoid computational burden.

Chung and Zhang (2014) run horseraces between the low-frequency proxies in the US stock markets, and assert that CPQS measure overrides the other low-frequency proxies.

Fong et al. (2017) compare the performances of low-frequency measures for 42 stock exchanges from 38 countries. They implement performance comparison via four dimensions: (i) Average cross-sectional correlation coefficients with the liquidity benchmarks, (ii) portfolio time series correlation coefficients with the liquidity benchmarks, (iii) individual time series correlation coefficients with the liquidity benchmarks, and (iv) RMSE against the liquidity benchmarks. According to their results, CPQS dominates the others at both monthly and daily frequencies as a relative measure, if quote data is available. HL (Corwin & Schultz, 2012) and FHT (Fong et al., 2017) dominate the others when quote data is not available. Amihud measure (Amihud, 2002) produces the best results at daily estimates as a cost-per-dollar-volume proxy, while CPQS Impact, LOT Mixed Impact, Amihud, High-Low Impact, and FHT Impact measures have close estimates for monthly forecasts.

Although, some proxies perform better than the others sometimes, there is no clear evidence that any of such proxies dominate the others every time, for stock markets. In addition to the researches which cover stock markets, other markets, such as foreign exchange, commodity, fixed income, ETF and cryptocurrency markets are investigated in the literature.

Marshall et al. (2011) compare 17 low-frequency proxies in the US commodity market. According to their results, the models with the highest correlation coefficients with the benchmark spread are Amihud measure (Amihud, 2002), Amivest (Amihud et al., 1997), and Effective Tick (Goyenko et al., 2009; Holden, 2009). These three measures perform better than the others. However, any of these measures is not able to represent the actual scale for the trade costs in the commodities markets.

Karnaukh et al. (2015) examine the bid-ask spread measures in the spot foreign exchange markets, and discover that low-frequency data is enough to adequately measure liquidity. Also, they make comparison between the low-frequency proxies, and find that HL measure (Corwin & Schultz, 2012) has the greatest correlation coefficient with the high-frequency spread benchmark.

Guloglu and Ekinci (2016) analyze the futures traded in BIST, and compare the low-frequency proxies' performances. According to their results, Effective Tick measure (Goyenko et al., 2009; Holden, 2009) has better forecast performance than the other proxies. Nevertheless, they report that a sole indicator of the bid-ask spread does not exist in the literature.

Marshall et al. (2018) compare the low-frequency proxies’ performances in the ETF markets, and they say that the low-frequency proxies are able to capture spread dynamics. According to their results, CPQS (Chung & Zhang, 2014), HL (Corwin & Schultz, 2012), and CHL (Abdi & Ranaldo, 2017) have greater estimation performance over the other spread proxies.

Langedijk et al. (2018) focus on the European sovereign bond markets in order to invetigate the proxies’ performances. They find that the relationships between the benchmark measures and the proxies are weak, and they become weaker with higher forecast frequency. However, for daily and monthly estimates, they report that CPQS (Chung & Zhang, 2014) dominates the others.

Schestag et al. (2016) extensively examines the high-frequency and low-frequency bid-ask spread measures, and compare their performances for the US corporate bonds. They implement six high-frequency measures, and discover that these six measures are highly correlated pairwise, and have results of trade costs within similar magnitude. Next, they apply 13 low-frequency proxies, analyze their results, and compare the forecasts with regard to the high-frequency measures. According to their results, most of the low-frequency proxies are able to reflect transaction costs. Additionally, they report that the time series and cross-sectional correlation coefficients between the high-frequency and low-frequency measures in the US corporate bond markets are higher than the correlation coefficients in Goyenko et al.’s (2009) research in stock markets. They assert that HL (Corwin & Schultz, 2012), Roll (1984), and Hasbrouck's Gibbs measure (Hasbrouck, 2009) have better forecast performance than the other proxies. Moreover, they report that the low-frequency measures could be preferable over the high-frequency measures, since they reduce the data requirements and calculation burden, even though minor accuracy loss in estimation performance may occur.

Diaz and Escribano (2019) examine 16 liquidity measures, and report that the bid-ask spread measures can comprehend the differences between the credit ratings in the US corporate bond markets.

Hameed et al. (2019) examine the Malaysian corporate bond market liquidity, and report that the quantity-based liquidity proxies perform better than the price-based liquidity proxies for their sample.

Bohmann et al. (2019) find that CPQS has the greatest performance as a proxy of the relative bid-ask spread. Additionally, Gao et al. (2020) compare the low-frequency liquidity proxies against the high-frequency bid-ask spreads for the Chinese stock markets. Their empirical results show that CPQS produces the smallest forecast error, and FHT has the greatest correlation coefficient with the high-frequency benchmarks. They say that CPQS and FHT outperform the other measures mostly, and assert that the measures based on Roll's (1984) model have inferior performance.

Brauneis et al. (2021) compare the low-frequency and high-frequency liquidity measures for the cryptocurrency markets. They find that CHL and CS measures are the best estimators of the bid-ask spread.

Different bid-ask spread measurement techniques are designed in order to understand features and mechnanisms of transaction costs in different ways. Even though different financial markets are analysed and the number of comparative studies rises, there is still no consensus about the best low-frequency measure that can replace high-frequency measures. There is no single best model that fit different asset classes and market structures. Nevertheless, the existing studies in the literature show that the usage of the low-frequency measures as a proxy of the bid-ask spread and liquidity is appropriate and efficient.

2.5.4 Bond features and liquidity

There are plenty of researches which attempt to define and evaluate the connections between asset features and bid-ask spread (or liquidity) in financial markets.

Chen et al. (2007) investigate the relations between liquidity, bid-ask spread, and bond yield for the corporate bond markets. According to their results, outstanding nominal amount of a bond has very limited effect on bond liquidity, and the relation between liquidity and bond volatility is significantly strong. They also assert that liquidity effect exists both for the investment and speculative-grade bonds. Moreover, there is strong correlation between changes in liquidity and yield spreads, and lower liquidity is significantly associated with greater yield spread. An increase in liquidity results in a decrease in yield spread. They find that the liquidity costs of speculative-grade bonds is greater than the liquidity costs of investment-grade bonds, and bonds with longer term to maturity have higher yield spreads and liquidity costs. Moreover, they report that liquidity and credit rating are significantly important for bond yield spreads, but liquidity effect is the most significant determinant. Additionally, Chang et al. (2021) investigates liquidity, yield and credit rating relations in the Chinese green bond market, and find that market liquidity has significant effect on bond yields.

Edwards et al. (2007) assert that an increase in outstanding nominal amount of a bond or a decrease in credit risk reduces the transaction costs. They also state that the lowest levels of transactions costs occur just after the issue date and just before the maturity date of a bond. Bao et al. (2011) report that an increase in outstanding nominal amount of a bond improves its liquidity, while a bond's age and maturity decrease its liquidity. Feldhütter and Poulsen (2018) say that longer term to maturity and higher credit risk lead to wider bid-ask spreads for the investment-grade corporate bonds. Chen and Stock (2021) assert that an additional issue of a preexisting bond has a positive effect on the bid-ask spread and liquidity, but that effect is temporary and decay over time. Addtionally, Helwege and Wang (2021) report that corporate bonds with greater outstanding nominal amount have higher liquidity, as expected.

Batten and Vo (2014) investigate liquidity and stock returns in Vietnam, and they discover that there is a positive correlation between them. Their findings do not confirm the typical relationships in the developed countries, and they argue that such a divergence could be a result of the lower levels of integration of an emerging country stock market with the global markets.

In certain bond markets, market makers (or primary dealers) are obliged to supply minumum amounts of buy and sell quotes within certain price limits. In other words, market maker presence entails a maximum of bid-ask spread. Bessembinder et al. (2020a) examine the market maker contracts for the US stocks, and find that market making activity leads to narrower bid-ask spread, higher price efficiency, and superior liquidity. Similarly, Beau (2014) asserts that less traded securities have greater bid-ask spreads than more actively traded securities. According to Remolona and Yetman (2019), a narrow bid-ask spread is an expected characteristic of a benchmark fixed income security.

Galliani et al. (2014) build a robust illiquidity index, and investigate the liasion between the bond properties and the robust illiquidity index for the European government bonds. They discover that outstanding nominal amount, term to maturity at issue date, issuer's credit rating, and bond duration have significant effects on bond liquidity. According to their results, lower credit rating, higher duration, and less outstanding nominal amount induce lower levels of bond liquidity.

Huang and Tian (2020) examine the Chinese corporate bond markets, and find that outstanding nominal amount, trade volume, bond return, and turnover are the proxies of bond liquidity premium, and the most significant characteristics is bond age. Mizrach (2015) reports that larger number of trades, smaller trade size and more extensive dealer networks result in higher liquidity for the US corporate bonds.

Ariyana et al. (2020) investigate the liquidity determinants of the government sukuks in Indonesia, and they find that higher outstanding nominal amount and lower yield to maturity are associated with higher liquidity. However, time to maturity has no effect on sukuk liquidity according to their results.

Arcuri et al. (2020) investigates the liquidity and bond spread relation for the European corporate bonds, and they assert that lower liquidity is associated with higher risk premia and wider bid-ask spread. They find a significant relation between the bid-ask spread and the default risk (or the corporate bond spread).

Phylaktis and Manalis (2013) examine whether the presence of futures on a stock has effects on the underlying stock liquidity. They discover that greater trading activity of futures leads to narrower bid-ask spread of the underlying stock. They report that trade of futures decreases the depth, and leads to lower inventory holding costs, which probably causes a decrease in the bid-ask spread.

Altay and Çalgıcı (2019) inspect asset returns and liquidity of the stocks traded in BIST. They assert that asset illiquidity and market illiquidity are positively correlated, while market return and asset illiquidity are negatively correlated.

Bedowska-Sojka (2019) examines the volatility and liquidity relation in the European stock markets. She finds that a rise in volatility causes lower liquidity and a greater bid-ask spread. She also reports that volatility and spread increase in case of a financial turmoil. In a similar manner, Bedowska-Sojka and Echaust (2019) analyzes five European emerging country stock markets in order to describe the volatility and liquidity relationship. They discover that global volatility and market liquidity are negatively correlated. Additionally, Bessembinder et al. (2020b) assert that macroeconomic announcements create uncertainty for sovereign fixed income securites which in turn results in decrease of trade volumes, immediate price changes, and increase in bid–ask spreads.

34

3. DATA AND METHODOLOGY

3.1 Data

The main focus of this thesis is on the Turkish and some other emerging country fixed income securities markets. Due to the data availability and trading frequency in these markets, only sovereign bonds are included in our data sets. In order to analyze the bid ask spread and its relations with the bond level factors, two different data sets are used in this study. First data set includes only Turkish Treasury bonds, while the second data set consists of sovereign bonds from a variety of emerging markets.

The first data set consists of bond data from BIST DSM. It is possible to trade Turkish Treasury bonds both at the OTC markets and the centrally organized exchange. BIST DSM is the only centrally organized exchange for Turkish bonds, and all trades handled electronically via the automated multiple price-continuous auction system. The sampling period for the first data set includes 231 trading days, and spans from 02.07.2018 to 31.05.2019. 48 instrument series with the highest trade volumes throughout the sampling period are selected as sample[1], and these instrument series constitute 79% of the total trade volume in the BIST Outright Purchases and Sales Market for the sampling period.

Both intraday and daily data is included in the first data set in order to calculate the quoted bid-ask spread and implement the other spread measures. Intraday data includes trade price, trade volume, trade timestamp, and bid-ask quotes, while daily data consists of daily low, high and close prices. Also, bond specific data, such as bond type, term to maturity, settlement date, outstanding amount, and market maker presence are included in order to use as the independent variables of the panel regression analysis. The first data set is obtained from BIST Data Store.

---

[1] It is possible to trade a bond with different settlement dates in a single business day, and a unique bond series is created for each settlement date. For example, there exist more than one instrument series for the bond TRT110226T13, and two of them are TRT110226T13_KESN_T0 and TRT110226T13_KESN_T1. As of business day 08/07/2019; the former series' settlement date is 08/07/2019, and the latter's is 09/07/2019.

The second data set consist of 33 LC sovereign bonds from six emerging countries, which are Turkey, South Africa, Russia, Indonesia, India, and Brazil. The sampling period for that data set includes 64 trading days, and spans from 21.06.2019 to 18.09.2019. The data set consists of daily data such as daily high, low, close prices, and daily quotes, also intraday quotes are included. We look for effects of credit rating on bid ask spread. However, there is only two minor rating change during the sampling period. Due to low number of credit rating change, we did not include it. Instead of credit rating changes, CDS of the issuer country is included, as both of them measure the same thing and are expected to follow each other. Additionally, bond specific data such as credit default swap, historical volatility, mid yield, basis point value, duration, option asset swap spread, market return, and issuer country, is included in the data set. These bond specific data has daily frequency. Thomson Reuters and Bloomberg are the data source for the second data set.

3.2 Methodology

The main objective of this thesis is to answer two research questions about emerging country bond liquidity and bid-ask spread: Do the low-frequency bid-ask spread measures are able to proxy the real-time bid-ask spread, and how the bid-ask spread is affected by the bond and country specific characteristics. The first data set is applied in order to answer both of these two questions, while the second data set is for the second research question.

For the first data set, we compute time weighted relative quoted bid-ask spread (TWS) with tick-by-tick order flow data throughout the sampling period for each bond. This high frequency data set contains the best bid and ask prices available at the market with timestamps. Since the data used for TWS includes high frequency real-time market data and represents the real quoted spread available in the market, TWS is selected as the benchmark for other spread measures and proxies.

TWS is represented in equation 3.1, and it is time weighted average of quoted spread that is mentioned in equation 2.2. The difference between the best available bid and ask prices at a certain point in time is called the quoted spread, and to have a relative measure it is divided by the average of the best bid and ask quotes. We applied time-weighted average for the intraday relative quoted spread instead of the simple average, since the best bid and ask quotes change within a day irregularly. The weight of each relative quoted bid-ask spread is calculated as the proportion of time that the spread is available in the market over the total trading time with a spread. TWS is a daily measure that incorporates all intraday updates in the available spread.

$$TWS = \frac{1}{T}\sum_{t=1}^{Q}\frac{Ask\ Price_t - Bid\ Price_t}{(Ask\ Price_t + Bid\ Price_t)/2} * (T_{t+1} - T_t) \tag{3.1}$$

where T is the total time with an available spread, Q is the total number of updates in the spread, T_t is the timestamp of $t - th$ spread update.

Daily TWS is computed for each bond, and the monthly benchmark spreads for each bond are calculated as the volume-weighted averages of the daily TWS.

Next, we estimate the relative bid-ask spread for each bond with seven different bid-ask spread measures. The measures are RT, CHL, Roll, FHT, CS, IQR, and CPQS, which are described in section 2. Each of these measures, except FHT, returns daily estimates of the relative bid-ask spread. In order to be able to compare all measures applied, we transform the daily estimates into monthly estimates via calculating volume-weighted averages of the daily results. So, the analysis with the first data set is based on the monthly bid-ask spread estimates. Additionally, data for term to maturity and volume are monthly frequent, while the other bond specific independent variables are time-invariant.

Afterwards, we run time series and cross-sectional regressions with the monthly estimates in order to compare the performances of the seven measures applied. For the time series regression analysis, a volume-weighted portfolio of bond series is constituted for each month. A monthly portfolio relative bid-ask spread is calculated as the volume-weighted average of the spread estimates available for that month for each measure. Monthly bond volumes are used as the portfolio weightings. For the cross-sectional regression analysis, a weighted average of the relative bid-ask spreads for each bond series is calculated via using monthly spread estimates with monthly volumes as the weightings. Each of the applied seven measures is regressed on TWS. Before implementing the regression analysis, in order to control stationarity of TWS and other bid-ask measures, a unit root test is applied. According to the Augmented Dickey-Fuller test result, TWS, RT, CHL, Roll, FHT, CS, IQR, and CPQS series are stationary, even at 1% significance level. So, the benchmark and other measures are included into regression without any transformation.

Results of the time series and cross-sectional regression models are evaluated with certain statistical metrics. In order to observe the patterns of measures against TWS, correlation coefficients are calculated. Additionally, some other measures are applied to evaluate the performance and accuracy of the measures, as magnitude is an important feature for transaction costs. Therefore, certain metrics such as MBE, MAE, and RMSE are calculated for both cross-sectional and time series regression results.

MBE refers to systematic error, and it is the average of forecast errors (equation 3.2). It shows the average model bias. MAE represents the average of magnitudes of error terms, and equals to average of absolute errors (equation 3.3). RMSE represents how broadly the observations spread around the estimated regression line, and equals to standard deviation of forecast errors (equation 3.4). It dedicates greater weights to larger errors in absolute term, and uses squared values of forecast errors. For MBE, MAE, and RMSE, the lower the result, the better fit.

$$MBE = \sum_{n=1}^{N} \frac{\overline{TWS_n} - Estimated\ Spread_n}{N} \tag{3.2}$$

$$MAE = \sum_{n=1}^{N} \left| \frac{\overline{TWS_n} - Estimated\ Spread_n}{N} \right| \tag{3.3}$$

$$RMSE = \sqrt{\sum_{n=1}^{N} \frac{(\overline{TWS_n} - Estimated\ Spread_n)^2}{N}} \tag{3.4}$$

For time series regression, *N* is the number of months, and $\overline{TWS_n}$ is the average of monthly TWS measures in month *n*. For cross-sectional regression, *N* is the number of bond series with a spread in a month, and $\overline{TWS_n}$ is the average of monthly TWS for a bond series. Both for the time series and cross-sectional regressions, MBE, MAE, and RMSE are computed for each of the applied measures.

For the second research question, the effects of bond specific characteristics on the bid-ask spread are analyzed. For this analysis, several regression methods are investigated. The pooled OLS regression is not choosen, since it has downsides when heterogeneity across groups exists. According to Hsiao (2014), the panel data method has advantages, as it helps to cope with heterogeneity and is able to discover effects that are not possibly detected via cross-sectional and time series regression. Additionally, time-invariant variables could be included into a panel regression model, with random-effects models. As a result, we choose a random-effects regression model, and apply certain statisical tests in order to validate the selected model. The random-effects regression model is represented in equation 3.5. Additionally, in order to control stationarity of TWS and independent regression variables, a unit root test is applied. According to the Augmented Dickey-Fuller test result, TWS, DTM, VOL, and NOM are stationary at 5% significance level. So, the dependent and independent variables are included into regression without any transformation.

For the panel regression model given in equation 3.5, TWS is the dependent variable, and the bond specific characteristics are the independent variables, which are described in Table 3.1. FRN, SET, NOM, and MM are time-invariant variables, while TWS, TTM, and VOL are time-variant.

$$\begin{aligned} TWS_{ij} = \mu + \beta_1 VOL_{ij} + \beta_2 TTM_{ij} + \beta_3 NOM_j + \beta_4 MM_j + \beta_5 SET_j \\ + \beta_6 FRN_j + u_i + e_{ij} \end{aligned} \tag{3.5}$$

where j refers to each bond series, i refers to month, u_i is the random effect (month specific), and e_{ij} is the error term.

Table 3.1 : Definitions of panel regression variables for the Turkish Sovereign Bond Market.

Variable	Definition
VOL	Trade volume of a bond series
TTM	Term to maturity of a bond (in months)
NOM	Outstanding amount of a bond at the issue date
MM	Dummy variable for market maker presence: Equals zero when there is no market maker for a bond, and equals one when at least one market maker exists.
SET	Dummy variable for settlement date: Equals zero if settlement date is T+0, and equals one if settlement date is at T+1.
FRN	Dummy variable for bond type: Equals zero for fixed-rate bonds, and equals one for floating-rate bonds.

For the second data set, we analyze the effects of the country-level and bond-level characteristics on the bid-ask spread for the selected emerging country treasury bonds.

Therefore, in order to establish a bid-ask spread measure, we compute average relative bid-ask spread for each bond for each business day. Similar to TWS given in equation 3.1, we calculate a relative spread with the intraday bid and ask quotes. However, we use simple average method, instead of time-weighted average, as quote timestamp is not available for the second data set. The benchmark relative spread is given in equation 3.6.

$$RS = \frac{1}{N}\sum_{n=1}^{N}\frac{Ask\ Quote_n - Bid\ Quote_n}{(Ask\ Quote_n + Bid\ Quote_n)/2} \tag{3.6}$$

where N is the number of changes in the spread within a business day.

Next, we implement a regression analysis to answer the second research question about the bid-ask spread and bond specific features. As mentioned earlier, panel data regression has advantages over OLS, as it allows to include time-invariant variables and helps to handle heterogeneity accross groups. The mixed-effect panel data regression we applied is given in equation 3.7. In order to control stationarity of RS and independent regression variables, the Augmented Dickey-Fuller test is applied before implementing the regression analysis. According to the Augmented Dickey-Fuller test result, the RS, MY, and CDS series are not stationary. So, differentiation method is applied in order to make these series stationary. After differencing, all regression variables are stationary at 10% significance level.

Daily RS of each bond, which is calculated with equation 3.6, is the dependent variable, and the bond-specific and country-specific variables are the independent variables of the model, which are represented in Table 3.2. AS, DUR, BPV, MY, VOL, CDS, and MKT are continuous variables and are daily frequent, while COU is a categorical and time-invariant variable.

$$\begin{aligned} \Delta RS_{ij} = \mu + \beta_1 AS_{ij} + \beta_2 DUR_{ij} + \beta_3 BPV_{ij} + \beta_4 \Delta MY_{ij} + \beta_5 VOL_{ij} \\ + \beta_6 \Delta CDS_{ij} + \beta_7 MKT_{ij} + COU_j + b_{0i} + b_{0j} + e_{ij} \end{aligned} \tag{3.7}$$

where j refers to each bond, i refers to each business day, b_{0i} and b_{0j} are the random intercepts, and e_{ij} is the random error.

Table 3.2 : Definitions of panel regression variables for the Emerging Country Sovereign Bond Markets.

Variable	Definition
RS	Relative Quoted Spread
AS	Asset Swap Spread
DUR	Duration
BPV	Basis Point Value
MY	Mid Yield
VOL	Historical Volatility
CDS	Credit Default Swap of Issuer Country
MKT	Daily Market Return in Issuer Country
COU	Country

4. EMPIRICAL FINDINGS

4.1 Findings on the Turkish Sovereign Bond Market

For the first data set, the descriptive statistics of the monthly bid-ask spread measures of the Turkish sovereign bonds are summarised in Table 4.1. TWS, our real-time benchmark spread, has a mean of 0.588% and a median of 0.411%. For example, a round-trip trade with one million TL volume has transaction cost of 5880 TL on average because of the bid-ask spread. The other measures' estimated means span between 0.224% and 1.011%.

Table 4.1 : Descriptive summary statistics of monthly spread measures for the Turkish Sovereign Bond Market.

	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
TWS	495	0.588%	0.520%	0.223%	0.411%	0.825%
RT	402	0.401%	0.651%	0.114%	0.216%	0.441%
CHL	456	0.761%	0.790%	0.226%	0.482%	1.015%
Roll	357	0.362%	0.360%	0.107%	0.228%	0.513%
FHT	451	1.011%	1.518%	0.241%	0.472%	1.070%
CS	397	0.224%	0.311%	0.052%	0.126%	0.267%
IQR	488	0.324%	0.341%	0.104%	0.223%	0.439%
CPQS	496	0.634%	0.633%	0.233%	0.435%	0.792%

CPQS gives the closest estimation results, and CHL is the second best, with regard to magnitude. RT, Roll, CS, and IQR underestimate the relative bid-ask spread, while CPQS, CHL, and FHT overestimate it. Additionally, the monthly averages are plotted in Figure 4.1 in order to demonstrate the patterns of bid-ask spread over the sampling period.

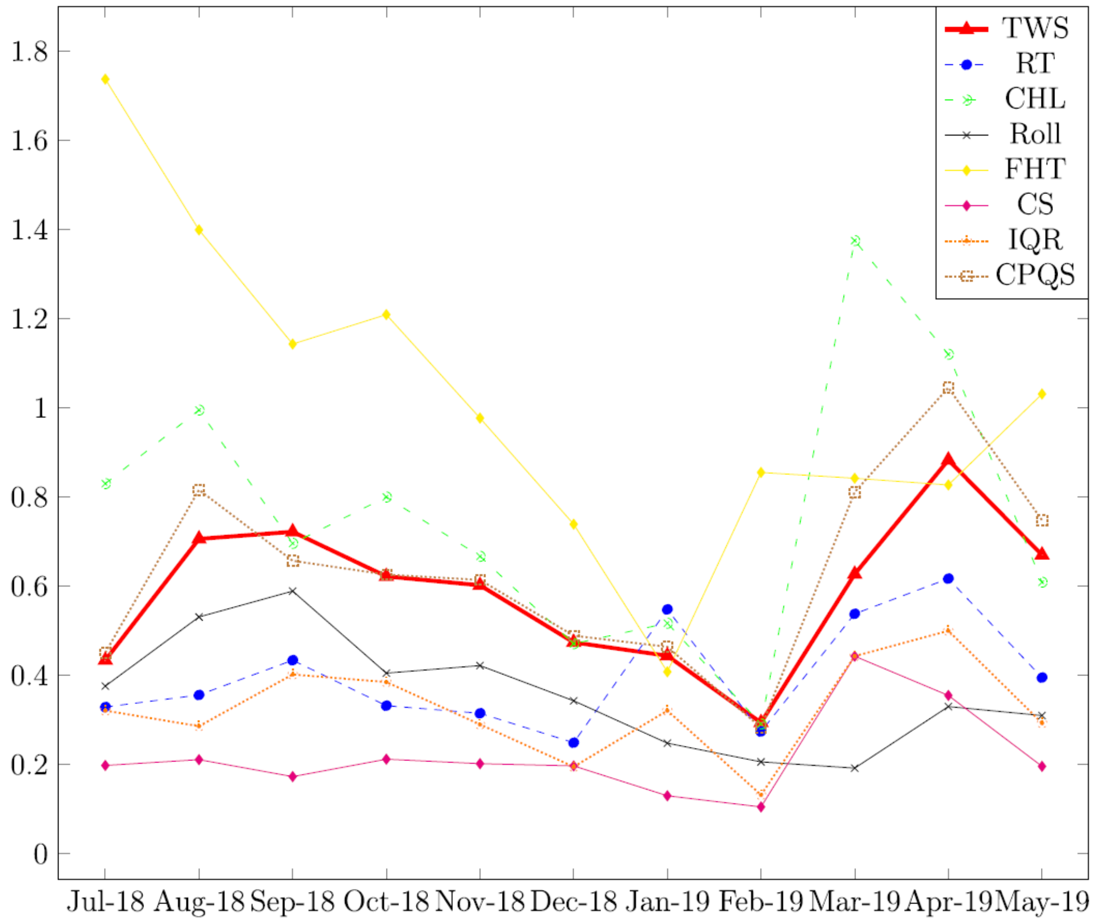

Figure 4.1 : Monthly averages of relative bid-ask spreads for the Turkish Sovereign Bond Market.

For the time series regression analysis, the monthly estimates for each measure are computed as the volume-weighted averages of the bond series' spreads. Then, these monthly averages are regressed on TWS. The correlation coefficients for the time series analysis are given in Table 4.2. According to the results, each measure is positively correlated with the benchmark, TWS. CPQS has the highest correlation coefficient of 0.603, while RT's coefficient is 0.522, and IQR has the third greatest coefficient with a value of 0.508. Moreover, we calculate each measure's pairwise correlation, and the pairs of CHL-CPQS, CHL-CS, CHL-IQR, CPQS-CS, and CPQS-IQR have high correlation coefficients. Nevertheless, many of the measures do not have significant or positive correlation coefficients with each other. Specifically, Roll and FHT are largely negatively correlated with the others. The main reason of that difference could be the data frequency, which these two measures use. Roll employs the daily return data of the last twenty one trading days, and FHT requires the close price volatility and the number of days with no trade within a month.

Table 4.2 : Time series correlation coefficients of monthly spread measures for the Turkish Sovereign Bond Market.

	TWS	RT	CHL	Roll	FHT	CS	IQR	CPQS
TWS	1							
RT	0.522	1						
CHL	0.361	0.727	1					
Roll	0.319	0.236	-0.164	1				
FHT	0.493	-0.042	-0.032	0.264	1			
CS	0.339	0.516	0.841	-0.355	-0.045	1		
IQR	0.508	0.705	0.801	-0.199	-0.082	0.744	1	
CPQS	0.603	0.659	0.903	-0.196	0.004	0.861	0.817	1

Further, the prediction errors of the time series regression for each measure are calculated, and the results are given in Table 4.3. All measures have results within a close range for MBE and MAE, except CHL. Similary, all measures, except CHL and FHT, have close results for RMSE, which range between 0.102% and 0.187%.

Table 4.3 : Prediction errors of time series regression of monthly spread measures for the Turkish Sovereign Bond Market.

	MBE	MAE	RMSE
RT	-0.092%	0.104%	0.115%
CHL	0.505%	0.505%	0.675%
Roll	-0.035%	0.091%	0.102%
FHT	0.157%	0.187%	0.225%
CS	-0.046%	0.091%	0.118%
IQR	0.092%	0.116%	0.149%
CPQS	0.120%	0.125%	0.187%

For the cross-sectional regression analysis, volume-weighted averages of the monthly estimates are used as the average bid-ask spread, and these averages are regressed on TWS. The correlation coefficients for the cross-sectional regression analysis is given in Table 4.4. The measure with the highest correlation coefficient is CPQS, with a value of 0.899. CPQS is followed by Roll, RT, CHL, and FHT, whose correlation coefficients span between 0.527 and 0.601. Moreover, we calculate each measure’s correlation pairwise, and discover that the pairs of Roll-CHL, IQR-CHL, CPQS-CHL, and IQR-CS have higher coefficients than the other pairs.

Table 4.4 : Cross-sectional correlation coefficients of monthly spread measures for the Turkish Sovereign Bond Market.

	TWS	RT	CHL	Roll	FHT	CS	IQR	CPQS
TWS	1							
RT	0.594	1						
CHL	0.574	0.583	1					
Roll	0.601	0.474	0.775	1				
FHT	0.527	0.308	0.456	0.516	1			
CS	0.305	0.168	0.691	0.527	0.201	1		
IQR	0.433	0.551	0.865	0.696	0.366	0.780	1	
CPQS	0.899	0.592	0.733	0.643	0.415	0.544	0.592	1

Then, the prediction errors of the cross-sectional regression for each measure are calculated, and the results are summarised in Table 4.5. The results of the cross-sectional regression are similar with the time series regression results. CPQS has the smallest mean bias. In terms of RMSE, CPQS has the best result with a value of 0.177%, while Roll, RT, and IQR have close RMSE values, which span between 0.353% and 0.374%.

Table 4.5 : Prediction errors of cross-sectional regression of monthly spread measures for the Turkish Sovereign Bond Market.

	MBE	MAE	RMSE
RT	-0.152%	0.230%	0.368%
CHL	0.267%	0.358%	0.514%
Roll	-0.195%	0.262%	0.353%
FHT	0.232%	0.447%	0.746%
CS	-0.280%	0.346%	0.458%
IQR	-0.151%	0.281%	0.374%
CPQS	0.063%	0.103%	0.177%

According to the results of the time series and cross-sectional regression analyses, CPQS dominates the other measures mostly. However, the results are different for the other measures, none of them dominates the others clearly every time. Our results, which indicate CPQS is the best proxy at capturing bid-ask spread dynamics, are similar with Chung and Zhang's (2014) results, who discover that CPQS outperforms other liquidity measures for the US stock markets. Also, our results are coherent with Fong et al. (2017), who assert that CPQS dominates both for daily and monthly estimations as a relative proxy. Moreover, Marshall et al.'s (2018) finding about the bid-ask spread on the ETF markets, which also denotes CPQS as the most competent measure, endorses our results. Gao et al.'s (2021) findings also support our results for superior performances of CPQS and FHT.

On the other hand, Karnaukh et al.'s (2015) findings on the commodity markets and Schestag et al.'s (2016) results on the OTC bond markets conflict with our findings. Additionally, our findings does not support the results of Corwin and Schultz (2012) and Lesmond (2005).

In addition to those findings, we examine the bond-specific characteristics' effects on the bid-ask spread for the Turkish sovereign bonds with the first data set. The panel regression, given in equation 3.5 is implemented with the variables described in Table 3.1.

The summary statistics of the regression variables are given in Table 4.6, and the pairwise correlation coefficients of the regression variables are represented in Table 4.7. TTM and NOM have positive coefficients against TWS, whereas the correlation coefficients of VOL, MM, SET, and FRN with TWS are negative.

Table 4.6 : Descriptive statistics of panel regression variables for the Turkish Sovereign Bond Market.

	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
TWS	495	0,006	0,005	0,002	0,004	0,008
TTM	495	654,194	1162,739	120,15	254,5	594,25
VOL	495	41,97	32,95	14,37	34,67	57,93
NOM	48	2436,83	890,66	1940,10	2300,40	2821,40
MM	48	0,265	0,442			
SET	48	0,541	0,499			

Table 4.7 : Correlation coefficients of panel regression variables for the Turkish Sovereign Bond Market.

	TWS	TTM	VOL	NOM	MM	SET	FRN
TWS	1						
VOL	-0,221	1					
TTM	0,514	0,108	1				
NOM	0,156	-0,01	0,298	1			
MM	-0,234	0,125	-0,249	-0,253	1		
SET	-0,025	0,198	0,031	0,053	-0,009	1	
FRN	-0,196	-0,081	-0,223	-0,161	0,521	0,045	1

The panel regression model results are given in Table 4.8. The model's adjusted R-squared value is 0.370. We implement the Breusch-Pagan test (Breusch & Pagan, 1979) to compare pooled OLS and panel regression methods. The test result indicates that appling pooled OLS is inconsistent on this data set, since the variances across groups are greater than zero, and there exist a panel effect. Next, we run the Hausman test to determine the correct panel data method, and the result suggests us to use a random-effects model instead of fixed-effects models, which implies that the unique errors and the regressors are not correlated.

Table 4.8 : Random effect panel regression results for the Turkish Sovereign Bond Market.

	TWS (Dependent variable)
TTM	8.36E-05*** (0.00000)
VOL	-1.3489E-06*** (0.00000)
NOM	-1.41E-07 (0.51770)
MM	-2.21E-04 (0.66353)
SET	2.23E-04 (0.55114)
FRN	-1.703-E03* (0.02616)
Constant	3.69-E03*** (0.00000)
Observations	495
R^2	0.378
Adjusted R^2	0.370
F Statistic	296.047***
Hausman Test	χ^2 (6) 3.3128
Breusch-Pagan	χ^2 (1) 5.6137**

Note: *p<0.1; **p< 0.05; ***p<0.01

VOL, NOM, MM, and FRN have negative regression coefficients, which means an increase in one of these variables lead to a narrower relative bid-ask spread. On the other hand, the independent variables with positive regression coefficients are TTM and SET. Only three of the regression coefficients are statistically significant, which are TTM, VOL, and FRN. There is a negative relation between MM and the bid-ask spread according to the results. In other words, a narrower bid-ask spread is associated with the market maker presence. Market makers supply quotes to an instrument and bring an upper limit to the bid-ask spread (Beau, 2014). Hence, it is expected that the presence of market makers increases liquidity, and decreases the explicit trading costs. Therefore, MM's regression coefficient is expected to be negative. Our results on the market-making activity is coherent with the expectations, but the effect is not statistically significant for the first data set.

Lower trade volume and fewer trade activity generally lead to wider bid-ask spreads (Beau, 2014). So, the expectation for VOL is to have a negative regression coefficient;

the greater volume, the lower transaction costs. According to our results, VOL's coefficient is significantly negative, and the expectation is relevant with the results.

For the TTM variable, the expectation is that shorter term to maturity leads to higher trade volume and lower trading costs. TTM has a positive coefficient, so the result is coherent with the expectations about term to maturity. Our results on TTM is also relevant with Feldhütter and Poulsen (2018), who assert that the bid-ask spread and the term to maturity are positively correlated for the investment-grade corporate bonds. On the other hand, our findings on TTM are partly endorse Edwards et al. (2007), who find that the transaction costs are at the lowest point just after the bond issue and just before the maturity date.

Additionally, our results are coherent with the findings of Galliani et al. (2014) to a certain degree. Their findings about term to maturity is coherent with our results on TTM. Nevertheless, NOM does not have a significant effect on the bid-ask spread according to our results, while outstanding amount is signifcantly related with the bid-ask spread according to their results.

Furthermore, FRN significantly affects the bid-ask spread. The fixed-rate bonds have wider bid-ask spreads than the floating-rate bonds on average, according to the regression results.

To summarize the findings for the first data set of Turkish sovereign bonds, the most capable bid-ask spread estimation method is CPQS, while the others are not able to dominate each other significantly. Shorter term to maturity or higher trade volume leads to narrower relative bid-ask spread. Bond type is also a significant variable on the relative bid-ask spread.

4.2 Findings on the Emerging Country Sovereign Bond Markets

For the second data set, the descriptive statistics of the monthly bid-ask spreads and the independent regression variables are summarised in Table 4.9. Additionally, the correlation coefficients of the regression variable are given in Table 4.10.

Table 4.9 : Descriptive summary statistics of regression variables the Emerging Country Sovereign Bond Markets.

	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
ΔRS	1670	-0.001	0.125	-0.015	0.000	0.015
AS	1670	64.078	77.893	8.615	68.295	134.673
BPV	1670	10.424	14.854	2.628	5.989	9.564
ΔMY	1670	-0.014	0.123	-0.045	-0.005	0.026
DUR	1670	4.944	3.256	1.919	4.296	7.750
VOL	1670	6.269	5.999	2.020	4.793	8.463
ΔCDS	1670	-0.451	4.923	-1.738	-0.424	0.710
MKT	1670	0.000	0.009	-0.005	0.001	0.006

Table 4.10 : Correlation coefficients of regression variables the Emerging Country Sovereign Bond Markets.

	ΔRS	AS	BPV	ΔMY	DUR	VOL	ΔCDS	MKT
ΔRS	1							
AS	0.004	1						
BPV	-0.001	0.057	1					
ΔMY	-0.016	0.016	0.049	1				
DUR	-0.005	0.546	0.165	0.097	1			
VOL	-0.011	0.437	0.089	-0.061	0.419	1		
ΔCDS	0.068	0.044	0.013	0.436	0.038	-0.059	1	
MKT	0.044	-0.017	0.006	-0.239	-0.029	0.031	-0.353	1

The results of the panel regression model, given in equation 3.7, are summarised in Table 4.11. CDS, MKT and MY are the statistically significant variables on the relative bid-ask spread, whereas AS, BPV, DUR, and VOL have only slight influences on the relative bid-ask spread. According to the results, the sovereign bonds of the countries with higher CDS or higher market return have wider bid-ask spreads. Additionally, yield of a bond has significant effect on bid-ask spread. There is no significant difference between countries.

Table 4.11 : Mixed effects panel regression statistics the Emerging Country Sovereign Bond Markets.

	Coefficient	Std. Error	t-Value	Pr(>|t|)
(Intercept)	-11.010	17.680	-0.623	0.5333
AS	0.105	0.108	0.976	0.3290
BPV	0.287	0.463	0.620	0.5356
ΔMY	-54.720	28.450	-1.924	0.0546*
DUR	-2.256	2.433	-0.927	0.3539
VOL	0.245	0.814	0.301	0.7637
ΔCDS	2.921	0.725	4.029	0.0000***
MKT	985.4	357.4	2.757	0.006***
COUID	11.080	18.990	0.584	0.5595
COUIN	20.430	24.480	0.835	0.4040
COURU	24.020	27.250	0.882	0.3781
COUTR	-3.070	18.420	-0.167	0.8676
COUZA	12.850	20.140	0.638	0.5235
AIC	-2078.87			
BIC	-1992.14			
Log Likelihood	1055.44			
Num. Observations	1670			
Num. Groups: Date	63			
Num. Groups: Bond	33			
Var: Date (Intercept)	0.00003			
Var: Bond (Intercept)	0.000000			
Var: Residual	0.01548			

Note: *p<0.1; **p< 0.05; ***p<0.01

ID: Indonesia, IN: India, RU:Russia, TR:Turkey, ZA:South Africa

Our results on yield and credit risk endorse Chen et al. (2007), who assert that bond liquidity, volatility, credit risk and yield are significantly related with each other. Bond yield and credit risk effects the bid-ask spread whereas volatility is not significantly related with the bid-ask spread, according to our results.

Our results about credit risk are coherent with Edwards et al. (2007), and our findings on yield support Ariyana et al. (2020), who report that term to maturity has no effect on liquidity, and yield has a significant effect on liquidity in the Indonesian sovereign sukuk market. Also, our results are similar with Galliani et al. (2014) and Feldhütter & Poulsen (2018) for credit risk. On the other hand, our finding on volatility does not confirm Bedowska-Sójka's (2019) results, and our results on bond duration does not support Bao et al.'s (2011) findings.

5. CONCLUSION

In this thesis, we examine two research questions about the sovereign bond liquidity in emerging countries, with an additional focus on the Turkish soverign bond market. The first question is about the low-frequency bid-ask spread measures' ability to proxy the real-time bid-ask spread, and the second question is about the bond and country specific characteristics' effects on the bid-ask spread. In order to answer these research questions, we implement empirical analysis with two distinct data sets.

The first data set consists of the data of the Turkish exchange-traded sovereign bonds, we investigate how effectively the low-frequency bid-ask spread measures are able to proxy bond liquidity, and how the bond-level features affect the bid-ask spread. Plenty of researches, which examine the efficiency of the bid-ask spread proxies for different asset classes and markets, exist in the literature. Nevertheless, there is no clear answer of the superior measure, and it is not obvious whether the previous results are valid for the exchange-traded bonds, because of the differences in asset features, market microstructures, and requirements of the market participants. So, we seek a suitable bid-ask spread measure, which is efficiently able to measure the bid-ask spread with less calculation burden and lower data requirement for the Turkish sovereign bonds, which are traded in BIST DSM.

According to the empirical results, CPQS is superior to the other proxies and able to represent the bid-ask spread dynamics. Less calculation burden and lower data requirement surpass the decline in the accuracy of estimates for the Turkish sovereign bonds traded in BIST DSM. The usage of daily data is adequate to accurately measure bond liquidity, in terms of the relative bid-ask spread. However, the results may not conform to distinct markets or asset classes, as bid-ask spread forecast generally differs across distinct market microstructures and ends up with different results.

Investment decisions and cost of capital is affected by the bid-ask spread, since asset liquidity imposes on transaction costs. Hence, these bond features may have significant implementations for the market participants. Therefore, we examine bond-level features' effects on the bid-ask spread with the first data set.

In order to analyze the bond-level features' effects on the bid-ask spread, a fixed-effect panel regression model is implemented. According to the panel regression results, the bonds features, which significantly affect the bid-ask spread, are trade volume, term to maturity, and bond type. Longer term to maturity or less trade volume results in a wider relative bid-ask spread. Additionally, the floating-rate bonds have narrower spreads than the fixed-rate bonds, on average.

For the second data set, which consists of sovereign bond data form six emerging countries, we examine the bond-level and country-level determinants of the relative bid-ask spread. Therefore, we implement a mixed effects panel regression model. According to the regression results, bond yield and issuer country CDS are significantly influential on the bid-ask spread. On average, higher country CDS, higher market return and lower bond yield result in a wider relative bid-ask spread.

Since liquidity costs are related with asset return and total trade cost, the results could be attractive for traders, issuers, and investors. Additionally, our empirical analysis has important results with regard to the relative bid-ask spread forecast, asset liquidity, and market structure, so the results could be beneficial for market regulators, as well. Moreover, as exchanges, market authorities, or other regulatory bodies intend to increase liquidity and reduce market frictions to enhance market efficiency and reduce costs, they could find the results appealing.

The evidence about the adequacy of the low-frequency relative bid-ask spread measures is coherent with some of the previous studies in the literature, such as the results of Goyenko et al. (2009), Chung and Zhang (2014), and Fong et al. (2017) on the US equity markets, Schestag et al.'s (2016) findings on the US fixed-income markets, the research of Marshall et al. (2018) on the ETF markets, and Gao et al.'s (2021) results on the Chinese stock markets.

Additionally, our findings on the relation between the bid-ask spread and bond features for the Turkish sovereign bond market are parallel to the findings of Edwards et al. (2007), and Feldhütter and Poulsen (2018) for the US OTC bond markets, and the findings of Galliani et al. (2014) in the EU bond markets.

Also, our results for the emerging country sovereign bond markets are coherent with the results of Feldhütter & Poulsen (2018) in the US bond markets. Besides, our findings are comparatively similar to Chen et al.'s (2007), and Edwards et al.'s (2007) results for the US OTC bond markets, Galliani et al.'s (2014) findings for the EU bond markets, and Ariyana et al. (2020) emerging country sukuk markets.

Even though, we achieve significant findings about the low-frequency measures, the findings may not be generalized for every financial markets, as market microstructure and asset specifications differ across asset classes. Additionally, short sampling period for the second data set may limit us to generalize our findings on emerging country sovereign bond markets to different bonds. Further research may involve a larger data set and include different aspects of bond markets, such as credit ratings, in order to obtain more robust and generalizable results.

REFERENCES

Abdi, F. & Ranaldo, A. (2017). A simple estimation of bid-ask spreads from daily close, high, and low prices. *Review of Financial Studies, 30,* 4437-4480. https://doi.org/10.1093/rfs/hhx084

Altay, E. & Çalgıcı, S. (2019). Liquidity adjusted capital asset pricing model in an emerging market: Liquidity risk in Borsa Istanbul. *Borsa Istanbul Review, 19,* 297-309. https://doi.org/10.1016/j.bir.2019.06.002

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5,* 31-56. https://doi.org/10.1016/s1386-4181(01)00024-6

Amihud, Y. & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17,* 223-249. https://doi.org/10.1016/0304-405X(86)90065-6

Amihud, Y., Mendelson, H., & Lauterbach, B. (1997). Market microstructure and securities values: Evidence from the Tel Aviv stock exchange. *Journal of Financial Economics, 45,* 365-390. https://doi.org/10.1016/s0304-405x(97)00021-4

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2006). *Liquidity and asset prices.* Now Publishers Inc.

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2012). *Market liquidity: asset pricing, risk, and crises.* Cambridge University Press.

Aquilina, M., & Suntheim, F. (2016). Liquidity in the UK corporate bond market: evidence from trade data. *The Journal of Fixed Income, 26*(3), 49-62.

Arcuri, M. C., Gandolfi, G., Monteux, M., & Verga, G. (2020). What Factors Influence European Corporate Bond Spread? *International Journal of Business and Management, 15*(4).

Ariyana, F. R., Arundina, T., & Kasri, R. A. (2020). Sukuk Market Liquidity Determinants: A Case Study on Sovereign Sukuk in Indonesia. *Institutions and Economies,* 66-90.

Baker, H. K. (1996). Trading location and liquidity: An analysis of US dealer and agency markets for common stocks. Blackwell Pubs..

Ball, C. A. & Chordia, T. (2001). True spreads and equilibrium prices. *The Journal of Finance, 56*(5), 1801-1835.

Bampinas, G., Panagiotidis, T., & Politsidis, P. (2020). Sovereign bond and CDS market contagion: A story from the Eurozone crisis.

Bao, J., Pan, J., & Wang, J. (2011). The illiquidity of corporate bonds. *The Journal of Finance, 66,* 911–946. doi:10.1111/j.1540-6261.2011.01655.x.

Batten, J. A. & Vo, X. V. (2014). Liquidity and return relationships in an emerging market. *Emerging Markets Finance and Trade, 50,* 5-21. https://doi.org/10.2753/ree1540-496x500101

Beau, D. (2014). Market-making and proprietary trading: Industry trends, drivers and policy implications. BIS.

Bedowska-Sojka, B. (2019). The dynamics of low-frequency liquidity measures: The developed versus the emerging market. *Journal of Financial Stability, 42,* 136-142. https://doi.org/10.1016/j.jfs.2019.05.006.

Bedowska-Sojka, B. & Echaust, K. (2019). Commonality in liquidity indices: The emerging European stock markets. *Systems, 7*. https://doi.org/10.3390/systems7020024.

Benos, E., Payne, R., & Vasios, M. (2020). Centralized Trading, Transparency, and Interest Rate Swap Market Liquidity: Evidence from the Implementation of the Dodd–Frank Act. *Journal of Financial and Quantitative Analysis, 55*(1), 159-192.

Bessembinder, H., Hao, J., & Zheng, K. (2020a). Liquidity provision contracts and market quality: Evidence from the New York Stock Exchange. *The Review of Financial Studies, 33*(1), 44-74.

Bessembinder, H., Spatt, C., & Venkataraman, K. (2020b). A survey of the microstructure of fixed-income markets. *Journal of Financial and Quantitative Analysis, 55*(1), 1-45.

Black, F. (1971). Toward a fully automated stock exchange, part I. *Financial Analysts Journal, 27*(4), 28-35.

Blau, B. M., Griffith, T. G., & Whitby, R. J. (2018). The maximum bid-ask spread. *Journal of Financial Markets, 41,* 1-16.

Bohmann, M., Michayluk, D., Patel, V., & Walsh, K. (2019). Liquidity and earnings in event studies: Does data granularity matter?. *Pacific-Basin Finance Journal, 54,* 118-131.

Bollen, N. P., Smith, T., & Whaley, R. E. (2004). Modeling the bid/ask spread: measuring the inventory-holding premium. *Journal of Financial Economics, 72*(1), 97-141.

Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2021). How to measure the liquidity of cryptocurrency markets?. *Journal of Banking & Finance, 124,* 106041.

Breusch, T. S. & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica, 47,* 1287. https://doi.org/10.2307/1911963

Chang, K., Feng, Y. L., Liu, W., Lu, N., & Li, S. Z. (2021). The impacts of liquidity measures and credit rating on corporate bond yield spreads: evidence from China’s green bond market. *Applied Economics Letters, 28*(17), 1446-1457.

Chen, F. & Stock, D. (2021). How a New Bond Issuance Affects the Liquidity of a Bond Portfolio. *The Journal of Fixed Income, 31*(2), 128-151.

Chen, L., Lesmond, D. A., & Wei, J. Z. (2007). Corporate yield spreads and bond liquidity. *The Journal of Finance, 62*, 119–149. doi:10.1111/j.1540-6261.2007.01203.x.

Choi, J. Y., Salandro, D., & Shastri, K. (1988). On the estimation of bid-ask spreads: Theory and evidence. *Journal of Financial and Quantitative Analysis, 23*(2), 219-230.

Chordia, T., Sarkar, A., & Subrahmanyam, A. (2004). An empirical analysis of stock and bond market liquidity. *Review of Financial Studies, 18,* 85-129. https://doi.org/10.1093/rfs/hhi010

Chung, K. H. & Zhang, H. (2014). A simple approximation of intraday spreads using daily data. *Journal of Financial Markets, 17,* 94-120. https://doi.org/10.1016/j.finmar.2013.02.004

Corwin, S. A. & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance*, *67,* 719-760. https://doi.org/10.1111/j.1540-6261.2012.01729.x

Darbha, M. & Dufour, A. (2013). Microstructure of the euro-area government bond market. In H.K. Baker, H. Kiymaz (Eds.), *Market microstructure in emerging and developed markets* (pp.39-58). Retreived from https://doi.org/10.1002/9781118681145.ch3.

Diaz, A. & Escribano, A. (2017). Liquidity measures throughout the lifetime of the U.S. treasury bond. *Journal of Financial Markets*, *33*, 42-74. https://doi.org/10.1016/j.finmar.2017.01.002

Diaz, A. & Escribano, A. (2019). Credit rating and liquidity in the US corporate bond market. *Journal of Fixed Income, 28,* 46-59. https://doi.org/10.3905/jfi.2019.28.4.046

Diaz, A. & Escribano, A. (2020). Measuring the multi-faceted dimension of liquidity in financial markets: A literature review. *Research in International Business and Finance, 51,* 101079. https://doi.org/10.1016/j.ribaf.2019.101079

Dick-Nielsen, J., Feldhütter, P., & Lando, D. (2012). Corporate bond liquidity before and after the onset of the subprime crisis. *Journal of Financial Economics, 103,* 471-492. https://doi.org/10.1016/j.jfineco.2011.10.009

Edwards, A. K., Harris, L. E., & Piwowar, M. S. (2007). Corporate bond market transaction costs and transparency. *The Journal of Finance, 62,* 1421-1451. https://doi.org/10.1111/j.1540-6261.2007.01240.x

Feldhütter, P. (2012). The same bond at different prices: Identifying search frictions and selling pressures. *Review of Financial Studies, 25,* 1155-1206. https://doi.org/10.1093/rfs/hhr093

Feldhütter, P. & Poulsen, T. K. (2018). What determines bid-ask spreads in over-the-counter markets? *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3286557

Fleming, M. J. (2001). Measuring treasury market liquidity. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.276289

Fong, K. Y. L., Holden, C. W., & Trzcinka, C. A. (2017). What are the best liquidity proxies for global research? *Review of Finance, 21,* 1355-1401. https://doi.org/10.1093/rof/rfx003

Foucault, T., Pagano, M., Roell, A., & Röell, A. (2013). *Market liquidity: theory, evidence, and policy.* Oxford University Press.

Galliani, C., Petrella, G., Resti, A., & Cazan, F. (2014). The liquidity of corporate and government bonds: Drivers and sensitivity to different market conditions. Technical Report Joint Research Centre of the European Commission. https://doi.org/10.2788/70146

Gao, Y., Zhao, W., & Wang, M. (2020). The comparison study of liquidity measurements on the Chinese stock markets. *Emerging Markets Finance and Trade,* 1-29.

George, T. J., Kaul, G., & Nimalendran, M. (1991). Estimation of the bid–ask spread and its components: A new approach. *The Review of Financial Studies, 4*(4), 623-656.

Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics, 21*(1), 123-142.

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92,* 153-181. https://doi.org/10.1016/j.jfineco.2008.06.002

Guloglu, Z. C. & Ekinci, C. (2016). A comparison of bid-ask spread proxies: Evidence from Borsa Istanbul futures. *Journal of Economics Finance and Accounting, 3,* 244-254. https://doi.org/10.17261/pressacademia.2016321992

Hameed, A., Helwege, J., Li, R., & Packer, F. (2019). Measuring corporate bond liquidity in emerging market economies: price-vs quantity-based measures. *BIS Paper,* (102e).

Han, S. & Zhou, H. (2007). Effects of bond liquidity on the nondefault component of corporate bond spreads: Evidence from intraday transactions data. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.981408

Harris, L. E., & Piwowar, M. S. (2006). Secondary trading costs in the municipal bond market. *The Journal of Finance, 61*(3), 1361-1397.

Hasbrouck, J. (2004). Liquidity in the futures pits: Inferring market dynamics from incomplete data. *Journal of Financial and Quantitative Analysis, 39*(2), 305-326.

Hasbrouck, J. (2009). Trading costs and returns for U.S. equities: Estimating effective costs from daily data. *The Journal of Finance, 64,* 1445-1477. https://doi.org/10.1111/j.1540-6261.2009.01469.x

Helwege, J. & Wang, L. (2021). Liquidity and price pressure in the corporate bond market: evidence from mega-bonds. *Journal of Financial Intermediation, 48,* 100922.

Ho, V. T. & Ho, T. N. P. (2020). Regulations in the US and bond market liquidity. *Journal of Economic Development, 45*(1).

Holden, C. W. (2009). New low-frequency spread measures. *Journal of Financial Markets, 12*, 778-813. https://doi.org/10.1016/j.finmar.2009.07.003

Hong, G., & Warga, A. (2000). An empirical study of bond market transactions. *Financial Analysts Journal, 56*(2), 32-46.

Huang, J., & Tian, Y. (2020). Research on Liquidity Risk of Corporate Bond Spread in China. *Engineering Letters, 28*(2).

Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies, 10*(4), 995-1034.

Hsiao, C. (2014). *Analysis of panel data* (No. 54). Cambridge university press.

Jahan-Parvar, M. R. & Zikes, F. (2019). When do low-frequency measures really measure transaction costs? *Finance and economics discussion series, 2019.* https://doi.org/10.17016/feds.2019.051

Karnaukh, N., Ranaldo, A., & Söderlind, P. (2015). Understanding FX liquidity. *Review of Financial Studies, 28,* 3073-3108. https://doi.org/10.1093/rfs/hhv029

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society,* 1315-1335.

Langedijk, S., Monokroussos, G., & Papanagiotou, E. (2018). Benchmarking liquidity proxies: The case of EU sovereign bonds. *International Review of Economics & Finance, 56,* 321-329. https://doi.org/10.1016/j.iref.2017.11.002

Lesmond, D. A. (2005). Liquidity of emerging markets. *Journal of Financial Economics, 77,* 411-452. https://doi.org/10.1016/j.jfineco.2004.01.005

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *Review of Financial Studies, 12,* 1113-1141. https://doi.org/10.1093/rfs/12.5.1113

Lybek, T. & Sarr, A. (2002). Measuring liquidity in financial markets. *IMF Working Papers, 2,* 1. https://doi.org/10.5089/9781451875577.001

Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do security prices change? A transaction-level analysis of NYSE stocks. *The Review of Financial Studies, 10*(4), 1035-1064.

Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2011). Commodity liquidity measurement and transaction costs. *Review of Financial Studies, 25,* 599-638. https://doi.org/10.1093/rfs/hhr075

Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2018). Do liquidity proxies measure liquidity accurately in ETFs? *Journal of International Financial Markets, Institutions and Money, 55,* 94-111. https://doi.org/10.1016/j.intfin.2018.02.011

Mizrach, B. (2015). Analysis of corporate bond liquidity. *Research Note, FINRA Office of the Chief Economist.*

O'Hara, M. (2007). Optimal microstructures. *European financial management, 13*(5), 825-832.

O'Sullivan, C. & Papavassiliou, V. G. (2020). On the term structure of liquidity in the European sovereign bond market. *Journal of Banking & Finance, 114,* 105777.

Phylaktis, K. & Manalis, G. (2013). Futures trading and market microstructure of the underlying security: A high frequency experiment at the single stock future level. *Borsa Istanbul Review, 13,* 79-92. https://doi.org/10.1016/j.bir.2013.10.012

Pu, X. (2009). Liquidity commonality across the bond and CDS markets. *Journal of Fixed Income, 19,* 26-39. https://doi.org/10.3905/jfi.2009.19.1.026

Remolona, E. & Yetman, J. (2019). The rise of benchmark bonds in emerging Asia. *BIS Paper,* (102g).

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance, 39,* 1127-1139. https://doi.org/10.1111/j.1540-6261.1984.tb03897.x

Schestag, R., Schuster, P., & Uhrig-Homburg, M. (2016). Measuring liquidity in bond markets. *Review of Financial Studies, 29,* 1170-1219. https://doi.org/10.1093/rfs/hhv132

Schultz, P. (2001). Corporate bond trading costs: A peek behind the curtain. *The Journal of Finance, 56*(2), 677-698.

Schwartz, R. A., & Francioni, R. (2004). *Equity markets in action: The fundamentals of liquidity, market structure and trading.* John Wiley & Sons. Inc., Hoboken, New Jersey.

Stange, S. & Kaserer, C. (2009). Market liquidity risk - an overview. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.1362537

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *the Journal of Finance, 44*(1), 115-134.

Wyart, M., Bouchaud, J. P., Kockelkoren, J., Potters, M., & Vettorazzo, M. (2008). Relation between bid–ask spread, impact and volatility in order-driven markets. *Quantitative finance, 8*(1), 41-57.

CURRICULUM VITAE

Name Surname : Emre SU

EDUCATION :

- **B.Sc.** : 2010, Istanbul Technical University, Management Faculty, Management Engineering Department
- **M.Comm.** : 2011, Macquarie University, Faculty of Business and Economics, Master of Commerce in Finance

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2017- … Specialist at Borsa Istanbul
- 2013-2016 Engineer at IETT

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Su, E., &** Tokmakcioglu, K. (2021) A comparison of bid-ask spread proxies and determinants of bond bid-ask spread. *Borsa Istanbul Review*, *21*(3), 227-238.
- **Su, E., &** Tokmakcioglu, K. (2023) Determinants of bid-ask spread in emerging sovereign bond markets. *Journal of Asset Management*, 1-7.

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- Ozcelebi, O., Tokmakcioglu, K., **& Su, E.** (2020). Revisiting the asymmetric impacts of the exchange market pressure on the inflation, interest rate and foreign trade balance in Eastern Europe. *Empirical Economics*, 1-22.