Universidade Federal do Rio Grande do Sul

Escola de Administração

Programa de Pós-Graduação em Administração

Henrique Pinto Ramos

Essays on liquidity in financial markets

Brazil

February, 2020

Ramos, Henrique Pinto
Essays on liquidity in financial markets/ Henrique Pinto Ramos. – Brazil, February, 2020-
111 p. : il. (algumas color.) ; 30 cm.

Supervisor: Marcelo Scherer Perlin

Tese (Doutorado) – Universidade Federal do Rio Grande do Sul
Escola de Administração
Programa de Pós-Graduação em Administração, February, 2020.

1. Liquidez. 2. Medidas de liquidez. 3. Microestrutura de mercado. 4. Negociação de alta frequência. I. Marcelo Scherer Perlin. II. Essays on Liquidity in Financial Markets.

Henrique Pinto Ramos

Essays on liquidity in financial markets

Dissertation submitted as a requirement for the Doctoral degree in Administration at the School of Administration at the Federal University of Rio Grande do Sul.

Marcelo Scherer Perlin, UFRGS
Supervisor

Marcelo Brutti Righi, UFRGS
Committee Member

Kelmara Mendes Vieira, UFSM
Committee Member

André Alves Portela Santos, UFSC
Committee Member

Alan de Genaro, FGV/EAESP
Committee Member

Brazil

February, 2020

Abstract

This dissertation is a collection of essays on liquidity in financial markets. As there are multiple facets of liquidity, I address this issue within three essays. The first essay entitled "Mispricing in the odd lots market in Brazil", and coauthored with professors Marcelo S. Perlin and Marcelo B. Righi, has been accepted for publication in the North American Journal of Economics and Finance (2018 JCR Impact Factor: 1.119, Qualis A2). We show that a peculiarity in the Brazilian financial market may harm investors depending on the platform on which they choose to negotiate. Investors pay higher prices for the same financial contract (i.e., shares of a company) if they trade at the odd lots market. This market is restricted to agents who trade between 1 and 99 shares of a company and, consequently, presents lower liquidity when compared to the round lots market. Using trade data we show that this mispricing is related to market returns, implied volatility and spreads. A trading strategy which takes advantage of the mispricing yields higher returns than our benchmark. Our results may be used both for regulators to rethink the objective of an odd lots market and for traders who seek to take advantage of this mispricing.

The second paper, entitled "Liquidity, implied volatility and tail risk: a comparison of liquidity measures" is coauthored with professor Marcelo B. Righi and has been accepted for publication in the International Review of Financial Analysis (2018 JCR Impact Factor: 1.693, Qualis A1). In this study, we present a comparison of liquidity variables in two empirical exercises using up to eight traditional liquidity measures and two new proposed variables in a sample of NYSE-listed stocks. The proposed proxies are based on semi-deviations of assets' returns. The first empirical exercise analyzes the relationship between liquidity and implied volatility, showing that increases in implied volatility increases illiquidity. Using a decomposition of the squared VIX, we show that both conditional variance and variance premium components affect liquidity. Our second empirical study investigates the relationship between common factors in liquidity and tail risk. Common factors increase individual stocks' Value-at-Risk and Expected Shortfall, although the effect is not significant for market risk. In both applications, most of the studied proxies present results aligned with the body of literature.

The third paper is "Does algorithmic trading harm liquidity? Evidence from Brazil", coauthored with professor Marcelo S. Perlin. The paper was submitted to the North American Journal of Economics and Finance (2018 JCR Impact Factor: 1.119, Qualis A2). The paper provides first evidence of algorithmic trading (AT) reducing liquidity in the Brazilian equities market. Our results are contrary to the majority of work, which has found a positive relationship between AT and liquidity. Using the adoption of a new data center for the B3 exchange as an exogenous shock, we report evidence that AT increases realized spreads in both firm fixed-effects and vector autoregression estimates for 26 stocks between 2017 and 2018 using high-frequency data. We also provide evidence that AT increases commonality in liquidity, evidencing correlated transactions between automated traders.

Keywords: Liquidity, Liquidity measures, Market microstructure, High-frequency trading.

Resumo

Esta tese é uma coleção de ensaios sobre liquidez nos mercados financeiros. Uma vez que há múltiplas facetas sobre o conceito de liquidez, o tema é investigado através de três artigos. O ensaio intitulado "*Mispricing in the odd lots market in Brazil*", escrito em coautoria com os professores Marcelo S. Perlin e Marcelo B. Righi, foi aceito para publicação na revista *North American Journal of Economics and Finance* (Fator de Impacto JCR 2018: 1.119, Qualis A2). No artigo, nós mostramos que uma especificidade no mercado financeiro brasileiro pode prejudicar investidores dependendo da plataforma que eles escolhem negociar. Investidores pagam preços mais altos para o mesmo contrato financeiro (ações de uma companhia) quando negociados no mercado fracionário. Este mercado é restrito a investidores que negociam entre 1 e 99 ações de uma empresa, consequentemente apresentando menor liquidez que o mercado tradicional (lote padrão). Usando dados de negociação, nós mostramos que este erro de precificação entre os dois mercados está relacionado com retornos do mercado, volatilidade implícita e *spreads*. Uma estratégia de negociação que se aproveita deste mecanismo gera retornos maiores do que o benchmark utilizado. Os resultados deste estudo podem ser usados tanto para reguladores repensarem o objetivo e a estrutura do mercado fracionário, bem como para *traders* que buscam aproveitar este erro de precificação.

O segundo manuscrito é intitulado "*Liquidity, implied volatility and tail risk: a comparison of liquidity measures*", tendo sido realizado em coautoria com o professor Marcelo B. Righi. O artigo foi aceito para publicação no *International Review of Financial Analysis* (Fator de Impacto JCR 2018: 1.693, Qualis A1). Nós apresentamos uma comparação de diversas variáveis em dois exercícios empíricos usando até oito *proxies* tradicionais para liquidez e duas *proxies* propostas com base nos semi-desvios dos retornos em uma amostra de ações listadas na NYSE. O primeiro exercício empírico analisa a relação entre liquidez e volatilidade implícita. Nós evidenciamos que aumentos na volatilidade implícita aumentam a illiquidez. Usando a decomposição do quadrado do VIX nós mostramos que tanto a variância condicional quanto o prêmio pela variância afetam a liquidez. No nosso segundo estudo empírico, nós investigamos a relação entre fatores comuns na liquidez e o risco de cauda. Fatores comuns na liquidez aumentam o risco de cauda individual quando medidos pelo *Value-at-Risk* e pela *Expected Shortfall*, ainda que o efeito não seja significativo para o risco de mercado. Em ambas as aplicações, a maior parte das *proxies* apresenta resultados alinhados com a literatura.

O terceiro artigo é intitulado "*Does algorithmic trading harm liquidity? Evidence from Brazil*". O paper foi submetido ao North American Journal of Economics and Finance (Fator de Impacto JCR 2018: 1.119, Qualis A2). O trabalho apresenta evidências sobre o efeito negativo de *algorithmic traders* (AT) sobre a liquidez de ações do mercado brasileiro. Os resultados são contrários a maioria dos trabalhos na área. Usando a data de início das operações de um novo datacenter da B3 como um choque exógeno, nós mostramos que AT aumenta *spreads* no mercado através de estimativas por efeitos fixos e vetores autorregressivos. A amostra é composta por 26 ações durante os anos de 2017 e 2018. Nós também evidenciamos o aumento da comunalidade na liquidez com o aumento da atividade de AT.

Palavras-chave: Liquidez, Medidas de liquidez, Microestrutura de mercado, Negociação em alta frequência.

Contents

Introduction

One of the several functions of financial markets is to channel capital into productive activities that demand resources in the economy. This simple transferring function promotes the country's economic development by improving competitiveness and employment (SHILLER, 2013). For this resource application to be efficient, both investors and invested agents emit and demand information regarding such investments. One of the main issues in financial literature is to understand how this information affects stakeholders and, consequently, affects asset prices in financial markets (FAMA, 1965; MALKIEL; FAMA, 1970). A challenging debate is whether prices reflect fundamentals of just noise originated by microstructure frictions. Seminal work such as that of Kyle (1985), Glosten e Milgrom (1985) and Easley e O'Hara (1987) highlight channels through which market frictions may affect the price formation process. Some of the issues studied in this topic are market design, asymmetric information, transaction costs, liquidity, order imbalance and price impact of trades. All of these are somehow related.

In this dissertation I study liquidity in financial markets through the presentation of three essays. Liquidity may be simply conceived as the ease of trading an asset without a penalty in price given the need for immediacy of an investor. In other words, liquidity is associated with how fast one can convert an asset to cash and vice-versa. In that sense, one can say that the notion of a financial market hinges on liquidity. Despite its relevance in finance, a formal and unique concept of liquidity is not easily defined. Some authors state that liquidity is a multidimensional variable, as the "ease of trading" may be represented in many ways (CHORDIA et al., 2000; FOUCAULT et al., 2013). For example, one can state that liquidity is the explicit cost of trading, namely spreads. Another can define liquidity as the ability for a given market to accommodate large orders without a penalty in price. Liquidity can also be related to a dynamic component, defined by how spreads and depth return to former levels of liquidity after large transitory shocks. None of these singular definitions is wrong, though none of them is complete and all of them are interconnected.

Even though liquidity is considered to be increasing through technological advances and facilitated trading, and its overall premium is diminishing (BEN-REPHAEL et al., 2015), liquidity still plays an important role in financial markets. The global financial crisis of 2007-2009 has demonstrated how a conjunction of factors, liquidity included, may harm investors (CHACKO et al., 2016). Usually liquidity is rather an effect itself than a cause of financial collapses. Thus, understanding liquidity is a major concern of policy makers and investors. Given the financial integration between markets and countries, shocks to assets' liquidity in one country or in different markets (stocks, bonds, etc) may disseminate to the entire market (AMIHUD et al., 2015).

Given the complexity of liquidity, I propose this dissertation as a collection of three essays on distinct empirical studies on distinct facets of liquidity. The first essay explores the difference of pricing in the odd lots market and in the round lots market in Brazil. The former allows one to trade between 1 and 99 shares of a company, whereas the latter allows only lots of multiples of 100 shares. Since the trade size constraint exists, the odd lots market has lower liquidity when compared to the round lots market. Also, it is expected that retail investors with capital constraints participate more in this market. Although on both platforms there are distinct order books, the financial contract that is being traded is the same (for example, common shares of a company).

In order to assess the difference between platforms, we have constructed a measure of mispricing. We compare the traded price for a given stock in the odd lots market against the closest traded price in the round lots market (at the maximum time difference of one second between platforms). To summarize, investors are paying higher prices in the odd lots market, where traded prices are on average 0.0188% higher than in the round lots market, through the 2013-2017 period. One can consider this a small mispricing, though the cumulative difference may be relevant over time as trading costs are considerably high in Brazil (SANVICENTE, 2012). Although financial theory would suggest that lower liquidity assets should have a discounted price (AMIHUD; MENDELSON, 1986), as their cost of capital is higher than that of a higher liquidity asset, we find the opposite. The mispricing is affected by implied volatility, market returns and spreads. In order to check the financial impact of this issue, we also propose a simple trading strategy that takes advantage of the mispricing. A rolling window (500, 250, 125 and 62 days) is defined in order to calculate mispricing averages and rankings through stocks in a given time window. In the following day, highly mispriced stocks are sold in the odd lots market and bought at the round lots market. Without including trading costs, the trading strategy outperforms the Ibovespa index in the sample period in all rolling windows. When trading costs of 0.1% per position are included, the trading strategy outperforms the benchmark only if a larger information set is considered (rolling window of 500 days). This result suggests that, although mispricing exists, an investor seeking to take profits with a similar trading strategy must consider the lack of liquidity in the odd lots markets.

Our objective with this paper is to evidence how a microstructure friction (two distinct platforms trading the same asset) may harm investors in the odd lots market. Such harm is assigned to liquidity differences and lot size constraints, as these are the fundamental differences between platforms. Therefore, the paper proposes rethinking the objective of an odd lots market in Brazil. The literature has been providing evidence on how odd lot trades may contribute to price discovery in financial markets (O'HARA et al., 2014), and thus we believe this issue may also receive attention in the Brazilian market. Recent studies also associate odd lot trades to HFTs (high-frequency traders) pinging for liquidity (DAVIS et al., 2017; JOHNSON et al., 2017).

The second paper of this dissertation proposes a comparison of multiple liquidity measures within two empirical exercises. Recent work in finance has faced the obstacle of liquidity measuring mostly in two ways: a) showing distinct aspects of liquidity regarding different specifications in asset-pricing related studies (ANTHONISZ; PUTNIŅŠ, 2016; AMIHUD et al., 2015; BELKHIR et al., 2018; BEN-REPHAEL et al., 2015; ABDI; RANALDO, 2017; GRILLINI et al., 2019; WU, 2019); and b) proposing new measures of liquidity and providing empirical applications using real data (CHACKO et al., 2016; FONG et al., 2018; NIETO, 2018). As new measures of liquidity are developed, new horseraces and comparisons between proxies are presented by researchers (GOYENKO et al., 2009; FONG et al., 2017; SCHESTAG et al., 2016). Therefore, our study addresses both strands of the literature: we propose new measures of liquidity based on the semi-deviation of assets' returns and study whether new and old liquidity measures are related to other strands of financial literature. Our two empirical studies are: a) connection with implied volatility and b) common factors in liquidity and tail risk.

In the first exercise, we estimate the effect of VIX over several illiquidity proxies and find a positive effect of the former in the latter. In order to detail the relationship between volatility and illiquidity, we present estimations switching the VIX measure for the conditional variance and the variance premium of the decomposed squared VIX as in Bekaert e Hoerova (2014). Using both time series and panel estimations, our results for this empirical exercise show that traditional liquidity measures are related to implied volatility, in which increasing implied volatility is related to an increase in illiquidity. The semi-deviation-based proxies also present this behavior. Additionally, we provide new insights on the dynamic relationship between liquidity and implied volatility using a vector autoregression (VAR) approach.

In the second application, we estimate the relationship between common factors in liquidity and tail risk. Our results show that direct measures of a stock's risk (Value-at-risk and Expected Shortfall) are related to common factors of our liquidity measures extracted by principal component analysis (PCA). We show that although common factors are not related to market risk, these common factors increase the individual risk of stocks. Therefore, we evidence a direct channel of how liquidity increases tail risk. Regarding our proposed proxies, semi-deviations capture the observations below the mean of returns, and therefore we believe they are sound proxies for understanding the relationship between common factors in liquidity and tail risk.

The third essay provides the first detailed evidence of how algorithmic trading (AT) affects liquidity in the Brazilian market. Although evidence on the issue in emerging markets is scarce, there is a large discussion in developed markets on whether this type of trader benefits or harms market conditions as a whole. On the positive side, algorithmic trading and high-frequency traders (HFTs) may increase liquidity as they may act as voluntary market makers, placing near-BBO (best bid and offer) orders (MENKVELD, 2013). In that sense, these traders may improve not only spreads, but also the depth of a market. These types of traders may also help stabilize the order book after the execution of large orders. The literature has provided evidence of liquidity-suppliers AT/HFTs increasing market quality in this sense (HENDERSHOTT et al., 2011; BROGAARD et al., 2014; BENOS; SAGADE, 2016). On the other side, AT/HFTs may also act as liquidity demanders by removing liquidity from the market. Hirschey (2017) shows that HFTs can anticipate price movements, therefore increasing trading costs for non-HFTs. If this consistently occurs, HFTs should reduce liquidity from the market and increase information asymmetry, acting as informed traders in the sense of Glosten e Milgrom (1985), and adversely select noisy traders' orders. HFTs may also reduce market quality in what is called 'low-latency arbitrage', in which these traders take advantage of order arrival time differences on distinct trading venues for the same (or correlated) assets. Another example of how fast traders may harm market quality includes the Flash Crash of 2010, where HFTs allegedly contributed to price drops as the algorithms understood the sell order of 75,000 E-mini S&P 500 contracts as a signal for a large sell-off (JONES, 2013; SEC, 2010b).

One may notice that fast traders hinge on a certain level of liquidity in order to execute their strategies (whether passive or aggressive). However, one may also think about whether these types of investors are present in emerging markets, where liquidity is low when compared to developed ones. On the other side, volatility is higher in emerging markets, which may attract aggressive AT/HFTs aiming to profit based on high price variations. Therefore, I believe understanding AT activity in a volatile market may be relevant in terms of regulation. Additionally, there is only one exchange in Brazil, so low-latency strategies cannot be executed. In order to assess AT activity, we use the proxy of Hendershott et al. (2011), which measures all traffic messages (order arrivals, changes, cancellations and trades) occurring for each US$ 100.00 of traded volume within a time interval, and the trades-message ratio of Malceniece et al. (2019). Data provided by the Brazilian exchange allows to track a specific order number, so it is possible to know if a specific order is being changed, canceled or traded. As previously defined, liquidity has many dimensions. Thus, we use proxies for the realized spreads and price-impact dimensions of liquidity.

Our results show evidence of algorithmic trading (AT) reducing liquidity in the Brazilian equities market. The results are contrary to the majority of previous work, which has found a positive relationship between AT and liquidity. Using the adoption of a new data center for the B3 exchange as an exogenous shock, we report evidence that AT increases realized spreads in both firm fixed-effects and vector autoregression estimates for 26 stocks between 2017 and 2018 using high-frequency data. We also provide evidence that AT increases commonality in liquidity, evidencing correlated transactions between automated traders.

We conduct tests to certify our exogenous shock as a valid instrument for the level of algorithmic trading. In order to remove noise from estimations, the level of AT is given by 2SLS estimates from a set of instrumental and control variables. Our firm fixed-effects estimations show the level of AT as increasing both realized spreads and price-impact variables. As most of the literature studies the effect of AT in liquidity, we also estimate the bidirectional effect using vector autoregression (VAR) models. Our results are consistent with lagged AT increasing spreads. Our results are weaker for the price-impact proxy, suggesting that AT do not trade based on private information (MESTEL et al., 2018). A methodological contribution from this study is the use of high-frequency data aggregated through 1-minute intervals. Most of the literature uses data aggregated on a daily basis. The very nature of AT is time-sensitive, and thus it is important to measure this variable on a high-frequency basis. Results are robust when data is aggregated in 5- and 15-minute intervals and on a daily basis.

Our study also addresses the relationship between AT and commonality in liquidity (CIL). On the one side, algorithmic traders could better parse firm-specific information. If information is quickly incorporated into prices, commonality is expected to drop (MORIYASU et al., 2018; MORCK et al., 2000). On the other side, if trading strategies from AT are correlated, an increase in CIL is expected as trades occur based on similar triggers for action. The literature provides evidence of correlated trading from HFT in the US market (BROGAARD, 2010) and in the FX market (CHABOUD et al., 2014). As one of the well-known strategies of AT/HFT is market-making, such traders are expected to trade not only one stock, but a basket of them. Thus, shocks in funding liquidity or in asset returns may force these voluntary market makers to create commonality through the liquidation of their positions (BRUNNERMEIER; PEDERSEN, 2009; HAGSTRÖMER; NORDEN, 2013; MENKVELD, 2013). Our approach is to measure intraday CIL through the R squared of a regression of liquidity proxies on market liquidity. AT activity increases commonality in realized spreads, suggesting that algorithms may present correlated trading strategies. The Brazilian market has a small number of liquid stocks compared to other markets, restricting options to trade, and therefore inducing CIL.

This dissertation intends to contribute to the understanding of how several aspects of liquidity may affect financial markets. Thus, one byproduct of this work is to generate high-impact academic papers. The first paper has been published in the North American Journal of Economics and Finance (JCR Impact Factor: 1.119), the second has been accepted for publication in the International Review of Financial Analysis (JCR Impact Factor 2018: 1.693). The third paper is in first round of review for minor revisions for the North American Journal of Economics and Finance (JCR Impact Factor: 1.119).

The remainder of this dissertation proceeds as follows: Part I presents the article "Mispricing in the odd lots market in Brazil", Part II presents 'Liquidity, implied volatility and tail risk: a comparison of liquidity measures" and Part III presents "Does algorithmic trading harm liquidity? Evidence from Brazil".

Part I

Mispricing in the odd lots market in Brazil

Abstract

We study the case of mispricing in the odd lots equity market in Brazil. Contrary to expectation, odd lot investors are paying higher prices than round lot investors. The pricing difference between markets is affected by market returns, volatility and spreads. Our main hypothesis is that; once the assets traded in the odd lot market are more illiquid than their counterparts, the mispricing is driven by liquidity factors. Additionally, we show that the mispricing yields an arbitrage opportunity that is not being traded away in the Brazilian market. Therefore, we propose regulators to review the market design for odd lots in Brazil. We argue that reducing the minimal trading unit in the round lots market would benefit investors.

Keywords: Odd lots, market microstrure, liquidity, BM&FBovespa.

Note: this article has been accepted for publication on the North American Journal of Economics and Finance (2018 JCR Impact Factor: 1.119, Qualis A2), volume 42, 2017.

1 Introduction

An odd lot trade refers to a negotiation in which the traded quantity is smaller than the standard (round) lot[1]. Many exchanges define the standard lot as 100 stocks of the equity pool of a certain company. The possibility of odd lot trading facilitates retail investor's access to stock markets by reducing the minimum amount of cash necessary to trade. In the US market, these trades were not reported in the consolidated tape data until December 2013. A long-dated belief has considered odd lotters as uninformed retail investors (LAKONISHOK; MABERLY, 1990; DYL; MABERLY, 1992; FOUCAULT et al., 2011; AHN et al., 2014).

However, after reporting odd lot trades in TAQ (Trade and Quote) databases, research has shown that maybe odd lot investors were not so uninformed and maybe they are not even retail investors, but high-frequency traders (HFT) with relevant participation in the market. O'Hara et al. (2014) show that odd lot trades contribute 35% of price discovery, which is consistent with the hypothesis of informed traders using odd lots to gather information. Davis et al. (2017) provide evidence that investors use 1-share trades to ping for hidden liquidity on NASDAQ. Other studies confirm the potential information contained in odd lots within the US market (BATTALIO et al., 2017; JOHNSON, 2014; JOHNSON et al., 2017; ROSEMAN et al., 2016).

Despite the recurring studies regarding odd lots in the US market, exchanges in many other countries such as Italy, Israel, South Korea, Canada, Taiwan, Philippines, Singapore, Mexico and Brazil allow this type of trading (GOZLUKLU et al., 2015). The microstructure of odd lots trading in these and other countries is, however, undocumented. Therefore, this study aims to shed light on a particular case of an emerging economy, namely Brazil. In the Brazilian equities market, odd lot trades have a completely separate platform. Symbols, order books and consequently traded prices differ from one market to another, though the financial contract is the same (i.e., a common share of a given company)[2].

Even though prices in both markets are highly correlated, it is possible that they diverge during some periods. If traded prices at both markets are different, such an effect should be due to microstructure effects. Despite that, empirical data of our study clearly shows that odd lot traders are consistently paying higher prices than round lot investors for the same financial contract. We use an intraday dataset from the end of 2013 to the beginning of 2017 to construct average mispricing measures for the market as whole and for individual assets. Our main finding is that mispricing occurs and it is positive in the Brazilian market for most of the assets in our sample. Simulating real-world frictions such as transaction costs, we show that this mispricing is not being traded away, yielding an arbitrage opportunity. The detachment of prices is affected by implied volatility, spreads and market downturns. This is consistent with a large reported hypothesis in literature of liquidity being reduced after increases in volatility and decreasing

[1] Sometimes referred as a board lot.

[2] For example, a round lot of Petrobrás common share (symbol: PETR3) is traded at the odd lots market with a different ticker (PETR3F). Every stock trading in the odd lots market has an additional F to its regular symbol. Additionally, an investor may buy cumulate odd lot shares and sell them as a round lot.

in market returns (CHORDIA et al., 2001; BRUNNERMEIER; PEDERSEN, 2009; NAGEL, 2012). Such discrepancies in prices could be reduced with merging the markets or, reducing the minimum trading unit in Brazil. Additionally, as a characterization of odd lot trades, we show that they are concentrated in small shares such as 1 and 10 lot sizes. Although small-share trades should be expected by retail investors with capital constraints, recent literature shows that small trades may also indicate high-frequency traders "pinging" for liquidity (DAVIS et al., 2017).

The Brazilian exchange, BM&FBovespa, is the largest exchange in Latin America and it is an interesting case to observe a pure effect of microstructure on prices. Assets with the same fundamental value may only differ in prices given frictions in the trading process. The main differences in the same stock negotiated at both odd lot and round lot markets are liquidity differences[3]. In short, if there were no market frictions or differences in liquidity, there would be no reason for prices to be different. Moreover, the existence of a single exchange, BM&FBovespa, makes it easier to study the structure of the whole market, unlike other highly segmented markets such as the US. News and other factors affecting fundamentals of a company should be the same for the asset traded in the odd lots market and in the round lots market. Also, as pointed out by Bekaert e Harvey (2002), studies on emerging markets are put aside many times, although there is a wide opportunity for research and, consequently, improvements in these markets. The access to information and to databases in these economies are sometimes difficult to obtain, hampering the progress of such fields of study. Therefore, we intend to contribute to the literature in market microstructure and in emerging markets in two ways: firstly by exposing that the market design for odd lots in Brazil implies in a pricing error when compared the same assets traded in the round lots market. Knowing and exposing different market designs for odd lots is the first step for improvement, which may have an impact over themes considered important for emerging markets literature, such as market efficiency and risk premium (KEARNEY, 2012). Apparently, other exchanges such as the Taiwan Stock Exchange and the Singapore Stock Exchange use a market design similar for odd lots. Therefore, the conclusions here can be extended to other scenarios. To the best of our knowledge, there is no research studying odd lot trades outside the US market apart from the current study. Our second contribution is to show that an arbitrage opportunity is present in the Brazilian market. Based on a simple trading strategy, we show that for highly mispriced stocks, the difference in closing prices can reach 150% over approximately 3 years. When trading costs are included, the strategy may also generate profits. This result brings up the question: is the odd lots market good for Brazilian investors? A suggestion for improvement of the Brazilian market discussed in the paper would be to reduce the minimum trading unit (MTU). For some countries, reducing MTU has improved traded prices and volume, and spreads have become lower (AMIHUD et al., 1999; HAUSER; LAUTERBACH, 2003; ISAKA, 2014; AHN et al., 2014). Another suggestion for regulators is to merge both markets.

The paper proceeds as follows: Section 2 defines both data and methodology. In Section 3, we present the results. Section 4 includes suggestions for the odd lots market in Brazil and our concluding remarks.

[3] For example, mean financial volume is much lower in the odd lots market (R$ 19,325.58) than the same measure calculated at the round lots market (R$ 24,655,600.00).

2 Data and Methodology

We gather data from BM&FBovespa's public ftp website using the R package *GetHFData* (PERLIN; RAMOS, 2017), which allows easy and free access to trade data from the Brazilian equity market. Our dataset contains all trades in the equities cash market. The sample comprises data from 2013-10-01 to 2017-03-14. We include all stock tickers available in the ftp website, accounting for 231 stocks. Daily data was also collected in the BM&FBovespa website. Data from market capitalization was retrieved using Economatica® software.

A simple methodology was employed to compare trade prices between markets: for each trade in the odd lots market of a given stock, a comparable trade was matched in the round lot market for the same asset. A trade is defined comparable if the absolute time difference between trades does not exceed one second. All trades with time differences larger than one second are removed from our sample, resulting in 3,956,837 observations. This approach assures that the possible price differences are not due to time effects[4]. Our mispricing variable is defined by the following equation:

$$Mispricing_{i,t,d} \quad = \quad \frac{POdd_{i,t,d} - PRound_{i,t^*,d}}{PRound_{i,t^*,d}} \tag{2.0.1}$$

where:

$POdd_{i,t,d}$	– trade price of stock i at time t of day d in the odd lots market
$PRound_{i,t^*,d}$	– trade price in the round lots market, also for stock i but for time t^* of day d
t^*	– closest time of a trade respecting the limit of one second.

The difference between markets is scaled by the price traded in the round lot market, and the difference between time periods t is not needed to be the same. Note that the measure is a percentage difference between markets, allowing cross section inferences between companies. Similar to other studies related to the liquidity literature (AMIHUD, 2002; KAMARA et al., 2008), we have removed observations below 1% percentile and above 99% percentile for the mispricing, traded prices and traded volume. The procedure aims to remove possible outliers which may bias our results.

The proposed approach also provides an intuitive understanding of the two markets. If equation 2.0.1 returns a positive value, it means that the traded price in the odd lots market is higher than the respective round lots price. A negative value means that the round lot stock has a higher price. The latter result is associated with the hypothesis widely accepted in the literature, stating that assets with higher illiquidity should have its price discounted at a higher rate and hence, present a lower price (AMIHUD, 2002; ACHARYA; PEDERSEN, 2005).

[4] We also test all procedures presented in this paper using time differences of 2 and 5 seconds, which also reduces the impact of time. The results remain qualitatively the same.

The first step in the research is to analyze the differences from one market to the other. Firstly, we perform simple t-tests and Wilcoxon Rank Sum tests in order to investigate whether the difference between traded prices is significantly different from zero. Hauser e Lauterbach (2003) and Gozluklu et al. (2015) used the same approach to investigate the impact of changes in prices after MTU reductions. Next, as in Hendershott et al. (2011), we divide the assets in five quintiles of asset's characteristics, where quintile 1 refers to smaller values and quintile 5 to the largest values of a given characteristic. This classification is made in terms of mean traded price, mean daily volume, mean daily number of trades and mean market capitalization. Apart from market capitalization, these measures use data from the odd lots market. After this division, we run t-tests and Wilcoxon Rank Sum tests for each quintile so we can understand if different assets' characteristics presents higher or lower mispricing.

In order to understand whether exogenous factors affect the difference between odd lot and round lot markets, we employ both time series and panel data regressions. Regarding the first approach, an equally-weighted index is created to account for overall mispricing. For each asset in our sample, the daily mean of intraday mispricing was calculated using:

$$Index_d = \frac{\sum_{i=1}^{N} \overline{Mispricing}_{i,d}}{N}, \tag{2.0.2}$$

where $\overline{Mispricing}_{i,d}$ refers to the average of mispricing calculated from trades in equation 2.0.1 for asset i at day d. Next, we estimate OLS regressions for our index against five pricing factors and a proxy for the Brazilian implied volatility (IVol) constructed by Astorino et al. (2017)[5]. We report robust standard errors from Newey e West (1986). As the difference of prices in odd lot and round lot trades has not been studied so far (to the best of our knowledge), we opt by testing the impact of variables that have been documented as affecting the stock market. In order to maintain the scale along with other variables, the natural logarithm was used in the estimations at section 3 for the IVol series. Although one could suggest the use of log-differenced implied volatility, we are interested in periods where the level of volatility is high or low, not when the variation of volatility is high or low. Other used factors are: MKT (market returns minus the risk-free rate), SMB (small minus big), HML (high minus low), WML (winners minus losers) and the IML (illiquid minus liquid) factor[6]. We also included a proxy for asset's spread in the round lot market, as defined by the Corwin e Schultz (2012) estimator (see Appendix A). Following the same procedure from the authors, all negative spreads, which correspond to a very small proportion of the dataset, are set to zero. The equation 2.0.3 describes our estimation by OLS:

$$\begin{aligned} Index_d = {} & \alpha + \beta_1 IVol_d + \beta_2 MKT_d + \beta_3 SMB_d + \beta_4 HML_d + \beta_5 WML_d + \\ & \beta_6 IML_d + \beta_7 Spread_d + \varepsilon_d, \end{aligned} \tag{2.0.3}$$

where α is the model intercept, and β_1 to β_7 are the coefficients regarding the volatility, the factors and the proxy for spread previously described. The notation ε_d refers to the error term

[5] The authors calculate the 2-month expected volatility of the Ibovespa index based on the next two expiration dates over the respective options' contracts. The approach is similar to the Carr e Wu (2008) model.

[6] Both the factors and the implied volatility data are retrieved from USP's NEFIN group (www.nefin.com.br).

at day d. For the second approach, we regress the average mispricing for every asset i in our sample against volatility, the pricing factors and the proxy for spread in a fixed effects panel regression. As usual in financial research, we believe that unobserved heterogeneity may be present within our sample (GORMLEY; MATSA, 2014). Although the research problem is not the same, O'Hara et al. (2014) and Roseman et al. (2016) also employed fixed effects estimations on odd lot trading[7]. The equation 2.0.4 shows the model to be estimated:

$$\begin{aligned}\overline{Mispricing}_{i,d} =& \alpha_i + \beta_1 IVol_d + \beta_2 MKT_d + \beta_3 SMB_d + \beta_4 HML_d + \beta_5 WML_d + \\ & \beta_6 IML_d + \beta_7 Spread_d + \epsilon_{i,d},\end{aligned} \tag{2.0.4}$$

where α_i is the level effect for asset i, and $\epsilon_{i,d}$ is the error term for asset i at day d. In addition to testing the effects of volatility, pricing factors and the spread to the mispricing in odd lots, our approach allows to compare results between time series and longitudinal data. We expect the results from both estimations to converge.

Finally, we estimate a simple trading strategy in order to verify the size and robustness of the mispricing in the Brazilian odd lots equities market. We define a rolling window scheme where w refers to the time window in days $w = \{500, 250, 125, 62\}$[8]. For a given value of w, we calculate the average mispricing for all assets in our sample, dividing them by quintiles and then selecting those with the highest quintiles (higher mispricing). Next, we sell the odd lot stocks with the highest mispricing and buy the same assets in the round lot using daily closing prices at $w + 1$. The simulated procedure would be equivalent to buy a round lot (i.e., 100 shares) and sell a combination in the odd lots market that sums up to 100 shares (i.e., 99+1 shares). We compare the results of such strategy to a buy and hold naive strategy, using the Ibovespa index. All calculations are net of the risk-free rate. We also test the profitability of this strategy when trading costs of 0.1% are considered. The estimation window rolls until the last day in our sample.

[7] Hausman (1978) tests were conducted in order to choose between random and fixed effects estimators. For the sake of brevity, the results are not exhibited in this paper.

[8] Roughly 2 years, 1 year, half year and a quarter of year, respectively.

3 Results

Table 1 presents summary statistics for the variables used in this study. The mean pricing error resulted by equation 2.0.1 is positive, with value of 0.0188%. While it may seem small at first sight, remember that the nominal effect of the mispricing accumulates over a high number of trades, for a large pool of investors. Financial theory suggests that, given higher iliquidity, the price in the odd lots market should be lower than in the round lot (AMIHUD; MENDELSON, 1986; AMIHUD, 2002; ACHARYA; PEDERSEN, 2005). However, Table 1 shows that, for BMF&Bovespa, trades occurring in the odd lots market have higher prices than its counterpart in the round lot market. After removing extreme cases, the pricing error ranges from the minimum value of -1.2689% to the maximum value of 1.5385%, which can also be seen in Figure 1. The equally-weighted index defined by equation 2.0.2 shows the same value for both mean and median (0.0182%) but with less extreme values compared to the percentage pricing error.

Table 1 also shows that lot size has a median value of 14, which evidences a large concentration of small-lot trades within odd lots. The volume traded at odd lots presents a large variation between extreme values, where the minimum value is R$ 3.90 and the maximum value is R$ 240,670.51 in a trading day.

Table 1 – Summary statistics

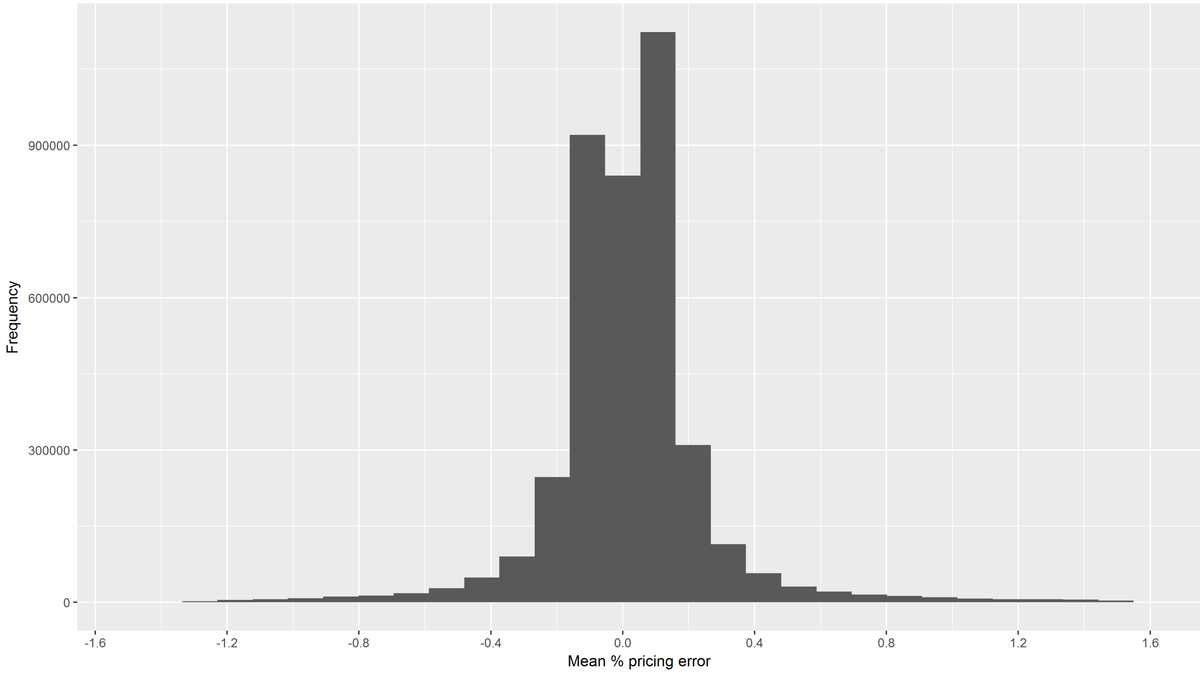
Table presents summary statistics for the variables related to odd lot trades. % pricing error refers to the result of equation 2.0.1, Index refers to the average mispricing for all assets resulted by equation 2.0.2. Lot size refers to the traded quantity of odd lot shares negotiated in comparable trades, Volume to the daily traded volume at the odd lot market. Volume is exhibited in Brazilian Reals (R$). MKT (market-minus-risk-free), HML (high-minus-low), IML (illiquid-minus-liquid), SMB (small-minus-big), WML (winners-minus-losers) are pricing factors for the Brazilian market. IVol is a proxy for implied volatility in Brazil constructed by Astorino et al. (2017). Spread is the Corwin e Schultz (2012) spread estimator constructed based on prices in the round lot market.

	Mean	Median	Min	Max	Sd
% pricing error	0.0188	0	-1.2689	1.5385	0.2375
Index	0.0182	0.0182	-0.2200	0.3150	0.0385
Lot size	24.3755	14	1	99	24.2861
Volume	19325.5787	4914.825	3.9	240670.51	35542.6605
Spread	0.0111	0.0063	0	1.086	0.0178
MKT	0.0001	-0.0005	-0.0443	0.0613	0.0139
HML	0.0001	-0.0002	-0.0387	0.0433	0.0098
IML	-0.0005	-0.0006	-0.0347	0.0312	0.0088
SMB	-0.0005	-0.0005	-0.0342	0.0362	0.0087
WML	0.0005	0.0009	-0.0819	0.0528	0.0109
IVol	24.4228	23.4502	13.8783	44.4336	4.7611

Figure 2 depicts the frequency of trades by lot size and their respective mean pricing error. Notably, one can see the majority of trades between 10 and 1 shares. This characteristic in the Brazilian market is different than what was found by O'Hara et al. (2014) on NASDAQ, where the most frequent odd lot size was 50. However, we note that 50 shares is the fifth most traded lot. In the right axis, the solid line exhibiting mean pricing error for each lot size shows a

Figure 1 – Mispricing histogram

The figure shows the histogram for the mispricing defined by equation 2.0.1. The mispricing is defined as the percentage difference between trades in the odd lots and round lots markets. Sample period ranges from 2013-10-03 to 2017-03-14.



curious pattern: the mispricing is positive and higher when it is close to lot sizes multiples of 5, but decreases as overall lot size increases. This pattern is valid until the lot size is close to 80, where an erratic behavior is presented afterwards. Specifically at lot size 91, the mispricing reaches its minimum value (-0.046%).

Based on Figure 2, evidence is unclear regarding the type of trader in the odd lots market. One could affirm that the concentration of trades below lot size of 10 is in line with retail investors facing financial restrictions to invest in a round lot. However, recent literature has shown that traders, mainly HFTs, may use small-lot trades to "ping" for hidden liquidity in the market (DAVIS et al., 2017). In this line, odd lot trading is associated with HFTs (JOHNSON et al., 2017) and aggressive traders (DAVIS et al., 2017). If HFT are engaging in the odd lots market, a simple strategy could rely on buying/selling a stock at the odd lot market whenever such asset is more cheaper/expensive than its counterpart at the round lots market. Other reasonable strategy would be for a HFT to place (and constantly update) limit orders with values higher/lower than the current values at the round lots order book. Thus, HFTs would wait for retail investors to send market orders with higher/lower prices than in the round lots market. As the positive mispricing exhibited at Table 1 means a higher buying pressure in odd lots, we hypothesize that HFT may be placing sell orders at slightly higher prices than the round lots market. Therefore, retail investors may act as aggressive traders paying for this higher price and HFTs end up earning the spread. If this occurs, HFTs act as market makers by providing liquidity at the cost of the spread plus a premium between odd and round lot markets. Although the evidence on HFT acting in Brazil is scarce, in developed markets HFTs are portrayed to provide liquidity as voluntary market makers (BROGAARD et al., 2014; MENKVELD, 2013).

Figure 2 – Lot size and mispricing

The figure depicts the frequency of lot size and mean pricing error over the sample of comparable trades. The barplot shows the frequency of trades to the respective lot size in the x-axis. The solid line draws the mean pricing error of the respective lot size.

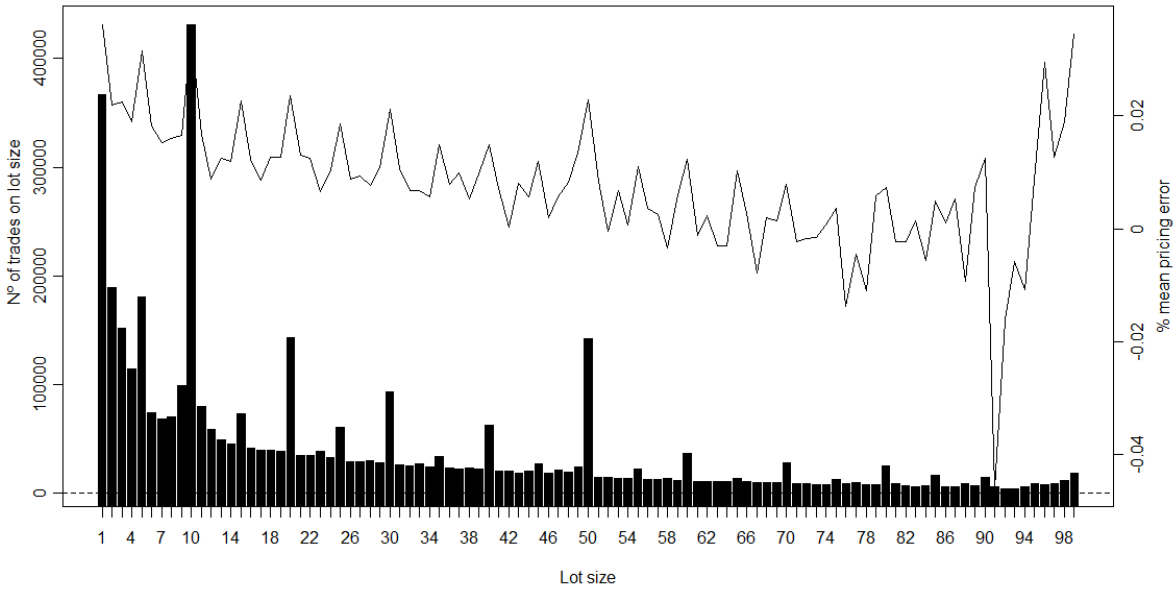


Table 2 reports test statistics for the null hypothesis of zero mean for the t-test for the overall (pooled) mispricing defined by equation 2.0.1. Although the mean value for the average pricing error is small (0.0188% per trade), it is highly significant with a t statistic of 157.7. Confirming the evidence on Table 1, the difference between odd lot and round lot prices is positive and highly significant. We also calculate the mean pricing error for each asset in our sample. About 80% (186 assets) present a positive mispricing from intraday data, where 65% of these assets (151) present positive and significant t and Wilcoxon Rank Sum statistics at the 5% level[9]. Note that the number of assets with significant positive mispricing is the same when t tests and Wilcoxon Rank Sum tests are employed. The percentage of assets presenting both significant negative mean and median mispricing which are significantly different than zero is close to 10% (23 assets for the t test and 25 for the Wilcoxon Rank Sum test). Thus, we have strong evidence that odd lot traders are paying higher prices on the same assets traded within round lots at almost the same time.

In order to understand which characteristics are related to the mispricing reported on Table 1 and Table 2, we divide our sample within quintiles of asset's characteristics: mean market capitalization, mean traded price, mean traded volume and mean number of trades. Table 3 shows the mean pricing error for these quintiles and their respective p-values for the t tests and Wilcoxon Rank Sum tests. Notably, the first two quintiles of market capitalization are increasing

[9] Results for each asset are available under request. This result remain qualitatively unchanged when the significance level is altered to 1%.

Table 2 – Summary of differences between odd lot trades and round lot trades

The table reports test statistics for the t-test (Wilcoxon Rank Sum) test for the null hypothesis of zero mean (median) for the $Mispricing_{i,t,d}$ variable (see Equation 2.0.1). Separating at the asset level, the table also shows the percentual with positive and negative mispricings. Additionally, we report the percentual of assets with positive and negative mispricings that present mean (median) statistically different than zero at the 5% level.

	Value
Mean (%)	0.0188
Median (%)	0.0000
T stat	157.7005
% above zero	80.52%
% below zero	19.48%
% above zero (t)	65.37%
% above zero (Wilcoxon)	65.37%
% below zero (t)	9.96%
% below zero (Wilcoxon)	10.82%
Total observations	3,956,837
Total assets	231

in pricing error but decreasing in the remaining quintiles. A similar effect occurs within price quintiles. We expect that larger and higher priced companies present a lower pricing error. Larger companies should be less affected by liquidity shocks (AMIHUD, 2002; BEN-REPHAEL et al., 2015), diminishing the effects studied in this paper. Furthermore, small priced stocks should be more affected by the mispricing in odd lots since the minimum tick size for the Brazilian market is R$ 0.01. So, a R$ 0.01 mispricing yields a higher percentage pricing error in small priced stocks when compared to more expensive ones.[10]

When mispricing is calculated in quintiles of volume and number of trades, we note a decreasing pattern: low-volume and low-negotiated assets (lower quintiles) present higher pricing errors. In the first quintile of both volume and number of trades, the pricing error is 0.193%, which is more than 10 times the mean pricing error for the overall sample (0.018%).

Table 4 reports the results for both time series and fixed effects panel regressions for the mispricing measures as dependent variables. Using equation's 2.0.2 index as response variable, column 1 shows the coefficients and standard errors for the OLS estimation. Notably, the proxy for implied volatility of the Brazilian market has a positive (0.014) effect on the mispricing index (at the 10% significance level). If the cause of mispricing is the lack of liquidity, when volatility rises a decrease in liquidity is expected, hence increasing the mispricing. This is in line with evidence that volatility dries up liquidity (BRUNNERMEIER; PEDERSEN, 2009; NAGEL, 2012). The coefficient for the market factor (-1.020) shows that when market returns decreases, the mispricing rises. We hypothesize that is also indirectly connected to liquidity, since when market returns are low, a decrease in liquidity is expected (CHORDIA et al., 2001). Also, the HML factor coefficient is significant (at the 10% level). Our proxy for spreads is positive (1.340)

[10] Suppose two stocks A and B presenting a pricing error the size of the minimum tick. If stock A is traded at the round lot market at R$ 1.00, the percentage pricing error is 1%. If stock B is traded at R$ 10.00, the percentage pricing error is 0.1%. All else equal, a small absolute pricing error yields a larger percentage pricing error in small priced stocks.

Table 3 – Mean percentual error by quintile

The table reports the mean mispricing over the sample period divided by assets' characteristics quintiles. Every asset is classified in one of the five quintiles of the distribution for the mean characteristic. Quintile 1 refers to the lowest measure and quintile 5 to the highest. Price refers to the mean traded price using intraday data, Market cap refers to the mean market capitalization, Volume refers to the mean daily traded volume in odd lots, Number of trades refers to the mean number of trades in the odd lots. T-test (Wilcoxon Rank Sum) p-values are reported for the null hypothesis of the respective mean (median) of average mispricing equals to zero.

Market cap	1	2	3	4	5
Mean % pricing error	0.029	0.066	0.036	0.020	0.015
T-test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Wilcoxon test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
N Obs	20,042	85,324	311,084	778,947	2,761,440
Price	**1**	**2**	**3**	**4**	**5**
Mean % pricing error	0.027	0.039	0.028	0.019	0.009
T-test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Wilcoxon test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
N Obs	77,305	189,728	961,400	1,451,633	1,276,771
Volume	**1**	**2**	**3**	**4**	**5**
Mean % pricing error	0.193	0.069	0.036	0.025	0.015
T-test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Wilcoxon test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
N Obs	7,128	33,475	173,319	719,512	3,023,403
Number of trades	**1**	**2**	**3**	**4**	**5**
Mean % pricing error	0.193	0.124	0.035	0.021	0.017
T-test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Wilcoxon test	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
N Obs	3,350	26,252	137,432	624,097	3,165,706

and significant, showing a direct measure of liquidity associated to our mispricing variable. Our main hypothesis that prices are higher in the odd lots market given its higher liquidity risk appears to fit our sample.

Column 2 reports results for the panel regression using fixed effects for equation's 2.0.1 pricing error as dependent variable. In comparison to the OLS estimations, the coefficient for the implied volatility (IVol) is also positive (0.025) but significant at a higher level (1%). With lower coefficients, the market factor (-1.340) and the HML factor (-0.310) also present negative and significant effects over the mispricing measured at the firm level. The SMB factor is negatively significant in the fixed effects estimations. This may be connected to the results of Table 3, which shows that larger firms (in terms of market capitalization) present lower mispricing. If the SMB factor decreases (whether by smaller firms presenting lower returns or larger firms presenting higher returns), a higher pricing error is expected. As in our time series regressions, the coefficient for the spread proxy is positive (0.282) and highly significant, evidencing the lack of liquidity affecting the difference between odd lot and round lot prices. One could expect a significant coefficient on the IML factor, measured by the Acharya e Pedersen (2005) measure. However, we believe that no effect is significant given that this is a price impact measure of liquidity. Notably, a transaction cost measure (spread) appears to be more relevant in affecting the mispricing.

Results from Tables 2, 3 and 4 show that, odd lots are traded at higher prices mostly in undesired moments of financial markets (high volatility, high spreads and market negative

Table 4 – Index (OLS) and mispricing (Panel Data) estimations

The table reports regression results for variables representing the mispricing between odd lot and round lot markets in Brazil. The explanatory variables are: the logarithm of the proxy for Brazilian implied volatility (IVol) constructed by Astorino et al. (2017), MKT (market excess return factor), SMB (small minus big factor), HML (high minus low factor), WML (winners minus losers factor) and IML (illiquid minus liquid factor). The Spread proxy is the Corwin e Schultz (2012) high-low spread estimator. For the OLS estimation, Spread is defined as the daily average from all assets' spreads. For the fixed-effect regression, the spread is measured for every asset. In column 1, the dependent variable is equally-weighted index for the daily mean of average mispricing using intraday data (equation 2.0.2). The estimation is performed by a simple OLS estimation. In column 2, we estimate a fixed effects panel regression in which the dependent variable is the mean mispricing over comparable trades for each asset generated by equation 2.0.1. Heteroskedasticity-autocorrelation consistent standard errors of Newey e West (1986) are reported in parenthesis.

	Index (OLS)	Mispricing (Panel Data)
	(1)	(2)
log(IVol)	0.014^{*}	0.025^{***}
	(0.008)	(0.007)
MKT	-1.020^{***}	-1.340^{***}
	(0.103)	(0.142)
HML	-0.246^{*}	-0.310^{**}
	(0.141)	(0.155)
IML	0.230	0.040
	(0.273)	(0.238)
SMB	-0.012	-0.473^{**}
	(0.277)	(0.236)
WML	0.164	-0.055
	(0.108)	(0.133)
Spread	1.340^{***}	0.282^{***}
	(0.464)	(0.109)
Constant	-0.040	
	(0.026)	
Observations	738	86,221
R^2	0.209	0.007
Adjusted R^2	0.201	0.004
Residual Std. Error	0.035 (df = 730)	
F Statistic	27.500^{***} (df = 7; 730)	87.800^{***} (df = 7; 85983)
Note:		*p<0.1; **p<0.05; ***p<0.01

returns) and within assets with low volume and low number of trades. Given this, one question arises: is this good for the odd lot investor? Financial rationale suggests that, since prices are higher and other factors except liquidity are the same, odd lot assets present a lower discount rate. This seems implausible given our results. For this reason, we consider the mispricing harmful to odd lot investors. The positive relationship between the spread proxy and the level of mispricing suggests that when liquidity is low in the round lot market (high spreads), a higher mispricing occurs in the odd lot market.

The next analysis investigates the financial impact of this mispricing. Table 5 shows the cumulative return, mean, maximum, minimum, standard deviation measures and the Sharpe ratios for our trading strategy detailed in section 2. We compare these results against a naive buy and hold strategy using the Ibovespa index as a proxy for the Brazilian market. For each day, both our benchmark as well as the returns of our strategy are deducted from the risk-free rate[11]. During three of four rolling windows, the Ibovespa index yields negative cumulative returns close to -12%; maximum and minimum values range from approximately 6% to -5%; and the standard deviation remains constant at 1.6%. Using a rolling window of 500 days, the Ibovespa has yielded a cumulative return of 10%. Using the same rolling windows, our strategy yields net cumulative returns around 98% to 153% when trading costs are not considered. The standard deviation of our strategy (between 1.3% and 1.4%) is lower than the Ibovespa's (1.6%). Our strategy also exhibits a positive Sharpe ratio.

Table 5 – Results of trading strategy based on odd lots mispricing

The table presents results for a trading strategy using the mispricing of odd lot and round lot trades. A rolling window scheme is defined where w refers to the time window in days $w = \{500, 250, 125, 62\}$. For a given time window, the average mispricing for all assets is calculated and divided by quintiles. The top-quintile assets are selected, ie., assets with the highest difference between odd lot and round lot trades within our intradaily sample. Next, we sell the odd lot stocks with the highest mispricing and buy the assets in the round lot using daily closing prices at $w + 1$. The result is subtracted by a risk-free daily rate. Trading costs are defined as 0.1% for each position (long and short). As a benchmark, a buy and hold strategy for the Ibovespa index net of the risk-free rate is reported.

Benchmark	Cumsum	Mean	Max	Min	Sd	Sharpe ratio
Window = 500	0.103	0.0003	0.063	-0.050	0.016	0.018
Window = 250	-0.125	-0.0002	0.063	-0.050	0.016	-0.013
Window = 125	-0.117	-0.0002	0.063	-0.050	0.016	-0.010
Window = 62	-0.121	-0.0002	0.063	-0.050	0.016	-0.010
No transaction costs	Cumsum	Mean	Max	Min	Sd	Sharpe ratio
Window = 500	1.449	0.004	0.095	-0.034	0.014	0.289
Window = 250	0.987	0.002	0.073	-0.053	0.013	0.122
Window = 125	1.529	0.002	0.095	-0.062	0.014	0.151
Window = 62	1.534	0.002	0.095	-0.069	0.014	0.137
Transaction costs: 0.1%	Cumsum	Mean	Max	Min	Sd	Sharpe ratio
Window = 500	0.742	0.002	0.092	-0.036	0.014	0.148
Window = 250	-0.218	-0.0004	0.071	-0.055	0.013	-0.027
Window = 125	0.072	0.0001	0.092	-0.064	0.014	0.007
Window = 62	-0.048	-0.0001	0.092	-0.071	0.014	-0.004

As one can see, the trading strategy neglects frictions present in financial markets such as transaction and liquidity costs. However, our main point is not to prove the profitability of such strategy, but to expose the impact of this anomaly created by allowing two distinct platforms trading the same financial contract. The strategy's profit may be seen as the price

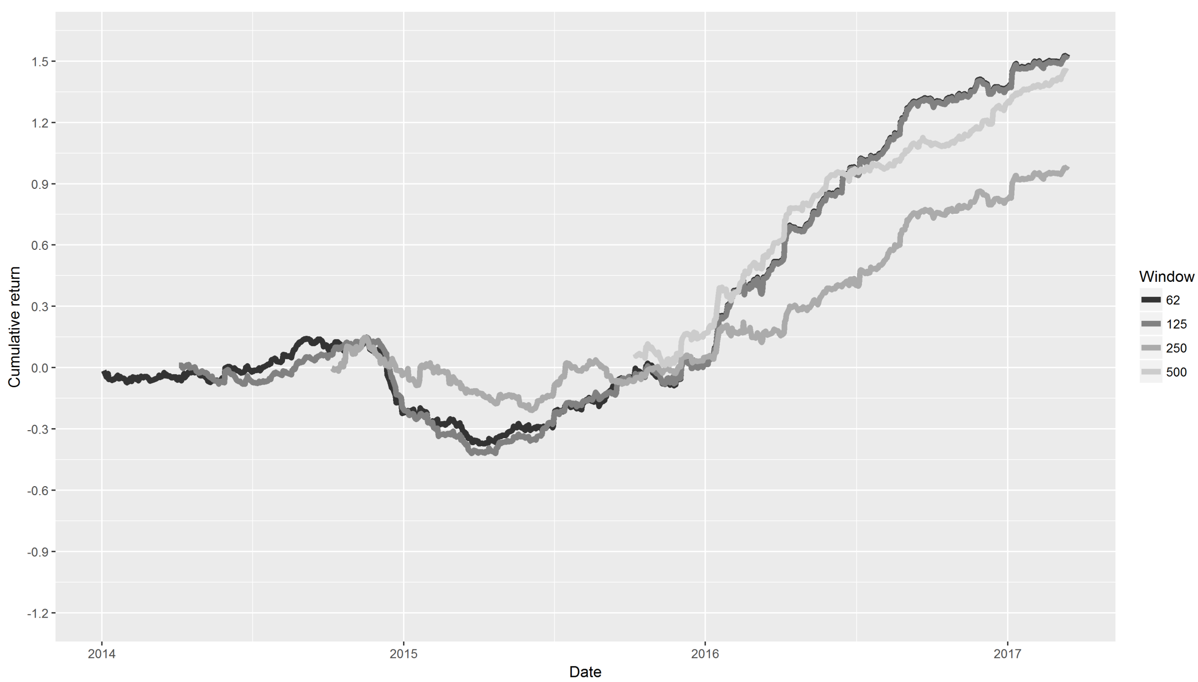
[11] We use the 30-day DI Swap as risk-free rate, which is the standard in the Brazilian market.

agents (probably retail investors) pay more to enter in a market with low liquidity (compared to the round lot market).

However, one could be interested in testing the real profitability of this strategy. In current days, HFTs may quickly identify situations where the detachment between odd/round lots may generate profits. As this strategy is short-lived (the investor would hold inventory for less than a trading day), is based on a low amount of cash and it hinges on quickly detect market conditions appropriated to profit, such strategy is suitable for HFTs. We present results when transaction costs of 0.1% are accounted for each overall position (long and short for all assets in the respective positions). This percentage also accounts for possible liquidity costs, mainly when the odd lots order book has not the depth necessary to sell at the closing prices of the day. At the end of Table 5, it is possible to note that transaction costs are prohibitive for most of the time windows used. On the other hand, using a rolling window of 500 days, the strategy yields significant cumulative returns (72%). Other values for w such as 125 days yields smaller values (7.2%), and the returns using rolling windows of 250 days (-21.8%) and 62 days (-4.8%) are overwhelmed by transaction costs. We believe that a longer time window may avoid stocks being incorrectly classified as high-mispriced assets due to transitory shocks, hence yielding higher returns in the rolling window of 500 days (72%). It is noteworthy that the mispricing exists within shorter rolling windows (as exposed by the returns without trading costs), but are not high enough to overwhelm the trading costs. Therefore, longer time horizons seem to better classify the most mispriced stocks.

Figure 3 – Strategy - No trading costs

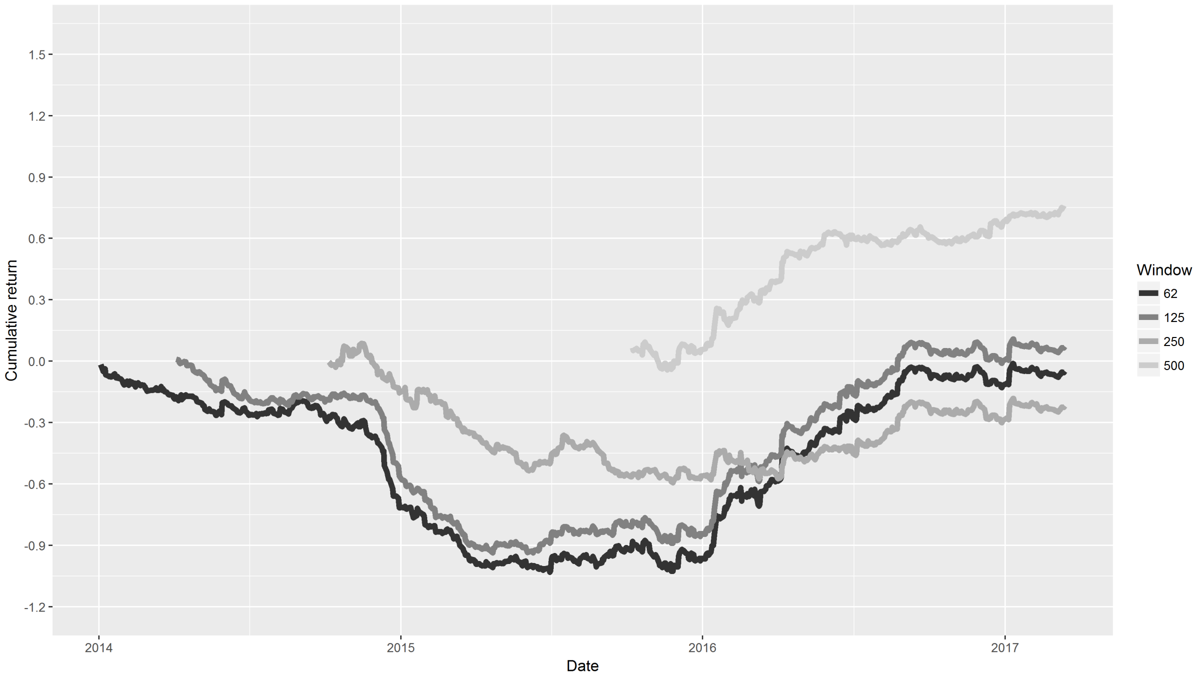
The figure depicts the results for the trading strategy that sells highly mispriced odd lot stocks and buys the respective round lot stocks. Rolling schemes are defined for time window $w = \{500, 250, 125, 62\}$. Results are deduced from the risk-free rate. No transaction costs are considered.



Figures 3 and 4 show the cumulative returns for the long-short strategy when not

Figure 4 – Strategy - Trading costs: 0.1%

The figure depicts the results for the trading strategy that sells highly mispriced odd lot stocks and buys the respective round lot stocks. Rolling schemes are defined for time window $w = \{500, 250, 125, 62\}$. Results are deduced from the risk-free rate and from trading costs of 0.1% per position (long and short).



considering transaction costs and the results accounting for transaction costs, respectively. Regardless of the rolling window chosen, Figure 3 shows that, for our sample period, the difference between the highest mispriced assets consistently yields positive returns from the beginning of 2016 until the end of our sample. Although the slope of the lines in Figure 4 are also ascending in this period, the transaction costs reduce the benefit of this strategy.

O'Hara et al. (2014) finds that most of the odd lot trades in NASDAQ come from HFTs, which can slice and dice larger orders into odd lot orders to ping for hidden liquidity, which is also documented in Davis et al. (2017). Although the use of high-frequency/algorithm trading is not well documented in Brazil, we expect this type of trading to be increasing over time. In 2010, the equities segment of BM&FBovespa started to rent *colocation* spaces within the exchange data center. The reduced latency time provided by *colocation* may attract the use of HFTs. In spite of not being the focus of this study, we believe that HFTs may take advantage of transitory mispricing during trading sessions. Literature reports that HFTs try to profit within small-capital and short-time operations (JONES, 2013; MENKVELD, 2016), which apparently fits the odd lot market in Brazil. Once again, we believe that retail investors will have to pay the bill for this. Our concluding section addresses our suggestions on how this cost may be reduced.

4 Discussion and Concluding Remarks

Our empirical evidence suggests that odd lotters pay higher prices for the same asset traded in round lots. This mispricing is higher when volatility and spreads are high, and market returns are negative. We consider this mispricing as unhealthy for the market. In short, it is not reasonable to offer two platforms for negotiating the exact same rights (i.e., common or preferred stocks) with distinct prices. As odd lot trades creates this anomaly to the market, a simple suggestion to solve the issue would be to end the round lot constraint for the Brazilian stock market, hence liquidating the odd lots market and merging it to the round lot one. This would inflict a high cost to shut down the odd lot platform and adjust the round lot one for allowing smaller trades.

The literature brings up evidence that a decrease in the Minimal Trading Unit (MTU) is beneficial to stock markets. Hauser e Lauterbach (2003) show that after the decrease in the MTU in 1999, stocks listed in the Tel Aviv Stock Exchange have increased price accuracy, volume traded, and overall prices. The evidence is similar in the Tokyo Stock Exchange, which also presented a lower PIN (probability of informed trading) after the reduction in MTU (AMIHUD et al., 1999; ISAKA, 2014; AHN et al., 2014). In Borsa Italiana, apart from higher prices and higher liquidity, Gozluklu et al. (2015) shows diminished bid-ask spreads after the reduction.

According to Amihud et al. (1999) and Hauser e Lauterbach (2003), the cited effects of MTU reductions are channeled via the raise of number of stockholders, since MTU constrained agents who could not trade the stock may enter in the market. Based on the model of Merton (1987), a higher number of shareholders should reduce problems of available information regarding the stock (HAUSER; LAUTERBACH, 2003). This increase in information reduces the cost of capital and, consequently, increases stock price. Since more investors may negotiate the asset, traded volume is expected to increase within investor base. An increase in volatility is expected as a drawback, since higher traded volume should imply in higher volatility (HAUSER; LAUTERBACH, 2003). However, as previously exposed, the benefits of MTU reduction seems to outperform its drawbacks.

The potential increase in traded volume may not be enough to affect thin traded stocks. The work of Hauser e Lauterbach (2003) documented thin traded stocks presenting lower prices and higher volatility after the reduction in the Tel Aviv Stock Exchange. This result may be just a stressing of the illiquidity premium. We believe that an unconstrained (or less constrained) market, with a lower MTU, may accentuate the effects of this premium.

Therefore, a suggestion for regulators is to rethink the role of the odd lot market in Brazil. We believe the MTU in round lots could be reduced from the 100-shares threshold or even consider the merge both markets. As this would be a major change in the market functioning, we suggest the MTU to be reduced to 50 shares. This would encourage constrained investors to enter in the market without bringing technical problems of matching odd lot orders. Therefore, the odd lot market would still be available for more constrained investors and for those willing to buy or sell stocks after splits. At first sight, one could foresee higher price volatility and small orders to be bringing technical problems for order execution. However, empirical evidence suggests the gains to overcome possible drawbacks. As exposed in Table 3, companies with lower traded volume and lower number of negotiations are the ones with higher price detachment regarding round lot trades. We suppose these assets to be the most benefited in a hypothetical MTU reduction.

Rethinking the odd lots market in Brazil seems to be necessary in order to reduce harm to small investors. Without considering trading costs, a simple trading strategy that sells overpriced odd lot assets, and buys the equivalent round lot presents consistent returns over our sample period. This profit is, in fact, a cost paid by the odd lotter. The reason for an odd lot market may resemble the times of oral auctions (open outcry system), when the local arrangement of brokers should be organized in a manner that brokers trading larger lots should be close to each other. In addition, before the advent of electronic trading sessions, the order matching for odd lots was possibly troublesome. However, none of these concerns are present in electronic trading exchanges.

Thus, we believe that a wide field of research is open both in emerging and in developed economies for better understanding of odd lot trades. For instance, markets such as in the Philippines and Singapore present a market design similar to the Brazilian market. Odd lot trades in the Taiwan Stock Exchange are only allowed between 1:30pm to 2:30pm. NYSE merged odd lot and round lot platforms in July 2010[12]. Certainly, a variety of countries differ in odd lot trading, leading to questions regarding these markets. Are odd lotters around the world uninformed investors or HFTs? Are HFTs increasing liquidity in already thin markets such as odd lots? Do odd lots contribute do price discovery in integrated markets? A better understanding of the distinct odd lot markets may answer some of these questions, enhance trading systems and, consequently, increase the benefits of financial markets to society.

[12] For details, see the SEC release No. 34-62302, File No. SR-NYSE-2010-43

Part II

Liquidity, implied volatility and tail risk: a comparison of liquidity measures

Abstract

Liquidity is easily perceived but not easily measured in financial markets. Researchers and practitioners develop and test new measures of liquidity which may be good candidates for measuring this elusive concept. In this study, we present a comparison of variables within two empirical exercises using up to eight traditional liquidity proxies and two proposed proxies based on semi-deviations in a sample of NYSE-listed stocks. The first empirical exercise analyzes the relationship between liquidity and implied volatility, showing that increases in implied volatility increases illiquidity. Using a decomposition of the squared VIX, we show that both conditional variance and variance premium components affects liquidity. Our second empirical exercise investigates the relationship between common factors in liquidity and tail risk. Common factors increase individual stocks' Value-at-Risk and Expected Shortfall, although the effect is not significant for market risk. In both applications, most of the studied proxies presents results aligned with the body of literature.

Keywords: Liquidity measures, Common factors, Tail risk, VIX, Semi-deviations

Note: this article has been accepted for publication in the International Review of Financial Analysis (2018 JCR Impact Factor: 1.693, Qualis A1). Previous versions of this paper were presented at the 2018 and 2019 Brazilian Finance Meetings and were submitted at the Journal of Empirical Finance (final decision: July 2018), Management Science (final decision: November 2018) and the International Journal of Finance and Economics (final decision: October 2019).

1 Introduction

Liquidity is a fundamental variable because the entire notion of a financial market is that it is a place (physical or virtual) where agents can easily trade financial assets. Negotiating illiquid assets can increase both explicit and implicit transaction costs, therefore reducing profits. Distinct literature streams exist regarding liquidity in finance, from asset pricing to portfolio optimization and many others. However, one of the main difficulties of these studies is measuring liquidity since it is a multidimensional concept that agglutinates distinct aspects (CHORDIA et al., 2000). In order to address these issues, we provide an encompassing approach to compare and relate different liquidity proxies for the US market. To achieve this goal, we study the relationship between liquidity and implied volatility and common factors in liquidity and tail risk. As a byproduct, we introduce a liquidity proxy based on the semi-deviations of stocks' returns.

Finance literature defines three main dimensions of liquidity. Namely: transaction costs, depth and resiliency (HASBROUCK, 2007; FOUCAULT et al., 2013). Transaction costs are associated to spread measures, which gauge both explicit and implicit costs to trade. Depth is associated with the impact of larger orders on prices. If a market is deep, one can negotiate larger quantities in distinct price ranges. Both dimensions are widely analyzed in literature and have been associated to asymmetric information (BROGAARD et al., 2014), the liquidity premium (AMIHUD, 2002), volatility (CHORDIA et al., 2005), high-frequency trading (CARTEA et al., 2019; HENDERSHOTT et al., 2011; HENDERSHOTT; RIORDAN, 2013) and many other topics in finance[13]. The third dimension of liquidity is resiliency, which is associated to a dynamic component: in a resilient market, liquidity shocks assigned to microstructure effects tend to quickly dissipate while changes in fundamentals should affect prices permanently. Research has shown the effects of resiliency on stocks (KEMPF et al., 2015) and bank bonds (BLACK et al., 2016). However, measuring liquidity for each of these dimensions poses as a challenge for both practitioners and researchers.

Many recent work in finance face the obstacle of liquidity measuring mostly by two ways: a) showing distinct aspects of liquidity regarding different specifications in asset-pricing related studies (ANTHONISZ; PUTNIŅŠ, 2016; AMIHUD et al., 2015; BELKHIR et al., 2018; BEN-REPHAEL et al., 2015; ABDI; RANALDO, 2017; GRILLINI et al., 2019; WU, 2019); and b) proposing new measures of liquidity and providing empirical applications using real data (CHACKO et al., 2016; FONG et al., 2018; NIETO, 2018). As new measures of liquidity are developed, new horseraces and comparisons between proxies are presented by researchers[14]. Therefore, our study addresses both strands of the literature: we propose new measures of liquidity based on the semi-deviation of stocks' returns and study whether new and old liquidity measures are related to other strands of financial literature. Our two empirical studies are: a)

[13] Gabrielsen et al. (2011), Amihud et al. (2012) and Holden et al. (2014) provide a comprehensive explanation of liquidity and its properties.

[14] Goyenko et al. (2009), Fong et al. (2017) and Schestag et al. (2016) report a set of comparisons between measures, although there are many more studies addressing the issue.

connection with implied volatility and b) common factors in liquidity and tail risk.

Both empirical applications are employed within NYSE stocks sampled from 1990 to 2016. Our goal is to compare whether usual liquidity proxies present similar effects when compared through empirical applications. Thus, we do not aim to specifically rank performance between measures. The liquidity proxies employed in this study are: 1) Dollar volume, 2) Turnover, 3) the Corwin e Schultz (2012) spread estimator, 4) the Roll (1984) spread estimator, 5) Range (high price minus low price), 6) Range-volume ratio, 7) the Amihud (2002) measure, 8) the Amihud-turnover measure, 9) the Amihud (2002) measure adapted to semi-deviations and 10) the Amihud-turnover measure adapted to semi-deviations. We aim to contribute to the financial theory and practice in three ways: 1) by comparing the effects of both new and old liquidity proxies in two new empirical exercises, 2) by proposing new measures of liquidity which possess empirical tractability and desired mathematical properties. The new measures respects the axioms of a Generalized Deviation Measure as proposed by Rockafellar et al. (2006) as well as the volatility over volume liquidity proxies of Fong et al. (2018). The third contribution is to provide new evidence of the relationship between liquidity and other issues in finance such as implied volatility and tail risk. Although the paper at hand uses data from US, it can be replicated/extended using both international and local data.

Regarding our first empirical application, liquidity and volatility are well documented to be related in financial markets. Relevant work such as from Stoll (1978), Stoll (2000) show that spreads are positively related to stock volatility. In a prevailing quote-driven market, stock volatility is expected to raise bid-ask spreads, since dealers will be expected to earn a compensation for bearing the risk of a volatile asset or to trade against a better-informed investor. Hence, in periods of low uncertainty regarding prices, a risk-averse dealer should charge lower spreads (AMIHUD, 2002; STOLL, 1978). Brunnermeier e Pedersen (2009) develop a model in which asset's liquidity is directly connected to market maker's funding liquidity. When a market maker buys an asset, he can use it as a collateral and borrow a part of the respective asset value to trade again based on a margin offered by the lender. During crises, the level of funding is diminished and margins are higher, so the level of liquidity is reduced. This implies in what the authors call "liquidity spirals", in which lower levels of funding diminish trading activity, increase margins and diminish both liquidity and returns, creating a vicious cycle[15]. Since margins required are higher for more volatile assets, when funding is low, market makers prefer to trade less volatile assets. This takes us back to the liquidity spirals, since more volatile assets will have their liquidity dried up.

However, most of the classical studies use proxies for volatility based on physical volatility. As improvements on options pricing theory have been made, in current days it is possible to retrieve an estimation of the future volatility of a given asset based on options trading; namely implied volatility. Regarding liquidity, the results of Chung e Chuwonganant (2014) show that the CBOE Volatility index (VIX) is positively related to measures of spread, while Chung e Chuwonganant (2018) evidences the effects of market volatility on stock returns through its impact on liquidity provision. Nagel (2012) mimics liquidity providers returns using a reversal

[15] For an intuitive vision, see Brunnermeier e Pedersen (2009) - Figure 2.

strategy and also finds that the VIX predicts the return of market makers. When VIX is high, market makers demand higher premiums to provide liquidity in market turmoil. Both listed studies rely on the model of Brunnermeier e Pedersen (2009) to explain these findings.

Given this panorama, we analyze whether the liquidity measures employed in this paper are related to implied volatility. In order to achieve this, we use the VIX index as a proxy, since it is widely used as the expectation of market volatility. Literature shows that VIX is related to spreads in sovereign credit default swaps (LONGSTAFF et al., 2011; PAN; SINGLETON, 2008), illiquidity of corporate bonds (BAO et al., 2011) and expected cross-sectional returns (ANG et al., 2006). We estimate the effect of VIX over several illiquidity proxies and find a positive effect of the former in the latter. In order to detail the relationship between volatility and illiquidity, we present estimations switching the VIX measure for the conditional variance and the variance premium of the decomposed squared VIX as in Bekaert e Hoerova (2014). Using both time series and panel estimations, our results for this empirical exercise show that traditional liquidity measures are related to implied volatility, in which increasing implied volatility is related to an increase in illiquidity. The semi-deviation-based proxies also present this behavior. Additionally, we provide new insights on the dynamic relationship between liquidity and implied volatility using a vector autoregression (VAR) approach.

Our second application is related to common factors in liquidity and tail risk. Although each given asset traded on financial markets has its own liquidity (which may vary depending on several factors), investors may notice that the level of liquidity during some periods may be more or less correlated between assets. The co-movements in liquidity are defined as commonality in liquidity (hereafter CIL). Chordia et al. (2000) were the first at documenting liquidity (measured by spreads and depth) as presenting commonalities across stocks. Intuitively, the idea of CIL is that the liquidity of a certain asset is sensible to market-wide liquidity. Literature has been presenting equilibrium models such as from Acharya e Pedersen (2005) and Anthonisz e Putniņš (2016), as well as evidences of both demand-side and supply-side explanations on the determinants of CIL. Demand-side theories allege that correlated trading activity of investors may create a common factor on liquidity (CHORDIA et al., 2000). Also, institutional investors and mutual funds engaging in similar strategies may present correlated trading activity, thereby inducing CIL (KAMARA et al., 2008; KOCH et al., 2016). Another source of commonality may arise from correlated trading due to incentives for diversification and investor sentiment (KAROLYI et al., 2012).

Supply-side explanations have a strong appeal from theoretical models. Intuitively, if funders (banks and financial institutions) suffer from shocks in their balance sheets they are likely to transfer such shocks to their customers (hedge funds, dealers and financial intermediaries in general) in the form of higher margins to suit for banks' risk management. If borrowers do not dispose of such capital, they will be forced to liquidate positions in order to meet margin calls, thereby inducing liquidity commonality. Brunnermeier e Pedersen (2009) define that such margins will be set by positions' Value-at-Risk (VaR), therefore implying an indirect link between tail risk and CIL[16]. If negative shocks increase tail risk, margin requirements should narrow,

[16] Brunnermeier e Pedersen (2009) provide the benchmark model for this mechanism, although many others

implying a common sell effect on prices. Such commonality should increase losses, therefore increasing risk and creating a vicious cycle. Given the impact on financial institutions' balance sheets, this effect is also magnified within higher volatility, increases on interest rates and poor market/economic conditions (SYAMALA et al., 2017).

Despite many studies that have been made to evidence distinct aspects of CIL, relative little attention has been given to the relationship between common factors in liquidity and tail risk, despite advances in each separated field. Ruenzi et al. (2016) model left tail dependence coefficients between liquidity and returns distributions. The authors find that stocks with high extreme downside liquidity risk (EDL) earn a significant return premium. Rubia e Sanchis-Marco (2013) use liquidity-related variables (trading activity and spreads) to forecast Value-at-Risk of portfolios formed by NYSE stocks in a dynamic quantile regression approach. Following these studies, we aim to illustrate how common factors in liquidity and tail risk relate each other in the US market. Specifically, we believe that the indirect hypothesis of tail risk affecting common factors (and vice-versa) of Brunnermeier e Pedersen (2009) has not been fully tested with empirical data. In the aforementioned model, margins provided by lenders are sensitive to the position's Value-at-Risk. If risk rises through a liquidity shock, so should margins, forcing financial intermediaries to costly liquidate positions in order to meet margin requirements. Thus, we provide empirical evidence for the predictions of Brunnermeier e Pedersen (2009) as our estimations show that when common factors rise, individual risk measures are amplified. Our results show that direct measures of stocks' risk (Value-at-risk and Expected Shortfall) are related to common factors of our liquidity measures extracted by principal component analysis (PCA). We show that although common factors are not related to market risk, these common factors increase individual risk of stocks. Therefore, we evidence a direct channel through liquidity increase tail risk. Regarding our proposed proxies, semi-deviation capture the observations below the mean of returns, therefore we believe they are sound proxies for understanding the relation between common factors in liquidity and tail risk. Our results relate to Adrian e Shin (2010), Adrian e Shin (2014) who show how funding liquidity of financial intermediaries are associated with leverage. As their studies focus on how funding liquidity and risk relate on intermediaries' balance sheet, we provide evidence on how liquidity relates to the risk of loss in stocks, an asset commonly used in collateralized borrowings.

The paper proceeds as follows: Section 2 describes the data and the liquidity proxies. Section 3 presents our study regarding liquidity and implied volatility and Section 4 presents our results regarding common factors in liquidity and tail risk. The final section ends this paper with concluding remarks.

provide contributions on this issue (HAMEED et al., 2010; ADRIAN; SHIN, 2010; COMERTON-FORDE et al., 2010; NAGEL, 2012).

2 Data and variables

The sample period comprises the years between 1990 and 2016 for companies listed on the NYSE. This period comprises different market conditions, such as the 2000 dot-com bubble and the 2008 global crisis. We use daily data from CRSP aggregated by monthly averages. As usual with similar studies, we include only CRSP share codes 10, 11 and 12. This accounts only for ordinary common shares and excludes American Trust Components (Primes and Scores), REITs and closed-end funds. Our liquidity proxies are described below. Notation is defined as follows: i refers to each stock and d accounts for each of the D days of month t.

First, we include simple spot measures often used as proxies for liquidity. The mean traded volume and turnover are defined as:

$$Volume_{it} = \frac{1}{D}\sum_{d=1}^{D} VOLD_{idt}, \tag{2.0.1}$$

where $VOLD_{idt}$ is the share traded volume multiplied by the closing price of asset i on day d of month t.

$$Turn_{it} = \frac{\sum_{d=1}^{D} Share.volume_{idt}}{Sh.out_{it}}, \tag{2.0.2}$$

where $Turn_{it}$ refers to the ratio of share turnover calculated by the sum of share volume of stock i at the month t to $Sh.out_{it}$, which refers to the shares outstanding of stock i at the beginning of month t. Volume and turnover are often used as proxies for liquidity, although literature shows that may not be the case. For example, Barinov (2014) provides evidence that turnover is unrelated to several liquidity proxies, therefore accounting for firm-specific uncertainty, not liquidity.

As a bid-ask spread proxy, we include the High-Low estimator developed by Corwin e Schultz (2012). This proxy is constructed based on consecutive two-day high and low values. It is calculated by the following equations:

$$CS_{it} = \frac{1}{D}\sum_{d=1}^{D} HighLow_{idt} \tag{2.0.3}$$

$$HighLow_{idt} = \frac{2(e^{\alpha} - 1)}{1 - e^{\alpha}} \tag{2.0.4}$$

$$\alpha = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}} \tag{2.0.5}$$

$$\beta = \sum_{j=0}^{1} \left(log \left(\frac{H_{id+jt}}{L_{id+jt}} \right) \right)^2 \tag{2.0.6}$$

$$\gamma = \left(log \left(\frac{max\,(H_{idt}, H_{id+1t})}{min\,(L_{idt}, L_{id+1t})} \right) \right), \tag{2.0.7}$$

where H_{idt} (L_{idt}) is the highest (lowest) price traded for asset i on day d of month t. Additionally, we include the spread estimator proposed by Roll (1984):

$$Roll_{it} = 2\sqrt{-Cov(R_{idt}, R_{i-1dt})}, \tag{2.0.8}$$

where R_{idt} is the log-return of stock i at day d of month t. This measure is set equal to zero in the case of $Cov(R_{idt}, R_{id-1t}) > 0$.

Next, we include a range-based measure that is given by the ratio of range on high and low traded prices to the closing price of stock i:

$$Range_{it} = \frac{1}{D} \sum_{d=1}^{D} \frac{H_{idt} - L_{idt}}{P_{idt}}. \tag{2.0.9}$$

H_{idt} and L_{idt} are previously defined and P_{idt} refers to the closing price of stock i on day d of month t. Dividing the range by the closing price allows comparison between stocks. This measure resembles the proxy of Garman e Klass (1980). The simple measure of *Range* may be used also as a proxy for stocks' volatility and it is expected to be correlated to our spread proxies. Additionally, resembling a price impact proxy, we include a version of *Range* divided by the traded volume, *Range.volume*:

$$Range.volume_{it} = \frac{Range_{it}}{Volume_{it}} \tag{2.0.10}$$

Also as a price impact proxy, the Amihud (2002) illiquidity measure is calculated by:

$$Amihud_{it} = \frac{1}{D} \sum_{d=1}^{D} \frac{|R_{idt}|}{VOLD_{idt}}, \tag{2.0.11}$$

where $|R_{idt}|$ is the absolute log-return of asset i on day d of month t. Given its simplicity and practical use, the Amihud (2002) proxy is one of the most used liquidity proxies. We also include this measure switching the traded volume by share turnover as in Brennan et al. (2013). The measure is defined by:

$$Amihud.turn_{it} = \frac{1}{D}\sum_{d=1}^{D}\frac{|R_{idt}|}{Turn_{it}}, \tag{2.0.12}$$

Finally, we propose semi-deviation-based proxies for liquidity. The absolute return in the Amihud's measure is switched by the semi-deviation at month t:

$$Semid_{it} = \frac{\sqrt{\frac{\sum_{d=1}^{D}((R_{idt} - E[R_{it}])^{-})^2}{D}}}{Volume_{it}}, \tag{2.0.13}$$

where the $^{-}$ sign refers to the negative part of R_{idt} - $E[R_{it}]$. In case of a positive value, the deviation is set to zero. Therefore, the proposed measure reflects the effect on a position price given a change in volume. Using semi-deviations in liquidity is somewhat new, although this type of deviation measure is widely studied in financial risk, portfolio optimization and engineering problems (GRECHUK et al., 2009; ROCKAFELLAR; URYASEV, 2013). By construction, this liquidity measure reflects the effect on price given a change in the asset negotiability (in this case, measured by volume). Semi-deviations are part of a wider class of measures which follow mathematical properties (axioms), assuring good theoretical results and empirical tractability. Such class os measures is named Generalized Deviation Measures and were proposed by Rockafellar et al. (2006). As semi-deviation measures present a natural adjustment for negative returns, they share mathematical properties of risk measures studied in the literature. Semi-deviation measures are closely related to risk measures, as it is possible to convert semi-deviation variables into risk measures (ROCKAFELLAR et al., 2006). Appendix A details the axioms of a Generalized Deviation Measure. The semi-deviation measures are also included in the general class of volatility over volume liquidity proxies proposed by Fong et al. (2018).

The denominator of our class of measures can be defined as a variable that reflects negotiability. One can replace $Volume$ for any type of related variable such as number of trades, quantity of shares traded and etc. Additionally, a proxy using semi-deviations was also calculated using the turnover as denominator:

$$Semid.turn_{it} = \frac{\sqrt{\frac{\sum_{d=1}^{D}((R_{idt} - E[R_{it}])^{-})^2}{D}}}{Turn_{it}}. \tag{2.0.14}$$

By proposing new liquidity proxies, our goal is not to replace well-known ones, but rather to provide a distinct perspective on the analysis of liquidity. Semi-deviation measures are higher when negative return dispersion is high (deviation below the mean). As liquidity may be a concerning issue mainly in down markets (HAMEED et al., 2010; ANTHONISZ; PUTNIŅŠ, 2016; BELKHIR et al., 2018), a measure that reflects both concepts (illiquidity and negative returns) may be beneficial to practitioners understand the relationship between liquidity and down markets.

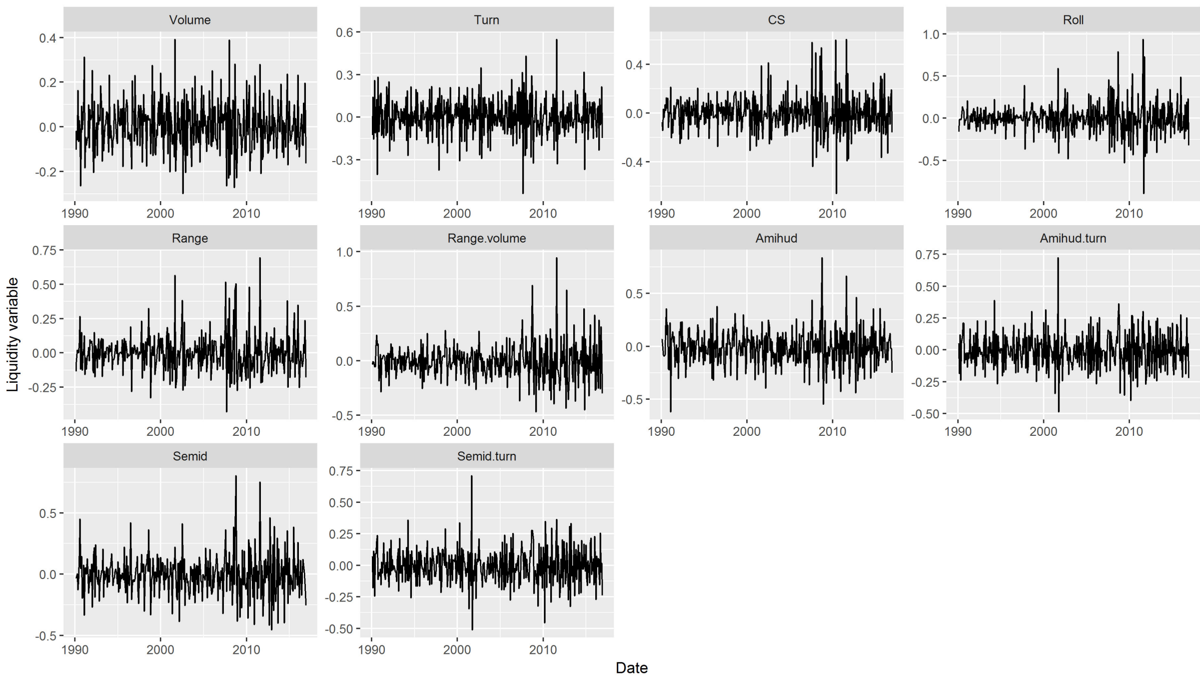
The proposed measures are similar to those used by Fong et al. (2018), who provide evidence that volatility over volume liquidity proxies may be more suitable to assess periods of unusual higher illiquidity when compared to the Amihud (2002) proxy. The aforementioned proxies resemble ours, with the difference on the numerator of both variables: volatility over volume proxies use standard-deviation while we use semi-deviations.

Though many of the listed proxies increase in assets' illiquidity, we denominate the set of variables as 'liquidity measures'. Since the variables are monthly averaged, we apply initial filters similar to Acharya e Pedersen (2005) and Pastor e Stambaugh (2003), who also use the same data frequency. For each month t, a stock i will be admitted to the cross-sectional calculation of liquidity measures if:

- it has at least 15 days with return and positive volume available,
- its beginning-of-month price is between \$5 and \$1.000.

After applying these filters, we calculate the liquidity measures and remove outliers at the 1% level for each month t. There are between 1215 and 2155 assets that fit in the filters applied. In order to avoid problems of non-stationarity, we follow Kamara et al. (2008) and report all estimations using our liquidity variables in log-differences. Figure 5 exhibits the equally-weighted monthly averages for the liquidity variables defined for the stocks in our sample period. Spread variables such as *Roll* and *CS* show clustering over some time periods. At the end of our sample, one can see a higher variability on measures such as *Range.volume* and *Semid*.

Figure 5 – Liquidity measures



The figure shows the monthly-averaged liquidity measures for all stocks over the sample period (1990-2016). All measures are presented in log-differences.

Table 6 presents descriptive statistics for the liquidity measures in log-differences. Price impact measures such as *Amihud*, *Amihud.turn*, *Semid* and *Semid.turn* present negative means, which is expected since trading activity is increasing over time, therefore decreasing the measures of illiquidity. Regarding spread measures, *Roll* (mean: -0.0008) and CS (mean: 0.0011) diverge in sign when their mean values are compared. Table 7 presents Pearson's correlation between the measures. Several pairs present correlations above 0.5 (in boldface). This is expected since many variables are somehow dependent on $Volume$. Still, one can note that $Volume$ is highly correlated to spread measures such as *Roll* (0.588) and CS (0.609). As the construction of *Amihud* and *Semid* is similar, they present high correlation (0.909), which also occurs within *Amihud.turn* and *Semid.turn* (correlation: 0.970). Still, *Semid* and *Semid.turn* are more robust from the mathematical point of view. The *Semid* proxy is also correlated with other proxies such as *Range.volume* (0.723) and *Semid.turn* (0.648).

Table 6 – Summary statistics

The table presents summary statistics for the variables employed in the estimations described in Section 2 using NYSE stocks. In order for a stock be included in the sample, its beginning-of-month price should be between \$5 and \$1,000 and be traded at least 15 days in a given month. Outliers are removed at the 1% level. The sample period spans from 1990 to 2016.

	Mean	Median	Min	Max	Sd
Volume	0.0064	0.005	-0.2994	0.392	0.1146
Turn	0.0042	0.0005	-0.539	0.5487	0.1402
CS	0.0011	-0.0012	-0.6605	0.6064	0.1595
Roll	-0.0008	-0.0022	-0.8932	0.9347	0.1922
Range	0.0002	-0.003	-0.4283	0.6945	0.1425
Range.volume	-0.0045	-0.0034	-0.4673	0.9476	0.1683
Amihud	-0.0084	-0.0078	-0.6183	0.8396	0.1751
Amihud.turn	-0.0051	-0.0075	-0.4856	0.7256	0.1466
Semid	-0.0068	-0.0071	-0.4555	0.8039	0.1695
Semid.turn	-0.005	-0.0112	-0.5118	0.7107	0.1456

As the correlation results show, one can think of replacing well-known proxies such as the Amihud (2002) by proxies such as *Semid*. As the calculation of simple linear correlations do not allow one to ascertain whether these proxies may be used interchangeably, we proceed to our empirical approach in order to prove this. Next, we test two distinct issues of interest of financial literature: a) connection between liquidity and implied volatility and b) common factors in liquidity and tail risk. In the following sections, we describe our methodological approach for each issue and present the results. The objective is to evidence of how (and if) the liquidity variables capture similar effects to proxies that are already used in practice, relating the results to previous findings in financial research.

Table 7 – Time series correlations

The table presents time series correlations between liquidity variables using NYSE stock data. Correlations higher than | ±0.50 | are reported in boldface. In order for a stock to be included in the sample, its beginning-of-month price should be between \$5 and \$1,000 and be traded at least 15 days in a given month. Outliers are removed at the 1% level for each month. The sample period spans from 1990 to 2016.

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Volume	1									
Turn	**0.635**	1								
CS	**0.609**	0.436	1							
Roll	**0.588**	0.422	**0.726**	1						
Range	**0.694**	**0.601**	**0.841**	**0.782**	1					
Range.volume	0.065	0.192	0.392	0.366	0.472	1				
Amihud	-0.01	0.211	0.33	0.356	0.468	**0.785**	1			
Amihud.turn	0.202	-0.258	0.45	**0.556**	**0.536**	0.441	**0.51**	1		
Semid	0.131	0.285	0.48	**0.53**	**0.619**	**0.732**	**0.909**	**0.609**	1	
Semid.turn	0.191	-0.219	0.435	**0.591**	**0.53**	0.424	**0.518**	**0.97**	**0.648**	1

3 Liquidity and implied volatility

Our estimation approach aims to compare whether our liquidity measures are sensible to implied volatility. In order to test this, we gather VIX data from the CBOE website. In our estimations, we use the log-differenced versions of the monthly aggregated VIX, which we refer to as $DVIX$. Next, we proceed for panel data estimations and time series regressions. For the panel data procedure, the following individual fixed-effects regression is estimated:

$$IL_{it}^{j} = \beta_0 + \beta_1 DVIX_t + \zeta Controls_t + \epsilon_t, \tag{3.0.1}$$

where IL_{it}^{j} refers to the liquidity measure j of stock i at month t, $DVIX_t$ to the log-difference of average daily VIX at month t and ϵ_t to the error term. Additionally, we include controls for the natural logarithm of the average market capitalization of firm i at month t and for the standard deviation of daily closing price returns of month t. Year dummies are also included and heteroskedastic and autocorrelation consistent covariance matrices are calculated.

Although VIX is commonly used as a proxy for volatility, recent studies have shown that the VIX may embed a conditional variance component and a variance premium component (BOLLERSLEV et al., 2009; BEKAERT; ENGSTROM, 2017). Additionally, we consider the decomposition of VIX as $VIX^2 = VP + CV$, where VP is the variance premium and CV refers to the conditional variance estimated. CV is referred to a proxy for uncertainty and VP to a proxy for risk aversion. VP may also be thought of as the payoff of the floating leg of a variance swap contract, where an agent accepts to pay a higher price to be protected from an increase in volatility (BEKAERT; HOEROVA, 2014). Bekaert e Engstrom (2017) also define the variance premium as "the difference between the risk neutral expected conditional variance and the actual expected variance under the physical probability measure". Although the effect of VIX in financial markets has been studied in recent years, the dynamics of its components (CV and VP) are far less studied. Therefore, we also estimate equation 3.0.1 switching $DVIX$ by the log-difference of the conditional variance (DCV) and the log-difference of the variance premium (DVP). CV is estimated through the model 8 of Bekaert e Hoerova (2014)[17].

$$IL_{it}^{j} = \beta_0 + \beta_1 DCV_t + \beta_2 DVP_t + \zeta Controls_t + \epsilon_t. \tag{3.0.2}$$

Following Chacko et al. (2016), we also estimate a time series regression of mean liquidity over $DVIX$, as shown in Equation 3.0.3:

$$AIL_{t}^{j} = \alpha_0 + \alpha_1 DVIX_t + \varepsilon_t, \tag{3.0.3}$$

[17] We are grateful to Marie Hoerova who kindly provided the time series for the conditional variance and the variance premium.

where AIL_t^j is the equally-weighted average of liquidity measure j for all stocks available at month t. Again, we test the procedure switching $DVIX$ by DCV and DVP:

$$AIL_t^j = \alpha_0 + \alpha_1 DCV_t + \alpha_2 DVP_t + \varepsilon_t \tag{3.0.4}$$

The procedure of analyzing both cross-sectional and time series relation between variables is widely used in this literature (AMIHUD, 2002; GOYENKO et al., 2009; ANTHONISZ; PUTNIŅŠ, 2016). In order to provide stronger evidence in the time series approach, we follow Chordia et al. (2005) and Karolyi et al. (2012) and conduct vector autoregressions (VAR) and Granger Causality tests. The baseline VAR model is defined by equations 3.0.5 and 3.0.6:

$$AIL_t^j = \beta_0 + \sum_{p=1}^{P} \beta_p AIL_{t-p}^j + \sum_{p=1}^{P} \alpha_p DVIX_{t-p} + e_t \tag{3.0.5}$$

$$DVIX_t = \alpha_0 + \sum_{p=1}^{P} \alpha_p DVIX_{t-p} + \sum_{p=1}^{P} \beta_p AIL_{t-p}^j + v_t, \tag{3.0.6}$$

where β_p refers to the liquidity coefficients, α_p to $DVIX$ coefficients, e_t and v_t are the error terms. The maximum lag P is defined by the lowest lag pointed between the Akaike and the Hannan-Quinn information criterion. Additionally, we estimate a VAR system switching $DVIX$ by DCV and DVP:

$$AIL_t^j = \beta_0 + \sum_{p=1}^{P} \beta_p AIL_{t-p}^j + \sum_{p=1}^{P} \gamma_p DCV_{t-p} + \sum_{p=1}^{P} \delta_p DVP_{t-p} + e_t \tag{3.0.7}$$

$$DCV_t = \delta_0 + \sum_{p=1}^{P} \gamma_p DCV_{t-p} + \sum_{p=1}^{P} \delta_p DVP_{t-p} + \sum_{p=1}^{P} \beta_p AIL_{t-p}^j + v_t \tag{3.0.8}$$

$$DVP_t = \gamma_0 + \sum_{p=1}^{P} \delta_p DVP_{t-p} + \sum_{p=1}^{P} \gamma_p DCV_{t-p} + \sum_{p=1}^{P} \beta_p AIL_{t-p}^j + u_t, \tag{3.0.9}$$

where, γ_p and δ_p are the coefficients for DCV and DVP, u_t is the error term for the third equation of the system and the rest remain similar to previous notation. A VAR approach is adequate since there is evidence that not only contemporaneous shocks in volatility affects illiquidity (CHUNG; CHUWONGANANT, 2014). Thus, the effect of volatility on liquidity is well documented, although the opposite effect is far less reported in literature (CHORDIA et al., 2005). One hypothesis in the model of Brunnermeier e Pedersen (2009) is that liquidity and volatility present a bidirectional effect. We test the predictions of this model.

3.1 Results

Table 8 reports the results for the fixed-effects estimations of equation 3.0.1. As expected, $DVIX$ presents positive and significant coefficients on all liquidity variables. As implied volatility increases, a decrease in liquidity is likely to occur (BRUNNERMEIER; PEDERSEN, 2009). This happens in terms of spreads (CS and *Roll*), price impact (*Range.volume*, *Amihud*, *Amihud.turn*, *Semid* and *Semid.turn*) and price volatility (*Range*). Direct negotiability variables such as $Volume$ and $Turn$ are positively affected by $DVIX$ (coefficients 0.032 and 0.179, respectively), which may be associated with investors switching positions when volatility is high. The effect of implied volatility over spread measures such as *Roll* and CS is positive and straightforward, since larger oscillations in prices, reflected by $DVIX$, should widen spreads (CHORDIA et al., 2005). The log-difference of *Amihud*, *Amihud.turn*, *Range*, *Range.volume*, *Semid* and *Semid.turn* are also positively affected by $DVIX$. The coefficient on the *Amihud* measure is the highest among all liquidity variables (0.648). Notably, the coefficients on *Semid* and *Semid.turn* present the same sign of resembling variables such as *Amihud* and *Amihud.turn*. All coefficients are significant on $DVIX$ at the 1% level.

Table 8 – Fixed-effects estimations - Liquidity and $DVIX$

The table reports results for the fixed-effects estimations of equation 3.0.1. We regress our liquidity measures presented in Section 2 against the log-differenced monthly average of VIX. Control variables include: $log(Mktcap)$ as the logarithm of market capitalization of stock i at month t and Sd as the standard-deviation of daily returns of stock i at month t. All estimations were made with heteroskedastic and autocorrelation robust standard errors. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016. Significance levels: *** (1%), ** (5%) and * (10%).

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DVIX	0.032***	0.179***	0.409***	0.503***	0.591***	0.534***	0.648***	0.455***	0.514***	0.380***
	(0.007)	(0.007)	(0.011)	(0.018)	(0.006)	(0.005)	(0.007)	(0.005)	(0.009)	(0.006)
log(Mktcap)	0.032***	0.014***	0.006***	0.038***	0.010***	−0.017***	−0.002	0.018***	0.027***	0.042***
	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Sd	4.988***	6.424***	3.591***	15.843***	4.801***	0.932***	5.665***	4.603***	11.209***	9.061***
	(0.229)	(0.359)	(0.224)	(0.601)	(0.295)	(0.120)	(0.261)	(0.097)	(0.512)	(0.168)
Observations	430,710	424,283	405,446	149,071	422,718	422,718	435,223	425,204	431,959	424,668
R^2	0.039	0.077	0.017	0.062	0.210	0.072	0.092	0.073	0.134	0.102
Adjusted R^2	0.031	0.068	0.008	0.038	0.202	0.064	0.085	0.065	0.126	0.094
F Statistic	438.156***	871.334***	172.366***	240.855***	2,776.098***	811.510***	1,098.712***	830.594***	1,653.883***	1,195.898***
	(df = 40; 426957)	(df = 40; 420531)	(df = 40; 401697)	(df = 40; 145341)	(df = 40; 418974)	(df = 40; 418974)	(df = 40; 431469)	(df = 40; 421452)	(df = 40; 428205)	(df = 40; 420916)

Next, we switch $DVIX$ by the conditional variance (DCV) and the variance premium (DVP), both log-differenced. Table 9 reports the results of equation 3.0.2. DCV present positive and significant coefficients at the 1% level for all liquidity measures as dependent variables. The results confirm some findings in financial literature, such as traded volume ($Volume$) being positively related to volatility (ROSSI; MAGISTRIS, 2013), as well as spreads (*Roll* and CS) being widened when conditional variance is increased. These spread measures have presented the highest coefficients on DCV (0.207 and 0.182, respectively). *Amihud*, *Amihud.turn*, *Semid* and *Semid.turn* present the expected results, as a variation on CV increases these illiquidity measures. As these measures are price impact proxies, an increase in CV must reduce depth, thereby increasing the price impact of larger incoming orders. This may also be associated with the results on the spread proxies, as higher price impact should widen spreads. Again, the semi-deviation-based measures present similar results to other usual proxies such as Amihud's.

Far less reported in the literature, the effects of DVP on liquidity measures are significant in many liquidity proxies. Stock's $Volume$ and $Turn$ (negotiability variables) present a decrease when the log-difference of the variance premium is positive (DVP coefficients of -0.046 and -0.033, respectively). As risk aversion increases, we assume that investors withdraw from trading as beliefs regarding future volatility may severely change short-term returns. As $Turn$ is also a volume-based measure, the rationale is similar. The spread proxies are negatively affected by DVP: CS (coefficient -0.041) and *Roll* (coefficient -0.081) are reduced when a change in VP is positive and increased when a VP change is negative. Although this seems counterintuitive, when investors' risk aversion perception is high, agents may place buy and sell orders with lower spreads so these orders are quickly traded. When the risk aversion is diminished, investors may place wide-spreads orders as the immediacy to trade before a high-volatility event is lower. Still, it is also possible that the spread proxies incur in a measurement error, since both are daily approximation of spreads. Corwin e Schultz (2012) state that their liquidity measure presents poor performance in illiquid periods, which may be related to periods when DVP is high. Also, the Roll (1984) estimator yields several cases when the autocovariance of returns is positive, in which the spreads are set to zero. Many of the price impact measures present positive and significant coefficients on the effect of DVP in most of the cases. As these variables are constructed by a negotiability variable at the denominator, as $Volume$ and $Turn$ diminishes, liquidity tend to present a positive variation when DVP is increased. DVP presents similar coefficients on *Range.volume* and *Amihud* (0.034 and 0.035, respectively), showing that price impact should be increased when risk aversion rises. As $Volume$ tends to be lower, larger orders tends to move prices more easily, as reported in the literature (AMIHUD, 2002). *Semid* and *Amihud.turn* are similarly affected by DVP (coefficients 0.012 and 0.016, respectively). *Range* is negatively affected by DVP (coefficient -0.014) and *Semid.turn* has not been reported as significantly affected by DVP.

Our findings regarding the variance premium and liquidity relate to those of Barras e Malkhozov (2016). The authors compare two versions of the variance risk premia measured in the equities and options market. The difference between the two variance premia is not affected by a liquidity risk factor. However, their study uses the liquidity risk factor of Pastor e Stambaugh (2003), which accounts for the sensitivity to market liquidity (liquidity risk). Our approach uses liquidity in levels.

Table 9 – Fixed-effects estimations - Liquidity, DCV and DVP

The table reports results for the fixed-effects estimations of equation 3.0.2. We switch $DVIX$ for the change in conditional variance (DCV) and the change in variance premium (DVP) proxies, as in Bekaert e Hoerova (2014). Control variables include: $log(Mktcap)$ as the logarithm of market capitalization of stock i at month t and Sd as the standard-deviation of daily returns of stock i at month t. All estimations were made with heteroskedastic and autocorrelation robust standard errors. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016. Significance levels: *** (1%), ** (5%) and * (10%).

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DCV	0.053***	0.086***	0.182***	0.207***	0.207***	0.138***	0.156***	0.121***	0.128***	0.098***
	(0.002)	(0.002)	(0.003)	(0.006)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)
DVP	−0.046***	−0.033***	−0.041***	−0.081***	−0.014***	0.034***	0.035***	0.016***	0.012***	0.0004
	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
log(Mktcap)	0.031***	0.014***	0.006***	0.036***	0.010***	−0.016***	−0.001	0.018***	0.027***	0.042***
	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Sd	4.999***	6.375***	3.378***	15.655***	4.528***	0.687***	5.477***	4.451***	11.113***	9.003***
	(0.232)	(0.362)	(0.216)	(0.611)	(0.283)	(0.109)	(0.256)	(0.097)	(0.515)	(0.169)
Observations	430,710	424,283	405,446	149,071	422,718	422,718	435,223	425,204	431,959	424,668
R^2	0.044	0.081	0.022	0.068	0.239	0.071	0.087	0.071	0.130	0.100
Adjusted R^2	0.036	0.073	0.013	0.044	0.232	0.063	0.079	0.063	0.122	0.092
F Statistic	483.642***	907.920***	218.286***	259.897***	3,207.987***	784.536***	1,003.317***	789.927***	1,561.982***	1,136.196***
	(df = 41; 426956)	(df = 41; 420530)	(df = 41; 401696)	(df = 41; 145340)	(df = 41; 418973)	(df = 41; 418973)	(df = 41; 431468)	(df = 41; 421451)	(df = 41; 428204)	(df = 41; 420915)

Table 10 exhibits the results of equation 3.0.3, in which time series averages are constructed for estimations. All liquidity variables present positive and significant coefficients (at the 1% level), confirming the results of Table 8. Adjusted R^2 ranges from 0.188 at the *Roll* measure to 0.483 at the *Range* measure. Table 11 reports the results when DCV and DVP are used as proxies for volatility. As reported in the panel regressions of Table 8, DCV has a positive and highly significant effect on all liquidity variables. Similar to previous results, DVP presents a negative coefficient at the $Volume$, CS and *Roll* measures, although coefficients' level of significance is reduced. *Amihud.turn*, *Range.volume*, *Amihud* and *Semid* exhibit positive and significant coefficients (ranging from the 5% to the 1% significance level).

Table 10 – Time series estimations - Liquidity and $DVIX$

The table reports results for the time-series estimations of equation 3.0.3. We regress our liquidity measures presented in Section 2 against the log-differenced monthly average of VIX. All estimations were made with heteroskedastic and autocorrelation robust standard errors. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016. Significance levels: *** (1%), ** (5%) and * (10%).

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DVIX	0.204***	0.258***	0.459***	0.582***	0.650***	0.544***	0.610***	0.499***	0.602***	0.471***
	(0.057)	(0.069)	(0.091)	(0.109)	(0.069)	(0.082)	(0.072)	(0.055)	(0.063)	(0.057)
Observations	323	323	323	323	323	323	323	323	323	323
R^2	0.388	0.349	0.354	0.284	0.544	0.354	0.428	0.512	0.479	0.461
Adjusted R^2	0.306	0.262	0.268	0.188	0.483	0.268	0.351	0.446	0.410	0.389
Residual Std. Error (df = 284)	0.095	0.120	0.136	0.173	0.102	0.144	0.141	0.109	0.130	0.114
F Statistic (df = 38; 284)	4.735***	4.005***	4.099***	2.965***	8.909***	4.101***	5.592***	7.827***	6.877***	6.391***

Table 11 – Time series estimations - Liquidity, DCV and DVP

The table reports results for the time-series estimations of equation 3.0.4. We switch $DVIX$ for the change in conditional variance (DCV) and the change in variance premium (DVP) proxies, as in Bekaert e Hoerova (2014). All estimations were made with heteroskedastic and autocorrelation robust standard errors. Year dummies are included in estimations but removed from the table for the sake of brevity. Sample period spans from 1990 to 2016. Significance levels: *** (1%), ** (5%) and * (10%).

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DCV	0.101***	0.109***	0.199***	0.236***	0.221***	0.147***	0.151***	0.135***	0.166***	0.128***
	(0.013)	(0.024)	(0.018)	(0.027)	(0.020)	(0.030)	(0.023)	(0.020)	(0.022)	(0.021)
DVP	−0.036**	−0.024	−0.036*	−0.048*	−0.005	0.030***	0.041***	0.022**	0.029***	0.016
	(0.014)	(0.015)	(0.021)	(0.026)	(0.011)	(0.011)	(0.008)	(0.009)	(0.009)	(0.011)
Observations	323	323	323	323	323	323	323	323	323	323
R^2	0.492	0.409	0.519	0.421	0.680	0.366	0.421	0.517	0.496	0.459
Adjusted R^2	0.422	0.327	0.452	0.341	0.636	0.279	0.341	0.450	0.427	0.384
Residual Std. Error (df = 283)	0.087	0.115	0.118	0.156	0.086	0.143	0.142	0.109	0.128	0.114
F Statistic (df = 39; 283)	7.029***	5.020***	7.820***	5.277***	15.419***	4.193***	5.279***	7.768***	7.144***	6.147***

In general, the contemporaneous effects of variations in $DVIX$ and its components on liquidity showed some patterns already found in the financial literature. Although the literature provides few results on the effect of DVP, our estimations show that this variable is related to liquidity measures. Semi-deviation-based measures have presented significant and coherent results so far. Next, we proceed to VAR estimations in order to check the dynamic relationship between liquidity and implied volatility. Table 12 reports the sum of coefficients regarding equations 3.0.5 and 3.0.6, the p-value for the Causality test (GRANGER, 1969), and the adjusted R^2 for each liquidity measure. Lag definition is made by the smallest lag pointed between the Akaike and Hannan-Quinn information criteria. As in previous estimations, year dummies are included.

The results for equation 3.0.5 show that all liquidity measures present strong negative autocorrelation, as shown by column *SumIlliq*. $Turn$ and *Roll* measures present the highest sum of coefficients; close to -0.92. As $Turn$ is based on the number of shares outstanding, which is a quantity that does not vary much from month to month, the measures based on $Turn$ are expected to present high autocorrelation. Volume-based proxies such as *Amihud* and *Semid* present lower sums of lagged coefficients (-0.236 and -0.539) than measures based on $Turn$. Although the evidence is scarce, Chordia et al. (2001) and Chordia et al. (2002) report liquidity measures as negatively autocorrelated. Column $SumDVIX$ reports the sum of coefficients for the impact of lagged $DVIX$ on liquidity measures. Both $Volume$ and $Turn$ present negative values (-0.155 and -0.001, respectively), which may be associated with a late response after an increase on implied volatility at the current time period. $DVIX$ causes (in the sense of Granger) $Volume$ with a p-value lower than the usual 0.01 cut-off. The VAR estimations show that, after spikes in negotiations, trading diminishes after a positive variation on VIX (and vice-versa). Note that results from Tables 8 and 10 show that the contemporaneous effect of $DVIX$ over $Volume$ and $Turn$ is positive but, after this spike, VAR estimations show that the traded volume tends to decrease. Spread measures such as CS and *Roll* present positive sum of coefficients, evidencing that a change in VIX inflicts on higher spreads up to two months ahead. P-values for the Granger Causality tests are significant at the 10% level for CS and at the 1% level for *Roll*. Price impact measures based on $Turn$ (*Amihud.turn* and *Semid.turn*) are also positively affected by lagged shocks on $DVIX$, with significant p-values for the Granger Causality tests. Also, volume-based proxies such as *Range.volume*, *Amihud* and *Semid* are Granger-caused by $DVIX$ at least at the 10% level. Overall, *Semid* and *Semid.turn* have presented the expected effects considering previous results reported on the literature and similar effects when compared to other measures. These results meet those found by Chung e Chuwonganant (2014), where changes in VIX (both contemporaneous and lagged) reduces liquidity measures on individual stocks and the results of Chacko et al. (2016) for bond ETFs. Our results are also in line with Nagel (2012), who shows that liquidity is diminished in market turmoil due to the fact that market makers may demand a premium for providing liquidity. Such increase in liquidity premium may also be associated with lower levels of funding liquidity, as in the model of Brunnermeier e Pedersen (2009). In the presence of shocks to balance sheets (in this case, measured by $DVIX$), financial institutions may reduce funding to dealers and hedge funds in order to mitigate possible losses. Such reduction diminishes the capability of liquidity provision, therefore reducing liquidity in a vicious cycle.

The results from equation 3.0.6 at Table 12 show the effects of liquidity measures affecting

$DVIX$. There are only two liquidity variables that present significant p-values for the Granger Causality test: $Volume$ (p-value: 0.001) and *Range* (p-value: 0.025) with the respective sum of coefficients being positive. The result is in line with $Volume$ as being connected to volatility (ROSSI; MAGISTRIS, 2013). *Range* itself is a proxy for price volatility, so it is expected to be related do $DVIX$. One can notice that other liquidity variables present a small sum of coefficients and insignificant p-values for the Granger Causality test. These results enforce the idea that liquidity has little effect on volatility.

Table 12 – VAR estimations - Liquidity and DVIX

The table reports results for the VAR systems described in equations 3.0.5 and 3.0.6 where the endogenous variables are the liquidity proxies and log-differenced implied volatility ($DVIX$). The sum of coefficients and the Granger Causality test p-value for equations 3.0.5 and 3.0.6 are reported. The maximum lag is defined by the lowest lag pointed between the Akaike and the Hannan-Quinn information criterion. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016.

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Equation 3.0.5										
Sum Illiq	-0.230	-0.923	-0.728	-0.925	-0.238	-0.518	-0.236	-0.859	-0.539	-0.792
Sum DVIX	-0.155	-0.001	0.257	0.407	0.043	0.256	0.165	0.443	0.308	0.389
GC DVIX	0.001	0.330	0.056	0.001	0.506	0.094	0.036	0.001	0.025	0.001
Adj. R Sqrd	0.384	0.452	0.298	0.280	0.115	0.168	0.172	0.353	0.238	0.320
Obs	322	321	321	321	322	321	322	321	321	321
Lags	1	2	2	2	1	2	1	2	2	2
Equation 3.0.6										
Sum DVIX	-0.057	-0.217	-0.238	-0.262	-0.160	-0.189	0.006	-0.169	-0.228	-0.216
Sum Illiq	0.280	0.171	0.164	0.179	0.242	0.020	-0.014	0.001	0.081	0.103
GC Illiq	0.009	0.474	0.258	0.145	0.025	0.977	0.827	0.129	0.738	0.139
Adj. R Sqrd	0.384	0.452	0.298	0.280	0.115	0.168	0.172	0.353	0.238	0.320
Obs	322	321	321	321	322	321	322	321	321	321
Lags	1	2	2	2	1	2	1	2	2	2

Table 13 reports the results when $DVIX$ is decomposed on DCV and DVP. Lagged DCV Granger-causes $Volume$ (at the 1% level) with a negative coefficient, following the rationale of a decrease in $Volume$ after a large swing on variance. DCV also Granger-causes *Amihud.turn* and *Semid.turn*, both at the 1% level, and *Amihud* at the 5% level. The sum of coefficients of these price impact variables is positive, reinforcing the results from Table 12 when using $DVIX$. However, one can notice that the coefficients are much smaller than those reported using $DVIX$. Lagged DVP does not appear to affect liquidity variables at all, as only $Volume$ is Granger-caused by lagged DVP at the 5% level. As this effect is negative, one can think that, when risk aversion (DVP) has increased in previous months, investors may withdraw from trading in the current period.

Results from equation 3.0.8 show the sum of coefficients and p-values for the Granger Causality tests when DCV is the dependent variable. Some liquidity variables present significant effects over DCV. *Roll*, *Range*, *Amihud.turn*, *Semid* and *Semid.turn* present p-values significant within at least the 5% level for the Granger Causality test. The sum of coefficients for these variables is positive (maximum: 0.711 on *Range*), showing that changes in spreads, price impact and stock volatility should affect the conditional variance of the market. The impact of DCV on DVP is not reported to be significant in any case.

At the bottom of Table 13, results from equation 3.0.9 show the sum of coefficients and Granger Causality tests p-values when DVP is the dependent variable of the system. One can notice that, except for the $Turn$ measure, all liquidity proxies present p-values lower than 0.10 for the Granger Causality test. Evidence is scarce on the relationship of the variance premium and liquidity, but we hypothesize that, when negative lagged variations in liquidity occur (higher spreads and higher price impact), risk aversion should increase for the following months. This behavior is endorsed by the estimations of equation 3.0.2. Also, lagged DCV is reported as Granger-causing the DVP. Although this relationship is not well documented by the literature, it is intuitively straightforward: when lagged changes in conditional variance occur, investors may be more risk averse and, therefore, pay the price to be protected against future variance changes (as in a variance swap contract)[18].

[18] The relationship between conditional variance and the variance premium is beyond the scope of this paper. We suggest Carr e Wu (2008) for an introduction to the issue.

Table 13 – VAR estimations - Liquidity, DCV and DVP

The table reports results for the VAR systems described in equations 3.0.7, 3.0.8 and 3.0.9 where the endogenous variables are the liquidity proxies, log-differenced conditional variance (DCV) and log-differenced variance premium (DVP). The sum of coefficients and the Granger Causality test p-value for equations 3.0.7, 3.0.8 and 3.0.9 are reported. The maximum lag is defined by the lowest lag pointed between the Akaike and the Hannan-Quinn information criterion. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016.

	Volume	Turn	CS	Roll	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Equation 3.0.7										
Sum Illiq	-0.157	-0.925	-0.384	-0.410	-0.144	-0.280	-0.233	-0.362	-0.240	-0.332
Sum DCV	-0.064	0.003	-0.012	-0.036	-0.023	0.031	0.053	0.035	0.028	0.019
GC DCV	0	0.761	0.801	0.514	0.569	0.085	0.023	0	0.215	0.009
Sum DVP	-0.001	-0.016	0.011	0.029	0.002	0.011	-0.006	0.008	0.005	0.012
GC VP	0.047	0.245	0.964	0.232	0.741	0.653	0.550	0.218	0.810	0.205
Adj. R Sqrd	0.414	0.453	0.281	0.252	0.113	0.150	0.173	0.308	0.212	0.276
Obs	322	321	322	322	322	322	322	322	322	322
Lags	1	2	1	1	1	1	1	1	1	1
Equation 3.0.8										
Sum DCV	-0.047	-0.150	-0.036	-0.072	-0.134	-0.011	-0.018	-0.048	-0.045	-0.062
Sum Illiq	0.700	0.320	0.289	0.394	0.711	0.216	0.254	0.510	0.396	0.641
GC Illiq	0.125	0.764	0.206	0.022	0.047	0.310	0.237	0.021	0.062	0.006
Sum DVP	0.009	0.011	-0.006	0.002	-0.011	-0.021	-0.025	-0.025	-0.026	-0.025
GC VP	0.494	0.989	0.494	0.494	0.494	0.494	0.494	0.494	0.494	0.494
Adj. R Sqrd	0.017	0.012	0.007	0.017	0.017	0.006	0.008	0.014	0.012	0.023
Obs	322	321	322	322	322	322	322	322	322	322
Lags	1	2	1	1	1	1	1	1	1	1
Equation 3.0.9										
Sum DVP	0.090	0.011	0.095	0.101	0.081	0.075	0.082	0.085	0.082	0.083
Sum Illiq	0.282	-0.481	0.450	0.486	0.079	0.182	-0.038	-0.218	-0.058	-0.140
GC Illiq	0.006	0.633	0	0	0.001	0.018	0.057	0.006	0.018	0.011
Sum DCV	0.254	0.290	0.193	0.167	0.264	0.255	0.287	0.310	0.290	0.299
GC DCV	0	0.001	0	0	0	0	0	0	0	0
Adj. R Sqrd	0.062	0.062	0.066	0.072	0.060	0.062	0.060	0.062	0.060	0.061
Obs	322	321	322	322	322	322	322	322	322	322
Lags	1	2	1	1	1	1	1	1	1	1

The objective of this application is to show how liquidity proxies perform in an empirical exercise relating liquidity and implied volatility. We shed some light on the relationship between liquidity and the components of implied volatility; namely conditional variance and the variance premium. As a byproduct of this application, our proposed measures ($Semid$ and $Semid.turn$) present results aligned with the body of literature on liquidity and asset pricing and are comparable with traditional liquidity proxies.

4 Common factors and tail risk

Following Korajczyk e Sadka (2008), we measure common factors in a principal components analysis (PCA) framework. Consider IL^j the k by t matrix of the liquidity measure j (in log returns) of k stocks on t months. As in Mancini et al. (2013), we decompose the covariance matrix as $IL^j IL^{jT} U^j = U^j D^j$ where U^j is the k by k eigenvector matrix and D^j is the k by k diagonal eigenvalue matrix. Our variable of interest is given by the principal component score matrix, $PC^j = (U^{jT} IL^j)^T$, for each measure. Before the decomposition, the liquidity variables are standardized. In order to be included in the estimations, a stock must pass the initial filters described in Section 2 and do not present more than 4% of monthly missing values (roughly one year). We are aware that such conditions may imply a survivorship bias in our results. However, relaxing the 4% constraint could cause our estimations to become highly unreliable and, thus, we opted to maintain this constraint[19]. The *Roll* spread proxy was removed from our estimations given the low number of stocks fulfilling the conditions to build the principal components. Missing values are estimated via the Expectation-Maximization (EM) algorithm of Korajczyk e Sadka (2008).

After calculating and describing our common factor measures, we run the following time series regression on each stock i on our sample:

$$IL_{it}^j = \alpha_0 + \alpha_1 PC1_t^j + \alpha_2 PC2_t^j + \alpha_3 PC3_t^j + \varepsilon_{it}, \tag{4.0.1}$$

which estimates the effect of common factors on stock liquidity. IL_{it}^j accounts for the j liquidity measure for stock i at month t, $PC1^j$, $PC2^j$ and $PC3^j$ account for the first three common factors for each j measure normalized to be in the 0-1 interval. We average the adjusted R-squared for all stocks for each measure in order to define whether the common factors help explain the liquidity variation. Furthermore, the residual ε_{it} is the idiosyncratic shock on liquidity which is not due to the first three common factors and it shall be used later in this application.

In order to relate common factors in liquidity and tail risk, we calculate risk measures for the assets in our sample. The first risk measure is the Value-at-risk (VaR), defined as:

$$VaR^\alpha(X) = -\inf\{x : F_X(x) \geq \alpha\} = -F_X^{-1}(\alpha), \tag{4.0.2}$$

where X is the log-return of a given asset, α refers to the significance level chosen, F_X is the probability function of X and F_X^{-1} its inverse. Intuitively, VaR^α informs the cutoff point such that a loss will not happen with probability greater than 1-α (JORION et al., 2007). One could point to the need to adjust VaR for liquidity risk, although we do not believe it is necessary, as usually

[19] All the equations proposed in this section are also estimated using a 0%, 10% and 15% missing values cutoff. Results are qualitatively the same.

this adjustment is made through the addition of a spread value[20]. Besides, the non-adjustment for liquidity underestimates risk and, hence, should reinforce our results.

Despite being widely used by practitioners, VaR fails to meet some desired mathematical properties; namely sub-additivity. In some specific cases, the VaR of a portfolio may be higher than the sum of individual assets composing the portfolio, which is contrary to modern portfolio theory (MÜLLER; RIGHI, 2018). In that sense, Artzner et al. (1999) developed an axiomatic framework of risk measures known as coherent risk measures. In short, these measures present desired mathematical properties (axioms) that a risk measure should possess in order to be reliable. The Expected Shortfall (ES) fulfills the proposed axioms by Artzner et al. (1999), therefore being defined as:

$$ES^{\alpha}(X) = -E[X|X \leq F_X^{-1}(\alpha)], \tag{4.0.3}$$

where the notation is the same as in VaR. ES measures the conditional expectation of a loss given that this loss exceeds VaR. In this paper, both VaR and ES are estimated through historical simulation (HS), which is a non-parametric procedure for estimating risk. This method is based on the empirical distribution of asset's returns and requires no distributional assumptions and it is widely used by practitioners. We follow Righi e Ceretta (2016) and estimate these variables by the following equations:

$$\widehat{VaR}^{\alpha}(X) = -(F_X^E)^{-1}(\alpha), \tag{4.0.4}$$

$$\widehat{ES}^{\alpha}(X) = -(N\alpha)^{-1}\sum_{d=1}^{N}(\{X\}_1^N * 1_{\{X\}_1^N} < -\widehat{VaR}^{\alpha}), \tag{4.0.5}$$

where $(F_X^E)^{-1}$ is the inverse distribution function of log-returns (or the quantile function), N is the window size to be used, 1_* is an indicator function that assumes value 1 if returns are lower than the estimated VaR and 0 otherwise.

Next, we estimate the relation of the first common factor for each liquidity measure and market tail risk in a VAR approach. We estimate Granger (1969) causality tests on the following system of equations:

$$PC1_t^j = \delta_0 + \sum_{p=1}^{P}\delta_p PC1_{t-p}^j + \sum_{p=1}^{P}\beta_p Risk_{mt-p}^k + v_t, \tag{4.0.6}$$

$$Risk_{mt}^k = \beta_0 + \sum_{p=1}^{P}\beta_p Risk_{mt-p}^k + \sum_{p=1}^{P}\delta_p PC1_{t-p}^j + e_t, \tag{4.0.7}$$

[20] Many of these studies adjust traditional risk measures such as Value-at-risk by adding a trading cost in order to account for the difficulty of unloading positions in an extreme event. For a review, see Bangia et al. (2001), Angelidis e Benos (2006), Acerbi e Scandolo (2008), Weiß e Supper (2013), Dionne et al. (2015), Soprano (2015) and Wagalath e Zubelli (2017).

where $PC1_t^j$ is the first principal component of liquidity measure j. $Rism_{mt}^k$ refers to tail risk estimates for the daily log-returns of S&P 500. These estimates are based on a 250-day rolling window of the empirical distribution of returns[21]. As an estimate of risk is calculated for every day, we average these variables monthly. The k superscript refers to each risk measure (VaR and ES) and the m subscript refers to the risk of the market. Lag selection is defined by the lowest lag pointed between the Akaike and the Hannan-Quinn information criterion.

The sum of δ_p coefficients on equation 4.0.6 should provide evidence on whether measures of market risk affect the common factor of liquidity. Although volatility measures such as standard deviation are far reported to affect liquidity, tail risk seems to better capture the effects associated with the funding liquidity literature. Additionally, this empirical exercise may have direct applications to portfolio allocation, since many hedge funds supposedly rely on liquidity risk (CHACKO et al., 2016; JAME, 2017). We estimate whether risk is affected by the common factor on equation 4.0.7 through coefficients β_p.

Our next step is to analyze whether the risk of a given asset is affected by the common factor in liquidity. For that end, we estimate the following equation through a fixed-effects estimation:

$$Risk_{it+1}^k = \beta_0 + \beta_1 PC1_t^j + \beta_2 IL^j *_{it} + \zeta Controls_t + \epsilon_t, \tag{4.0.8}$$

where $Risk_{it+1}^k$ refers to VaR and ES estimations for stock i at month $t+1$, $PC1_t^j$ to the first principal component of measure j. The variable IL^j* is the idiosyncratic shock on liquidity given by the residual ε_{it} of equation 4.0.1, accounting for idiosyncratic variation in liquidity which is not due to the first three common factors. As $Controls$, we include the standard deviation of returns and the logarithm of market capitalization of stock i, and the log-difference of VIX, which is related to funding liquidity (CHEN; LU, 2017).

4.1 Results

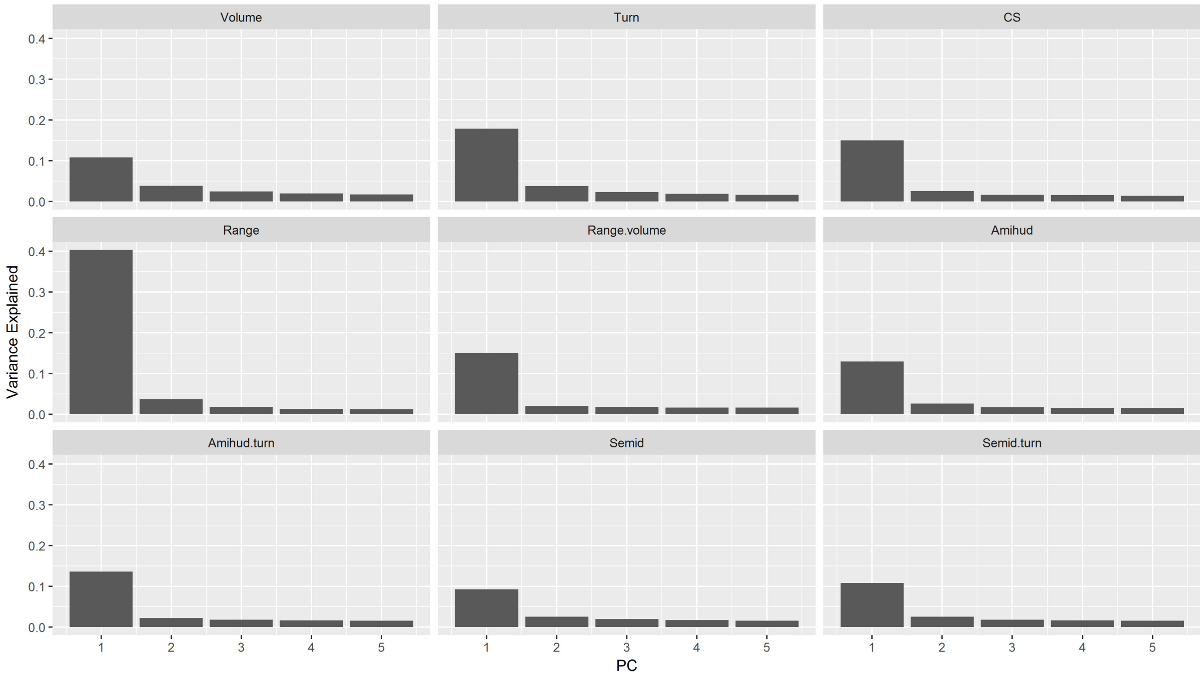
Figure 6 shows the variance explained by the first five common factors extracted from PCA. One can notice the first common factor as explaining a larger part of variance against the other factors. The first PC explains approximately 9% to 18% of total variance in most of our measures. Negotiability variables such as $Volume$ and $Turn$ presented the first common factor explaining 10.8% and 17.9% of total variance, respectively. These values are close to others found in the remaining measures. The spread measure CS (15%), *Range.volume* (15.1%), *Amihud* (12.9%) and *Amihud.turn* (13.6%) have shown similar magnitudes on the first common factors of liquidity. The semi-deviation measures *Semid* (9.3%) and *Semid.turn* (10.8%) have presented marginally lower values than other proxies. The relative small values for the variance explained by the first common factors may be considered expected, once we are modeling liquidity in log-differences, not in levels[22]. The *Range* measure presents the highest level of explained variance

[21] This window size was chosen to mimic Basel III risk parameters commonly used by financial institutions in order to manage tail risk (BIS, 2017).

[22] PCA analysis using liquidity in levels has reported the first common factor explaining a larger part of variance (ranging from 40% to 60%). Results are available under request.

in the first common factor (40.3%), clearly contrasting from the pattern of other measures. This may evidence the *Range* measure as more of a volatility proxy than a liquidity one, although this result is curious and may be related with volatility spillovers among stocks.

Figure 6 – Common factors in liquidity



The figure shows the percentage of variance explained by the first five principal components of the log-differences of liquidity measures described in Section 2. The Roll spread estimator was removed from the estimations given the amount of missing observations. The sample period spans from 1990 to 2016.

Table 14 reports the variance explained by first common factor for each liquidity measure (plotted in Figure 6), the ratio of variance explained by the first PC to the sum of explained variance on the five first principal components of each liquidity proxy and the eigenvalue of the first common factor. The second row allows us to ascertain the relevance of the first common factor relative to other PCs. With the exception of *Range*, the first PC accounts for approximately 50% to 65% of variance explained by the first five principal components, which is reflected by the eigenvalues with similar values, ranging from 5.018 (*Semid*) to 6.787 (*Turn*). Again, the exception is for *Range* in which the importance of first PC is predominant relative to the other four PCs. 83.4% of the explained variance of the first five PCs is accounted for the first one. Accordingly, the eigenvalue of *Range* (9.874) is higher than the other liquidity variables.

Next, we estimate equation 4.0.1 and average the adjusted R-squared of all stocks. Table 15 reports how much common factors can explain idiosyncratic variation in liquidity. As in previous analysis, with the exception of *Range*, the effects of common factors are similar in most of the liquidity variables. Mean adjusted R-squared ranges from 9.5% (*Semid*) to 15.1% (*Turn*) with significant t-statistics. As expected given previous results, common factors in *Range* explain a larger part of variation in stocks' *Range* (28.4%). Mean values do not seem to be affected by outliers as medians are close to means for all liquidity proxies. Wilcoxon rank sum tests were also conducted to test the hypothesis of median values statistically different than zero. In general, *Semid* and *Semid.turn* have presented similar results compared to other proxies. This suggests

Table 14 – PCA statistics

The table reports statistics of Principal Component Analysis of liquidity measures defined in Section 2. Stocks with more than 4% of monthly values (roughly one year) are removed from the sample. For the remaining stocks, missing values are replaced by estimations generated by the Expectation-Maximization (EM) algorithm described by Korajczyk e Sadka (2008). The Roll estimator was removed from the estimations given the amount of missing observations. The first row in the table describes the variance explained by the first common factor extracted. The second row describes the percentage of the variance explained by the first common factor divided by the variance explained by the first five common components. The third row informs the highest eigenvalue of each liquidity variable. The sample period spans from 1990 to 2016.

	Volume	Turn	CS	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Var explained	0.108	0.179	0.150	0.403	0.151	0.129	0.136	0.093	0.108
Relative var exp	0.519	0.653	0.678	0.834	0.679	0.635	0.654	0.545	0.594
1st eigenvalue	5.390	6.787	6.079	9.874	6.042	5.981	5.868	5.018	5.272

that the use of this variables may at least yield comparable results when semi-deviation measures are used.

Table 15 – Adjusted R-Squared of illiquidity on commonality

The table reports statistics on the mean adjusted R-squared from the regressions of all stocks in the sample against the first common factor on each liquidity measure. Mean, median, t-statistics and statistics for the Wilcoxon rank sum test are presented.

	Volume	Turn	CS	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mean	0.100	0.151	0.100	0.284	0.127	0.115	0.114	0.095	0.101
Median	0.080	0.138	0.088	0.281	0.121	0.108	0.110	0.083	0.092
T-stat	49.880	65.671	50.863	83.071	64.763	61.129	61.618	50.513	55.058
W-stat	5e+06	6e+06	5e+06	5e+06	5e+06	5e+06	5e+06	5e+06	5e+06

Table 16 reports results for the VAR system defined in equations 4.0.6 and 4.0.7 when using the S&P 500 5% VaR as a measure of market risk. Information criteria has pointed two lags. By the results of equation 4.0.6, one can notice a negative autocorrelation for the first principal component in all liquidity variables, in which the highest sum of lagged coefficients is reported in *Range.volume* (-0.044) and the lowest to *Turn* (-1.059). When analyzing the effect of VaR at the first common factor, one can notice the sum of lagged coefficients as being positive, suggesting that increases in market risk in the past 2 months may increase the common factor in all variables. This originates distinct interpretations regarding different liquidity measures, although most of them are straightforward in finance. *Volume* (sum of coefficients: 6.202) and *Turn* (9.595) should be increased in severe market downturns as investors may unload positions. Trading activity may increase volatility, thus increasing risk (RUBIA; SANCHIS-MARCO, 2013). The *CS* (5.515) measure is also expected to increase in risk, as is moments of distress volatility effects should widen spreads (CHORDIA et al., 2005). Although this effect is usually reported in contemporaneous relations, Table 16 suggests a temporal dependence. The *Range* (9.323) measure as a proxy for volatility is also positively affected by market risk. The relation between risk and volatility is straightforward. Our measures of price impact, namely *Range.volume* (4.298), *Amihud* (4.874) and *Amihud.turn* (3.622) are also positively affected by VaR. As risk increases, higher volatility should empty assets' depth and, hence, increase spreads. The semi-deviation proxies *Semid* (3.359) and *Semid.turn* (1.090) present lower sums of coefficients, although both are positive. The fact that semi-deviations already account for negative returns in their construction may be a reason why the coefficients are smaller, although this can be investigated in further studies.

Overall, the liquidity proxies have presented expected coefficients on the relationship between VaR and common factors. As the model of Brunnermeier e Pedersen (2009) predicts, if a negative shock occurs in the economy (in this case, the shocks are represented by VaR), intermediaries should increase margins on borrowers (dealers, hedge funds and other investors), forcing them to liquidate long positions originated by leverage. Table 16 suggests that such an effect occurs lagged up to two months. To the best of our knowledge, this is the first evidence on the issue. Although appealing, one should be aware that the Granger-causality tests present low significance in their p-values reported at Table 16. Common factors of *Range* are Granger-caused by VaR at the 10% level, whereas the first common factors of $Turn$ and *Range.volume* are Granger-caused at the 5% significance level.

Results from equation 4.0.7 at Table 16 show VaR as a highly persistent variable where the sum of lagged coefficients is close to unit in all liquidity variables. This is expected as the calculation of risk in a 250-day rolling window should imply time dependence[23]. The sum of coefficients of $PC1$ exhibits a negative effect, in which Granger causality tests show evidence of the first common factors in *Range* (sum of coefficients: -0.002), *Semid* (-0.002) and *Semid.turn* (-0.003) as Granger-causing VaR at the 10% level and *Range.volume* (-0.003), *Amihud* (-0.002) and *Amihud.turn* (-0.003) at the 5% level. This result is somehow counterintuitive to the model of Brunnermeier e Pedersen (2009) and overall supply-side explanations of commonality, as it suggests that higher common factors (measured by $PC1$) decrease market risk. Thus, the sum of coefficients is small.

[23] This may raise concerns about the stability of the VAR system. Nevertheless, in unreported results the roots of the coefficient matrices are all reported as less than one. Additionally, we have estimated the VAR equations using the difference in $Risk_t^m$. Results are qualitatively the same and are omitted for the sake of brevity.

Table 16 – VAR estimations - Common factors and VaR

The table reports results for the VAR system described in equations 4.0.6 and 4.0.7, where the endogenous variables are the first common factor extracted by PCA ($PC1$) and the monthly average of the 5% Value-at-risk of S&P 500 returns on a 250-day rolling window ($Risk^m$). The sum of coefficients estimated through VAR and the Granger Causality test p-value for equations 4.0.6 and 4.0.7 are reported. The maximum lag is defined by the lowest lag pointed between the Akaike and the Hannan-Quinn information criterion. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016.

	Volume	Turn	CS	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Equation 4.0.6									
Sum PC1	-0.645	-1.059	-0.747	-0.525	-0.044	-0.166	-0.622	-0.298	-0.668
Sum VaR	6.202	9.595	5.515	9.323	4.298	4.874	3.622	3.359	1.090
GC VaR	0.223	0.039	0.151	0.054	0.033	0.134	0.256	0.245	0.270
Adj. R Sqrd	0.103	0.311	0.196	0.103	0.006	-0.010	0.128	0.018	0.149
Obs	322	322	322	322	322	322	322	322	322
Lags	2	2	2	2	2	2	2	2	2
Equation 4.0.7									
Sum VaR	0.976	0.981	0.985	0.994	0.993	0.991	0.987	0.992	0.988
Sum PC1	0	-0.001	-0.001	-0.002	-0.003	-0.002	-0.003	-0.002	-0.003
GC PC1	0.767	0.673	0.451	0.081	0.048	0.036	0.038	0.063	0.059
Adj. R Sqrd	0.984	0.984	0.984	0.984	0.985	0.985	0.984	0.985	0.984
Obs	322	322	322	322	322	322	322	322	322
Lags	2	2	2	2	2	2	2	2	2

Table 17 reports coefficients and statistics from equations 4.0.6 and 4.0.7 using the 5% ES as $Risk^m$. The results are similar to those reported at Table 16 despite small differences in the Granger causality p-values. Results from equation 4.0.6 show that ES Granger-causes common factors in $Volume$ (sum of coefficients: 4.947), $Turn$ (7.884) and $Range$ (8.612) at the 1% level, $Range.volume$ (5.124) at 5% and $Amihud$ (4.550) at the 10% level, enforcing previous results. Using ES in equation 4.0.7 also yields similar outcomes when compared to VaR. Common factors of $Amihud.turn$ (sum of coefficients: -0.003) and $Semid.turn$ (-0.003) Granger-causes ES at the 5% level, whereas $Amihud$ (-0.002) and $Semid$ (-0.002) Granges-causes ES at 10%. As in Table 16, the sum of coefficients is small. Tables 16 and 17 presents further evidence on the time series relation between common factors of liquidity measures and market tail risk. Although results of equation 4.0.6 using both traditional proxies and the semi-deviation proposed proxies supports supply-side explanations of commonality, results from equation 4.0.7 present curious evidence of lagged common factors negatively affecting market risk, whereas the expected effect would be the opposite. Therefore, we have investigated the effect of common factors on market risk, not on idiosyncratic risk of individual stocks.

Table 17 – VAR estimations - Common factors and ES

The table reports results for the VAR system described in equations 4.0.6 and 4.0.7, where the endogenous variables are the first common factor extracted by PCA ($PC1$) and the monthly average of the 5% Expected Shortfall of S&P 500 returns on a 250-day rolling window ($Risk^m$). The sum of coefficients estimated through VAR and the Granger Causality test p-value for equations 4.0.6 and 4.0.7 are reported. The maximum lag is defined by the lowest lag pointed between the Akaike and the Hannan-Quinn information criterion. Year dummies are included in estimations but removed from the table for the sake of brevity. The sample period spans from 1990 to 2016.

	Volume	Turn	CS	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Equation 4.0.6									
Sum PC1	-0.626	-1.069	-0.771	-0.534	-0.072	-0.191	-0.673	-0.343	-0.713
Sum ES	4.947	7.884	4.856	8.612	5.124	5.522	4.550	4.695	2.773
GC ES	0.004	0.003	0.138	0.007	0.028	0.074	0.218	0.345	0.509
Adj. R Sqrd	0.123	0.321	0.199	0.105	0.003	-0.012	0.129	0.015	0.144
Obs	322	322	322	322	322	322	322	322	322
Lags	2	2	2	2	2	2	2	2	2
Equation 4.0.7									
Sum ES	0.971	0.975	0.975	0.990	0.980	0.979	0.974	0.980	0.977
Sum PC1	0.000	-0.001	0.000	-0.002	-0.002	-0.002	-0.003	-0.002	-0.003
GC PC1	0.375	0.343	0.388	0.128	0.203	0.089	0.042	0.053	0.037
Adj. R Sqrd	0.979	0.979	0.979	0.980	0.979	0.979	0.979	0.979	0.979
Obs	322	322	322	322	322	322	322	322	322
Lags	2	2	2	2	2	2	2	2	2

Table 18 reports results for equation 4.0.8 in which a fixed-effects panel was estimated to assess the effect of common factors on idiosyncratic risk (measured by stocks' VaR). One can note a pattern as all common factors (with the exception of $Volume$) present significant positive coefficients on $PC1$ (ranging from 0.001 on $Turn$ to 0.017 on *Range.volume*). This implies that higher values of the common factor at month t are associated to higher risk at month $t+1$. This could be caused by both demand-side and supply-side explanations of CIL. Correlated trading in moments of assets' distress may force prices down, causing flights to quality, or to liquid assets. However, such correlated trading may also be a result of higher funding constraints of investors (dealers, hedge or even leveraged mutual funds). The result is different than those reported at Table 16, in which market risk has been modeled instead of idiosyncratic risk. As in the model of Brunnermeier e Pedersen (2009), margins are probably set by assets' (or portfolio's) VaR, not on market VaR. Again, the *Semid* and *Semid.turn* variables have coefficient values (0.011 and 0.009, respectively) similar to proxies such as *Amihud* and *Amihud.turn* (0.015 and 0.009, respectively). The liquidity shock measured by IL^* presents a negative effect in all liquidity variables. Market capitalization has a negative effect on all estimations as small stocks should present more risk (FAMA; FRENCH, 1993). As expected, Sd (standard-deviation) at month t is positively associated to risk in the next month, with all coefficients presenting similar values (approximately 0.40) in all variables. The sign of $DVIX$ coefficients is different when liquidity measures are $Volume$ and $Turn$. Apart from these variables, both spreads, price volatility and price impact proxies report positive coefficients, which is *a priori* expected as a change in market uncertainty should increase risk. In unreported results, the use of VIX in levels instead of differences yield similar results. Still, we opt to keep this variable in log-differences in order to avoid non-stationarity. Main results of Table 18 hold when VaR is changed by ES at Table 19. $PC1$ coefficients are also positive for all liquidity proxies except for $Volume$. Effects of controls IL^*, $log(Mktcap)$ and Sd are similar to those reported in Table 18. A major difference when changing risk measures is the coefficient on $DVIX$, which becomes more unstable as coefficients vary on their sign depending on the liquidity variable used.

Table 18 – Individual risk (VaR) and common factors

The table reports results for the fixed-effects estimations of equation 4.0.8. We regress the monthly average of the 5% daily Value-at-risk calculated within a 250-day rolling window ($Risk^k$) at month $t+1$ against the first common factor extracted by PCA ($PC1$) at month t for each liquidity measure j. IL^* is the residual of idiosyncratic change in liquidity against PC1, as in equation 4.0.1. Control variables include: $log(Mktcap)$ as the logarithm of market capitalization of stock i at month t, Sd as the standard-deviation of daily returns of stock i at month t, $DVIX$ as the monthly average of log-difference of the VIX index. All estimations use heteroskedastic and autocorrelation robust standard errors. The sample period spans from 1990 to 2016. Significance levels: *** (1%), ** (5%) and * (10%).

	Volume	Turn	CS	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PC^1	−0.001***	0.001***	0.008***	0.008***	0.017***	0.016***	0.009***	0.011***	0.009***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
IL^*	−0.002***	−0.003***	−0.001***	−0.008***	−0.0002***	−0.002***	−0.002***	−0.004***	−0.003***
	(0.0001)	(0.0001)	(0.00004)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
log(Mktcap)	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Sd	0.411***	0.419***	0.407***	0.437***	0.412***	0.423***	0.413***	0.447***	0.433***
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
DVIX	−0.003***	−0.002***	0.001***	0.002***	0.006***	0.007***	0.001***	0.004***	0.0003**
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0002)
Observations	356,899	356,899	356,899	356,899	356,899	356,899	356,899	356,899	356,899
R^2	0.181	0.182	0.180	0.186	0.183	0.184	0.181	0.187	0.184
Adjusted R^2	0.173	0.175	0.173	0.178	0.175	0.177	0.174	0.180	0.177
F Statistic (df = 5; 353752)	15,585***	15,732***	15,532***	16,115***	15,811***	16,002***	15,687***	16,276***	15,992***

Table 19 – Individual risk (ES) and common factors

The table reports results for the fixed-effects estimations of equation 4.0.8. We regress the monthly average of the 5% daily Expected Shortfall calculated within a 250-day rolling window ($Risk^k$) at month $t+1$ against the first common factor extracted by PCA ($PC1$) at month t for each liquidity measure j. IL^* is the residual of idiosyncratic change in liquidity against PC1, as in equation 4.0.1. Control variables include: $log(Mktcap)$ as the logarithm of market capitalization of stock i at month t, Sd as the standard-deviation of daily returns of stock i at month t, $DVIX$ as the monthly average of log-difference of the VIX index. All estimations use heteroskedastic and autocorrelation robust standard errors. The sample period spans from 1990 to 2016. Significance levels: *** (1%), ** (5%) and * (10%).

	Volume	Turn	CS	Range	Range.volume	Amihud	Amihud.turn	Semid	Semid.turn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PC^1	0.0001	0.003***	0.014***	0.014***	0.024***	0.022***	0.013***	0.016***	0.012***
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0003)
IL^*	−0.005***	−0.006***	−0.002***	−0.013***	0.001***	−0.004***	−0.003***	−0.006***	−0.006***
	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)
log(Mktcap)	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Sd	0.790***	0.805***	0.778***	0.830***	0.783***	0.800***	0.787***	0.846***	0.826***
	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)
DVIX	−0.008***	−0.007***	−0.002***	0.001**	0.004***	0.005***	−0.003***	0.001***	−0.004***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Observations	356,899	356,899	356,899	356,899	356,899	356,899	356,899	356,899	356,899
R^2	0.211	0.214	0.209	0.218	0.211	0.213	0.210	0.219	0.216
Adjusted R^2	0.204	0.207	0.202	0.211	0.204	0.206	0.203	0.212	0.209
F Statistic (df = 5; 353752)	18,970***	19,228***	18,670***	19,718***	18,906***	19,152***	18,842***	19,882***	19,512***

5 Concluding remarks

In this paper we provide an encompassing approach to compare up to ten liquidity measures in two empirical applications. The objective of our applications is not to rank performance between the measures, but rather to show that the distinct proxies capture similar effects of liquidity. This is achieved through two studies related to the finance literature. Two proposed proxies capture similar effects when compared to usual liquidity variables.

In the first empirical exercise, we show the relationship between changes in liquidity, implied volatility, conditional variance and the variance premium. Our results show that changes in VIX precede decreases in trading activity and increases in illiquidity. Using the decomposition of Bekaert e Hoerova (2014), we notice that the conditional variance component of VIX is positively related to illiquidity variables. This effect is also pronounced within the VAR estimations for the equally-weighted average of liquidity for the stocks in our sample. Still, one must be aware that the decomposition used may present some noise. Thus, a further path for this issue would be to test different modeling. Also, we focused on time series estimations, whereas cross-section analysis for the sensibility of liquidity measures may provide interesting insights to asset pricing.

The second empirical exercise addressed the relationship between common factors in liquidity and tail risk. We based our estimations on the model of Brunnermeier e Pedersen (2009), which predicts that high-risk assets (in their model, measured by Value-at-Risk) should have higher margins for funding. Thus, we hypothesize that higher risk should entail increases in common factors. When we model S&P 500 risk using both Value-at-risk and Expected Shortfall, there is moderate evidence of risk increasing common factors of liquidity. When we analyze idiosyncratic risk and common factors, we note that the latter has a positive effect on the former. A suggestion for further research would be to test distinct risk measuring as our estimations rely on historical simulation. Additionally, our approach on measuring common factors turns difficult to investigate cross-section relations so a path for further research would be to analyze how risk is related to the sensibility of stocks' liquidity to market liquidity (liquidity risk) in an LCAPM-fashion (ACHARYA; PEDERSEN, 2005; ANTHONISZ; PUTNIŅŠ, 2016; RUENZI et al., 2016).

Another path for future research would be to run comparisons of liquidity measures in order to check correlation between semi-deviation-based and transaction data proxies for liquidity, as in Goyenko et al. (2009). Also, we believe that studying our proposed measures using intraday data may yield interesting insights given its natural adjustment for downside/negative returns. Due to its mathematical properties, portfolio optimization using semi-deviations should also produce interesting insights. Through the results achieved in this paper, we believe there is a strong appeal for the use of our proposed measures within both theoretical and empirical work in finance.

Given the close relationship between liquidity, implied volatility and tail risk, our findings have implications for both investors and policy makers. Exogenous shocks to volatility may suddenly dry up liquidity as dealers cannot finance their activities, creating a vicious cycle to asset prices. As we highlight the connection between liquidity, implied volatility and tail risk, policy makers may act to facilitate funding in such cases of shocks to volatility, preventing a widespread effect to markets. For portfolio managers, we show that common factors in liquidity may increase assets' risk. If a fund is highly leveraged, an increase in risk may trigger margin calls causing a spillover effect. Thus, we highlight the importance of managing not only individual risks, but overall exposure attributed to leverage.

Part III

Does algorithmic trading harm liquidity? Evidence from Brazil

Abstract

This paper provides the first evidence of algorithmic trading (AT) reducing liquidity in the Brazilian equities market. Our results are contrary to the majority of work which finds a positive relationship between AT and liquidity. Using the adoption of a new data center for the B3 exchange as an exogenous shock, we report evidence that AT increased realized spreads in both firm fixed-effects and vector autoregression estimates for 26 stocks between 2017 and 2018 using high-frequency data. We also provide evidence that AT increases commonality in liquidity, evidencing correlated transactions between automated traders.

Keywords: Liquidity, Algorithmic trading, Spreads, Commonality in liquidity.

Note: this article was firstly submitted to the Emerging Markets Review and now is in first round review (minor revisions) in the North American Journal of Economics and Finance (2018 JCR Impact Factor: 1.119, Qualis A2). An earlier version of this paper was presented at the 2019 Brazilian Finance Meeting.

1 Introduction

Liquidity is assumed to be systematically increasing in capital markets. Financial integration between different countries, new regulations and technology improvements have made access to financial markets easier (EVANS; HNATKOVSKA, 2014). One channel which through liquidity is increasing is the advent of algorithmic trading (AT) and high-frequency traders (HFT)[24]. These traders can easily detect order imbalances or transitory arbitrage opportunities through fast analysis of market data, news and other public announcements (HOLDEN et al., 2014). As O'Hara (2015) cites, "markets are different now, transformed by technology and high-frequency trading", as any trading strategy may be automated and executed in time frames not distinguishable by human eyes. Much discussion has been had to understand whether this controversial type of trading benefits or harms financial markets. Our paper sheds light on the Brazilian equities market, highlighting the negative effect of AT on liquidity after the adoption of a data center with increased capacity for the Brazilian exchange.

AT/HFT activity has drawn attention in recent years, generating conflicting views and evidence regarding its effects on financial markets. A variety of studies present benefits for market quality as these types of traders should engage in market-making strategies that supply liquidity through fast quote updates, therefore acting as voluntary market makers. Hendershott et al. (2011) provide the first evidence of AT improving price discovery and reducing adverse selection costs in the US market. Hasbrouck e Saar (2013) propose a measure of low-latency activity as a proxy for AT/HFT. As automated trading should be based on a large amount of message traffic, the authors calculate a proxy based on the number of order submissions, order changes and cancellations within millisecond intervals. Quoted spreads and price impact of trades are reduced within increased low-latency activity on NASDAQ stocks. Additionally, a large body of literature has shown positive effects of AT/HFT on spreads and overall transaction costs (ANAGNOSTIDIS; FONTAINE, 2018; BENOS; SAGADE, 2016; BROGAARD; GARRIOTT, 2019; CONRAD et al., 2015; MENKVELD, 2013; MENKVELD, 2016)[25]. If algorithms are able to acquire and process information faster than a human trader, market quality is expected to increase as rapid agents can reduce noise in news, public announcements and corporate reports. If monitoring costs are reduced through machines, transaction costs are expected to fall as well. ATs may act as informed traders, setting prices more efficiently, thus reducing transaction costs (JOVANOVIC; MENKVELD, 2016; MORIYASU et al., 2018).

On the other hand, the media has portrayed high-frequency traders as the starting agents of the May 2010 Flash Crash, a market event that lasted for approximately thirty minutes and resulted in more than a trillion dollar decrease in prices. Although there is evidence of participation of HFTs, Kirilenko et al. (2017) show that theses traders participated in the event

[24] Although both concepts are not entirely unambiguous, HFT is considered a subset of algorithmic trading. Empirical studies have found that AT activity is highly correlated to HFT activity (HAGSTRÖMER; NORDEN, 2013; SKJELTORP et al., 2015).

[25] For a full review on studies in this literature, we suggest Jones (2013), Linton e Mahmoodzadeh (2018) and Virgilio (2019).

by responding to market volatility, but did not initiate the event itself. The overall effect of these traders is still unclear, since AT/HFTs may engage in strategies based on adversely selecting slow investors given their speed advantage (BIAIS et al., 2015). If HFTs anticipate short-period direction of prices, they might act as aggressive traders, therefore reducing liquidity.

In that sense, Cartea et al. (2019) provide evidence that ultra-fast traders reduce intraday market quality in NASDAQ stocks. Since high-frequency traders can receive information and learn about order flow incredibly fast, these agents may detect a benefit in prejudice of slow traders. Empirically, Hirschey (2017) provides evidence of anticipatory trading of HFT, potentially increasing non-HFT trading costs. Thus, informed traders should dry liquidity from the market, since they are expected to trade aggressively in order to anticipate price movements. In short, fast traders may act as both demanding and supplying liquidity, which could be both negative and positive to the market (BROGAARD et al., 2014). One difficulty lies in the fact that even in detailed databases, there is no way to ascertain which trading strategy the HFT is engaging in. Biais et al. (2015) have developed a theoretical model in which the interaction of slow and fast traders induce adverse selection costs for the former, creating an advantage for the latter.

Using a sample of 42 countries, Boehmer et al. (2018) show that AT activity increases volatility worldwide despite a reduction in spreads. Theoretical models where HFT may potentially harm other investors are present in the literature (BUDISH et al., 2015). Biais et al. (2015) and Foucault et al. (2017) provide theory that fast traders may cause negative externalities originated from adverse selection costs to non-HFT traders[26].

Previous research on the subject of AT is centered in developed markets, mainly North America and Europe, where detailed data is available. Studies regarding AT/HFT in emerging markets have been partially made by Boehmer et al. (2018), who investigated the effects of co-location on market liquidity of 42 trading venues. Although financial volume in emerging markets is much lower than in developed economies, returns can be significantly higher (LESMOND, 2005). Additionally, emerging markets may present more price inefficiencies (HULL; MCGROARTY, 2014) and may be more dependent on foreign capital flows. This may draw attention of automated traders searching for transitory arbitrage opportunities. Lee (2015) finds no evidence of AT/HFT reducing spreads in the Korean futures market, but rather that it hampers the price discovery process. Jawed e Chakrabarti (2018) study the speed of information adjustment and persistence in different indexes for the Indian stock market after the introduction of co-location services.

Given the puzzling effects of algorithmic trading on liquidity and the lack of research on emerging markets, we contribute by addressing the issue in the Brazilian equities market. We use the date when a new data center started to operate in the Brazilian exchange (B3) as an exogenous shock to algorithmic trading activity. Our proxies for AT are the volume-message ratio proposed by Hendershott et al. (2011) and the message-trades ratio used by Malceniece et al. (2019). Our sample spans 320 trading sessions from 2017 to 2018 and comprises 26 stocks which do not have designated market makers. Liquidity is measured through realized spreads and the high-frequency Amihud (2002) price impact measure.

We conduct tests to certify our exogenous shock as a valid instrument for the level of

[26] Hoffmann (2014) provides theory with similar conclusions for order-driven markets.

algorithmic trading. In order to remove noise from estimations, the level of AT is given by 2SLS estimates from a set of instrumental and control variables. Contrary to many studies for developed markets, our firm fixed-effects estimations show the level of AT as increasing both realized spreads and price-impact variables. As most of the literature studies the effect of AT in liquidity, we also estimate the bidirectional effect using vector autoregression (VAR) models. Our results are consistent with lagged AT increasing spreads. Our results are weaker for the price-impact proxy, suggesting that ATs do not trade based on private information (MESTEL et al., 2018). A methodological contribution from this study is to use high-frequency data aggregated through 1-minute intervals. Most of the literature uses data aggregated on a daily basis. The very nature of AT is time-sensitive; thus, it is important to measure this variable on a high-frequency basis. Results are robust when data is aggregated in 5 and 15 minute intervals and on a daily basis.

Our study also addresses the relationship between AT and commonality in liquidity (CIL). On one side, algorithmic traders could better parse firm-specific information. If information is quickly incorporated to prices, commonality is expected to drop (MORIYASU et al., 2018; MORCK et al., 2000). On the other side, if trading strategies from AT are correlated, an increase in CIL is expected as trades occur based on similar triggers for action. The literature provides evidence of correlated trading from HFT in the US market (BROGAARD, 2010) and in the FX market (CHABOUD et al., 2014). As one of the well-known strategies of AT/HFT is market-making, such traders are expected to trade not only one stock, but a basket of them. Thus, shocks in funding liquidity or in asset returns may force these voluntary market makers to create commonality through the liquidation of their positions (BRUNNERMEIER; PEDERSEN, 2009; HAGSTRÖMER; NORDEN, 2013; MENKVELD, 2013). Our approach is to measure intraday CIL through the R squared of a regression of liquidity proxies on market liquidity. AT activity increases commonality in realized spreads, suggesting that algorithms may present correlated trading strategies. The Brazilian market has a small number of liquid stocks compared to other markets, restricting options to trade, therefore inducing CIL.

The main contribution of our paper is to provide the first detailed evidence of the effect of AT on liquidity for the Brazilian market. The Federal Bank of Brazil started to cut interest rates in 2016, leading retail investors to migrate from fixed income investments to the stock market. With a higher number of slow investors, understanding the effects of fast trading is imperative. Our results are contrary to the majority of findings for developed markets in which AT is beneficial to liquidity. As emerging markets tend to behave differently than developed ones, our results evidence the necessity of studies focusing on emerging markets. We also show that commonality in liquidity (CIL) is higher when AT activity is high. Our methodological contribution presents estimations through high-frequency data not aggregated on a daily basis, but rather using the effects within intraday frequencies as in Jain et al. (2016). Thus, our research contributes to the large body of literature studying the effects of AT on liquidity and to the growing number of papers addressing commonality and AT (JAIN et al., 2016; KLEIN; SONG, 2017; MALCENIECE et al., 2019; MORIYASU et al., 2018).

The paper proceeds as follows: Section 2 describes data and variables, Section 3 presents our methodology and results. The final section concludes the paper with our remarks and directions for future work.

2 Data and variables

Before 2008 the Brazilian financial market was concentrated in two main exchanges: Bovespa (São Paulo Stock Exchange), which traded mainly equities, and BM&F (Brazilian Mercantile and Futures Exchange), which negotiated commodities, futures and other derivatives. In 2008, Bovespa and BM&F merged as BM&FBovespa. Later, in 2017, BM&FBovespa merged with the clearing house CETIP, forming the B3 company, one of the largest exchanges in Latin America. According to B3's website, in March 2019 there were 430 companies listed on the stock market[27]. The B3's 2018 Annual Report shows an average daily trading value of R$ 12.3 billion in equities and equities derivatives[28].

The Brazilian case is an interesting market to be studied since there is only one trading venue for equities. Many of the previously presented studies allege that the profitability of AT/HFT comes from market fragmentation, since these traders should engage in low-latency arbitrage strategies by taking advantage of spread differences between trading venues[29]. Unlike other exchanges, there is no market fragmentation in Brazil, so one can conjecture as to whether AT/HFT are present in this market and to what extent they contribute for liquidity. Also, few studies have approached liquidity in the high-frequency world in Brazil (VICTOR et al., 2013; PERLIN, 2013; RAMOS et al., 2017). In 2010, the equities segment of BM&FBovespa started to rent co-location spaces within the exchange data center. Although the use of AT/HFT is not documented in Brazil, it is expected that the reduced latency time provided by co-location has attracted fast traders within recent years as the use of this type of trading is a reality in developed markets.

The B3 exchange provides open access to trade and order data. For each day and market (equities, odd lot equities, options and futures), there are two main categories of files hosted at the B3 FTP[30]. The first category reports all trades within a specific market at the nanosecond time stamp. The second category of files provides the limit order book (LOB) for a given day and a given side of the LOB. For example, there is one file for the buy-side order book and one file for the sell-side order book. These files report order changes, submissions, cancels and trades. Each order is reported with a unique code for identification. The files are downloaded and organized through the R package *GetHFData* from Perlin e Ramos (2017).

Even with increasing computer power and tools to handle high-frequency data, recent related studies have used samples with small time-frames in research regarding AT/HFT activity[31]. Our sample includes 320 trading sessions from 2017-04-03 to 2018-07-30 for stocks traded on the

[27] Source: <http://www.b3.com.br/en_us/products-and-services/trading/equities/listed-companies.htm>.

[28] R$ denotes Brazilian Reals. According to the Central Bank of Brazil, the USDBRL exchange rate was R$ 4.15 = US$ 1.00 as of October 2019.

[29] These type of strategies aim to profit from the time difference between order arrivals in different trading venues for the same asset. For details, see SEC (2010a) and Foucault et al. (2017).

[30] File Transfer Protocol.

[31] For example, Hirschey (2017) used a dataset of one year (2016). Hasbrouck e Saar (2013) used two samples of one month each. Upson e Ness (2017) analyzed three months of trading data.

B3 exchange. We removed May 18 of 2017 so as not to bias our estimates. This trading session was subsequent to the event known as "Joesley Day", on which a tape recording of the possible involvement of the former president in a corruption scandal was published. On 2017-05-18, the B3 index melted more than 10%, activating the circuit breaker protocol during the trading session, and the USDBRL rose more than 8%.

Our initial sample started with 430 tickers. We filtered our sample as in Acharya e Pedersen (2005) and Pastor e Stambaugh (2003). A stock was included in the estimations if:

- it is a common or a preferred stock (excluding other classes such as units, REITs, ETFs and Brazilian Depositary Receipts);
- its closing price was higher than R$ 5.00 (around 1.20 USD) during the sample period;
- its number of trades was higher than the median for all stocks available in the sample period;
- it was negotiated in all trading sessions through the sample period;
- it does not have a registered market maker.

The final sample includes 26 stocks (out of 430) traded on the B3 exchange. These filters are required since many stocks lack liquidity and therefore may bias the results. When both common and preferred stocks passed the filters for the same company, we selected the stock with the highest mean traded volume over the sample. Not having a designated market maker is important for our objectives, since higher AT activity could be assigned to a market maker who is forced to place quotes within price ranges defined by contract. This restriction is not explicitly described in other studies in the literature. It is worth noting that all market making contracts are available on B3's website.

2.1 Proxies for AT and Liquidity

We use two proxies for AT activity. The first is calculated by the ratio of traded volume (in hundreds) to messages (new order changes or submissions, cancellations and trades) as in Hendershott et al. (2011):

$$\text{ATmsg}_{itd} = -\frac{\text{R\$100.00 of trading volume}_{itd}}{\text{Msg}_{itd}}, \tag{2.1.1}$$

where Msg_{itd} refers to the sum of new order submissions, cancellations and trades for stock i at time t of day d. The time interval t is set to 1 minute. The nominator is equal to the total volume traded at time t divided by R$ 100.00 so the measure can be comparable between stocks. We take the negative sign for the measure, so higher values (closer to zero) reflect an increase in algorithmic trading activity.

As in Malceniece et al. (2019), we include the $ATtrades$ proxy as the ratio of the number of messages of a given asset to the number of trades:

$$\text{ATtrades}_{itd} = \frac{Msg_{itd}}{Trades_{itd}}, \tag{2.1.2}$$

where $Trades_{itd}$ refers to the number of trades for the asset i at minute t of day d. Both proxies evidence the number of messages as a main characterization of algorithmic trading.

In order to investigate the effect of AT activity on liquidity, measures for liquidity are calculated. We define the *RealizedSpread* as in Goyenko et al. (2009):

$$\text{RealizedSpread}_{itd} = \begin{cases} 2(ln(P_{itd}) - ln(P_{it+jd})), \text{if the trade at t is a buy.} \\ 2(ln(P_{it+jd}) - ln(P_{itd})), \text{if the trade at t is a sell.} \end{cases} \tag{2.1.3}$$

where ln refers to the natural logarithm, P_{itd} refers to the traded price of stock i at time t of day d. In this case, the spread is measured by the price difference between trades in a j interval set to 5 minutes.

For a price-impact proxy, we use the high-frequency Amihud (2002) measure:

$$\text{Amihud}_{itd} = \frac{|R_{itd}|}{Volume_{itd}}, \tag{2.1.4}$$

where R_{itd} is the log-return of traded prices for stock i at time t of day d. $Volume_{itd}$ accounts for the sum of traded volume for stock i during the t time interval of day d. All variables have their top (bottom) 1% of observations removed for each day and each stock. This filter is necessary given the low liquidity of the Brazilian market.

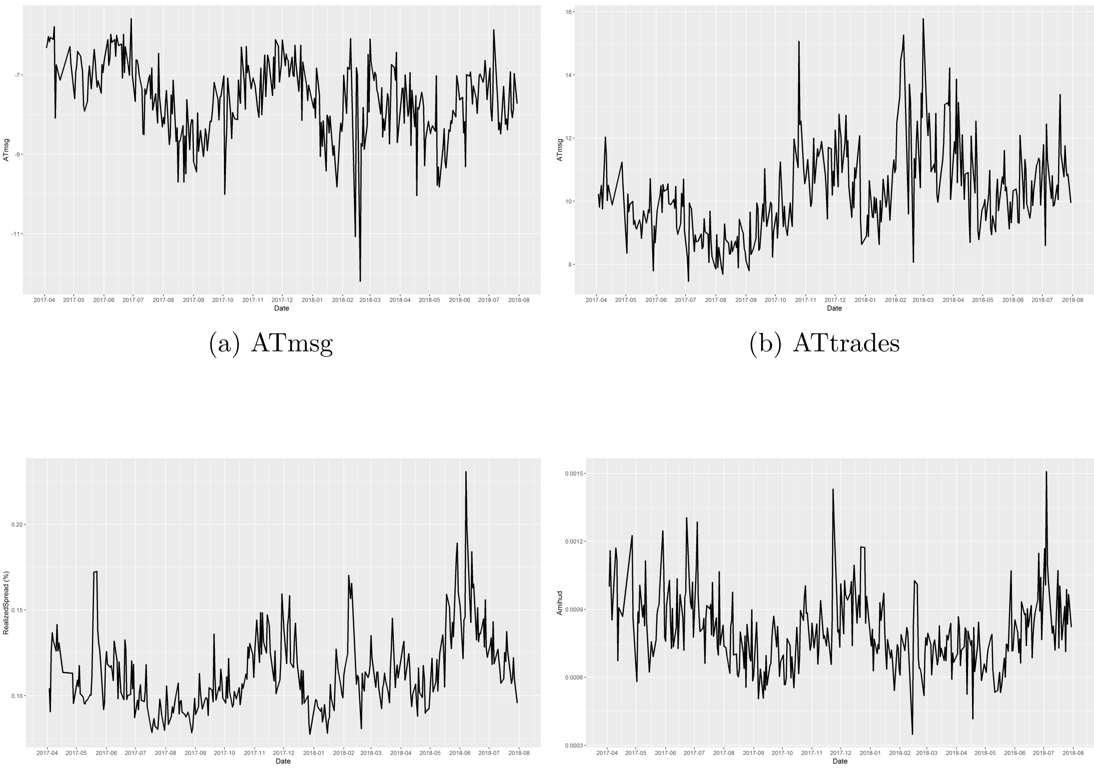
Figure 7 shows the daily equally-weighted averages of the main variables. Panels a), b), c) and d) report $ATmsg$, $ATtrades$, *RealizedSpread* and *Amihud*, respectively. One can note that $ATmsg$ is highly volatile during the sample, therefore not presenting trending behavior as well as $ATtrades$. *RealizedSpread* presents spikes at the end of the sample (as in between February and March of 2018 and between June and July of 2018). The high-frequency Amihud (2002) measure also presents an increase during June and July of 2018[32].

The literature shows that many other variables may affect liquidity and AT activity. Therefore, we include in this study several control variables in order to isolate the effects of AT on liquidity as in Moriyasu et al. (2018), Malceniece et al. (2019) and many others (BOEHMER et al., 2018; HENDERSHOTT et al., 2011; MESTEL et al., 2018):

- $MRet_{itd}$ is the market return proxy for the equally-weighted return for the 26 stocks. This variable is calculated excluding stock i for each estimate at minute t of day d.
- $Invprice_{itd}$ is the inverse of closing price of stock i at minute t of day d.

[32] Unit root tests were conducted for these variables. The results suggests no existence of unit roots and are available upon request.

Figure 7 – Time series of ATmsg, ATtrades, Realized Spreads and high-frequency Amihud (2002)



(a) ATmsg (b) ATtrades

(c) RealizedSpread (%) (d) Amihud

The figure shows ATmsg (a), ATtrades (b), RealizedSpread (c) and high-frequency Amihud (2002) (d) over the sample period. Daily averages are calculated across the 26 selected stocks for the period between 2017-04-03 and 2018-07-30.

- Vol_{itd} is the realized volatility, measured by the squared 1-minute return for stock i at minute t of day d.
- $MVol_{itd}$ is market the volatility calculated through the mean Vol. This variable is calculated excluding stock i for each estimate at minute t of day d[33].

Table 20 reports descriptive statistics for all variables. The average (negative) number of messages per R\$ 100.00 ($ATmsg$) amounts to -8.031 with a median value of -7.595. The mean number of trades per message ($ATtrades$) is 10.673, while both proxies for AT present a standard deviation (Sd) close to 2.5 (2.44 and 2.318 for $ATmsg$ and $ATtrades$, respectively). The number of messages per minute (Msg) is reported with a mean (301.431 per minute) higher than its median (275). The difference between the first (211.318) and the third (361.435) quartiles of Msg suggests significant differences among the stocks in our sample. *RealizedSpread* presents mean value of 0.129% with a standard deviation of 0.048%. The small variability of spreads may be explained through the sampling process, as only stocks with high volume are included in the

[33] For the sake of brevity we have omitted the subscripts *itd* when referring to the control variables and other variables previously defined. We show subscripts whenever necessary.

sample. Although it may raise questions of a liquidity bias, Brazilian stocks overall present small liquidity. Therefore, including thinly traded stocks may severely bias the results. The *Amihud* measure is reported with a mean (0.091) higher than its median (0.079).

Table 20 – Summary statistics

The table reports the mean, median, first quartile (Q1), third quartile (Q3) and standard deviation (Sd) for the algorithmic trading variables, liquidity proxies and control variables for the 26 selected stocks. ATmsg and ATtrades are proxies for algorithmic trading, Msg is the sum of messages (new order submissions, changes, cancels and trades). RealizedSpread and Amihud are liquidity proxies. Control variables include MRet as the average return of stocks available, Invprice as the inverse of last traded price, Vol as the squared return and MVol as the average Vol. Sample period spans from 2017-04-03 to 2018-07-30.

	Mean	Median	Q1	Q3	Sd
ATmsg	-8.031	-7.595	-9.126	-6.407	2.440
ATtrades	10.673	10.318	9.031	11.909	2.318
Msg	301.431	275	211.318	361.435	130.985
RealizedSpread	0.129	0.118	0.094	0.152	0.048
Amihud	0.091	0.079	0.055	0.113	0.053
MRet	0.000	0.000	-0.019	0.019	0.041
Invprice	0.051	0.05	0.047	0.055	0.005
Vol	1.004e-04	7.033e-05	4.451e-05	1.186e-04	1.008e-04
MVol	1.696e-05	3.469e-06	7.039e-07	1.332e-05	5.076e-05

Table 21 presents time series correlations of our variables. All absolute values higher than 0.5 are boldfaced. One can note the high correlation between Msg and *RealizedSprad* (0.835), suggesting that high message traffic may induce higher spreads. However, there is a negative correlation between Msg and *Amihud* (-0.33). Msg is also highly correlated with both Vol (0.791) and $MVol$ (0.81). One can note that the proxies for AT activity are positively correlated to the liquidity variables. Evidence in the literature suggests endogeneity in this relationship, as higher algorithmic trading activity may affect spreads or, higher spreads may repel AT activity. The proxies for volatility are also highly correlated to *RealizedSpreads* (0.972 and 0.933 for Vol and $MVol$, respectively).

Table 21 – Time series correlations

The table presents time series correlations for algorithmic trading variables, liquidity proxies and control variables for the 26 selected stocks. ATmsg and ATtrades are proxies for algorithmic trading, Msg is the sum of messages (new order submissions, changes, cancels and trades). RealizedSpread and Amihud are liquidity proxies. Control variables include MRet as the average return of stocks available, Invprice as the inverse of last traded price, Vol as the squared return and MVol as the average Vol. Absolute values higher than 0.5 are boldfaced. Sample period spans from 2017-04-03 to 2018-07-30.

	ATmsg	ATtrades	Msg	RealizedSpread	Amihud	MRet	Invprice	Vol	MVol
ATmsg	1								
ATtrades	0.279	1							
Msg	0.002	0.45	1						
RealizedSpread	0.246	0.421	**0.835**	1					
Amihud	**0.563**	0.113	-0.33	0.07	1				
MRet	-0.022	-0.135	-0.067	-0.08	0.031	1			
Invprice	**0.595**	-0.273	-0.112	0.186	**0.528**	0.038	1		
Vol	0.228	0.285	**0.791**	**0.972**	0.079	-0.047	0.263	1	
MVol	0.159	0.372	**0.81**	**0.933**	0.026	-0.015	0.111	**0.939**	1

Liquidity is highly subject to noise given intraday seasonality and other microstructure effects. Thus, we adjust liquidity for both weekday and trading hour seasonality based on the procedure of Hameed et al. (2010) and Moriyasu et al. (2018):

$$LIQ^k_{itd} = \alpha_0 + \sum_j^4 d_j WDay + \sum_j^7 h_j Hour + Res\_liq_{itd}, \tag{2.1.5}$$

where LIQ^k_{itd} refers to one of the two liquidity measures, $WDay$ are weekday dummies and $Hour$ trading hour dummies. The residual of the OLS regressions estimated for each stock, Res_liq_{it}, are used for the estimations as liquidity proxies.

In order to assess the effect of AT activity in liquidity, it is necessary to use an exogenous shock of AT as the relationship between these two variables is potentially endogenous. On November 13 of 2017, the B3 exchange started to operate the PUMA trading system from the new data center located in Santana de Parnaíba, state of São Paulo[34]. According to B3, the new facility offers faster connectivity to co-location. Therefore, we use the starting date of operation of the new data center as an exogenous shock. In order to test the validity of the instrument, we estimate the following equation:

$$\begin{aligned} AT_{itd} =& \beta_0 + \beta_1 Dummy_t d + \beta_2 \overline{AT}_{-itd} + \beta_3 MRet_{-itd} + \\ & \beta_4 InvPrice_{itd} + \beta_5 Vol_{itd} + \beta_6 MVol_{-itd} + \epsilon_{itd}, \end{aligned} \tag{2.1.6}$$

where AT_{it} refers to one of the two proxies, $ATmsg$ or $ATtrades$, and $Dummy_{td}$ is the dummy that takes the value of 1 if the observation is on or after November 13 of 2017 and 0 otherwise. As in Mestel et al. (2018), we include $\overline{AT}_{-itd}$ as the average of the AT proxy for the assets available at minute t excluding stock i as an additional instrumental variable and the set of control variables. All the variables used on estimations henceforth are normalized.

The results of Equation 2.1.6 are reported in Table 22. The coefficients of the $Dummy$ variable are positive and statistically significant for both $ATmsg$ (0.007^{***}) and $ATtrades$ (0.047^{***}), providing evidence that AT activity has increased after start of the operation of the new B3 data center. $\overline{AT}^{msg}$ and $\overline{AT}^{trades}$ are reported with positive coefficients (0.117^{***} and 0.170^{***}, respectively), evidencing that overall AT activity may be correlated among stocks. This is in line with the idea of correlated trading activity and the fact that algorithmic traders usually have similar strategies (BENOS et al., 2017; BOEHMER et al., 2015). Market returns $MRet$ have a small positive effect on both proxies of AT activity (0.002^{***} and 0.003^{***}, respectively), which may be expected since the euphoria of positive returns may attract AT/HFT in seeking opportunities. $Invprice$ and the Vol proxy presented distinct signals of coefficients for both proxies. As the results in Table 22 report the new data center dummy as a valid instrument, we use the fitted values of Equation 2.1.6 for our estimations.

[34] Source: B3's Circular Letter 070/2017-DP.

Table 22 – First stage regression of the impact of new data center on AT

The table reports the impact of the beginning of new B3 data center operations on AT activity. The estimated regression is: $AT_{it} = \beta_0 + \beta_1 Dummy_t + \beta_2 \overline{AT}_{-it} + \beta_3 MRet_{-it} + \beta_4 InvPrice_{it} + \beta_5 Vol_{it} + \beta_6 MVolat_{-it} + \epsilon_{it}$. AT is one of the two AT proxies ($ATmsg$ and $ATtrades$), $DCenter$ is the dummy variable that takes value of one when trading started to occur in the new data center of B3. $\overline{AT_{-i}}$ is the average market AT excluding stock i. $MRet$ is the average return of stocks available, $Invprice$ is the inverse of last traded price, Vol is the squared return and $MVol$ is the average Vol. All variables are standardized before estimations. Newey e West (1986) standard errors are reported in parenthesis. Time aggregation is set to 1 minute. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	$ATmsg$	$ATtrades$
	(1)	(2)
Dummy	0.007***	0.047***
	(0.001)	(0.001)
$\overline{AT}$	0.117***	0.170***
	(0.001)	(0.001)
$MRet$	0.002***	0.003***
	(0.001)	(0.001)
$Invprice$	0.190***	−0.163***
	(0.002)	(0.002)
Vol	−0.051***	0.002***
	(0.001)	(0.001)
$MVol$	−0.009***	−0.0004
	(0.001)	(0.001)
Observations	2,347,706	2,347,706
R^2	0.033	0.050
Adjusted R^2	0.033	0.050
F Statistic (df = 6; 2347674)	13,399.670***	20,396.960***

3 Liquidity and algorithmic trading

In order to assess the effect of AT activity on liquidity, we estimate Equation 3.0.1 using the residual from 2.1.5 and the fitted values based on Equation 2.1.6:

$$Res\_liq_{itd}^{k} = \alpha_0 + \alpha_1 \widetilde{AT}_{itd} + \alpha_2 MRet_{-itd} + \alpha_3 InvPrice_{itd} + \alpha_4 Vol_{itd} + \alpha_5 MVol_{-itd} + \epsilon_{itd}, \tag{3.0.1}$$

where $Res\_liq_{itd}^{k}$ refers to the residual liquidity for each one of the k measures ($RealizedSpread$ and $Amihud$), $\widetilde{AT}_{itd}$ refers to fitted values for one of the AT proxies estimated in Equation 2.1.6. Individual fixed effects are included in the estimations.

Table 23 reports the results for Equation 3.0.1 when time aggregation is set to 1 minute. Columns (1) and (2) show the results using $RealizedSpread$ as the dependent variable. Both AT proxies present positive and significant coefficients of the effect on $RealizedSpread$ (0.054*** in $ATmsg$ and 0.095*** in $ATtrades$), evidencing that AT activity harms market quality in the form of spreads. Columns (3) and (4) show the results with $Amihud$ as the dependent variable. Both $ATmsg$ and $ATtrades$ present positive and significant coefficients (0.018*** and 0.023***, respectively), evidencing that AT activity increases high-frequency price impact. Our evidence is contrary to the majority of work relating AT activity to an increase in liquidity. Most of the documented benefits of AT/HFT come from voluntary market making, although many studies do not explicitly exclude stocks with market makers. Therefore, the effects of AT/HFT may be biased as designated market makers may use algorithms to meet their obligations. As our sample excludes stocks with designated market makers, we are able to evidence an unbiased effect of AT on liquidity.

The selected control variables present the expected coefficient signals in most of the cases. Market returns ($MRet$) have negative coefficients for $RealizedSpread$ (-0.001^{***} and -0.002^{***}), which is expected as when returns diminish, an increase in spreads is expected (HAMEED et al., 2010). Coefficients of this variable have the opposite sign when $Amihud$ is the dependent variable (0.007***). $Invprice$ has a positive effect for both Columns (1) and (2) using $RealizedSpread$ as dependent variable and for Columns (3) and (4) when $Amihud$ is used. This suggests that cheaper stocks present higher illiquidity. Stocks' realized volatility (Vol) is reported with positive and significant coefficients. The effect of market volatility ($MVol$) is positive when $RealizedSpread$ is the dependent variable and negative when $Amihud$ is used. Adjusted R-squared estimates for both $RealizedSpread$ and $Amihud$ (close to 16% and 2%, respectively) evidence a higher explicative power for spreads instead of price impact. This is consistent with Mestel et al. (2018), who find a small effect of AT on price impact for the Austrian market.

Table 23 – Effect of algorithmic trading on stock liquidity

The table reports the effect of two AT proxies in stock liquidity. The estimated regression model is: $Res\_liq^k_{itd} = \alpha_0 + \alpha_1 \widetilde{AT}_{itd} + \alpha_2 MRet_{-itd} + \alpha_3 InvPrice_{itd} + \alpha_4 Vol_{itd} + \alpha_5 MVolat_{-itd} + \epsilon_{itd}$, where Res_liq^k is one of the k adjusted liquidity proxies (*RealizedSpread* and *Amihud*) from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 (*ATmsg* and *ATtrades*). *MRet* is the average return of stocks available, Invprice is the inverse of last traded price, Vol is the squared return and MVol is the average Vol. All variables are standardized before estimations. Newey e West (1986) standard errors are reported in parenthesis. Time aggregation is set to 1 minute and individual fixed effects are included. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	RealizedSpread		*Amihud*	
	(1)	(2)	(3)	(4)
ATmsg	0.054***		0.018***	
	(0.002)		(0.002)	
ATtrades		0.095***		0.023***
		(0.001)		(0.002)
MRet	−0.001**	−0.002***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Invprice	0.123***	0.177***	0.438***	0.454***
	(0.002)	(0.002)	(0.003)	(0.003)
Vol	0.382***	0.376***	0.038***	0.036***
	(0.001)	(0.001)	(0.001)	(0.001)
MVol	0.079***	0.075***	−0.044***	−0.045***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	2,347,706	2,347,706	1,670,635	1,670,635
R^2	0.163	0.164	0.018	0.018
Adjusted R^2	0.163	0.164	0.018	0.018
F Statistic	91,316.940*** (df = 5; 2347675)	92,140.180*** (df = 5; 2347675)	6,230.839*** (df = 5; 1670604)	6,247.329*** (df = 5; 1670604)

3.1 Vector autoregression estimates

The majority of studies analyzing the relationship between AT/HFT and liquidity have focused on the effect of the former on the latter. However, little attention has been given to the effect of liquidity on AT/HFT activity. Lee (2015) investigates the connection between market quality and HFT activity in the Korean futures market through vector autoregressions (VAR), finding a small effect of liquidity on HFT activity. Contrary to the majority of literature, the author finds that HFT do not improve market quality, but worsens the price discovery process.

In order to test the bidirectional relationship of AT and liquidity, we run VAR estimations on the previously defined variables. Equations 3.1.1 and 3.1.2 describe the restricted VAR model[35]:

[35] Chaboud et al. (2014) estimate similar VAR equations on AT activity and arbitrage opportunities in the foreign exchange market.

$$Res\_liq_{itd}^{k} = \alpha_0 + \sum_{p=1}^{P} \alpha_p Res\_liq_{it-pd}^{k} + \sum_{p=1}^{P} \gamma_p \widetilde{AT}_{it-pd} + \zeta X + \upsilon_{itd} \tag{3.1.1}$$

$$\widetilde{AT}_{itd} = \beta_0 + \sum_{p=1}^{P} \beta_p \widetilde{AT}_{it-pd} + \sum_{p=1}^{P} \delta_p Res\_liq_{it-pd}^{j} + \zeta X + e_{itd} \tag{3.1.2}$$

where $Res\_liq_{itd}^{k}$ refers to each one of the k liquidity measures ($RealizedSpread$ and $Amihud$), $\widetilde{AT}_{itd}$ refers to one of the AT activity proxies for asset i at time t of day d. X refers to the set of control variables: $MRet$ is the average return of stocks available, $Invprice$ is the inverse of last traded price, Vol is the squared return and $MVol$ is the average Vol. The number of p lags is set to 20 for the endogenous variables. The first 20 observations for each day are set to zero so variables in the previous days will not affect current ones. This is plausible since the effect of AT activity yesterday should not affect liquidity today. Data for each asset is pooled through the sample and all variables are standardized. We also perform causality tests as in Granger (1969).

Table 24 shows the average results for Equations 3.1.1 and 3.1.2 when $RealizedSpread$ is the liquidity variable. Columns (1) and (2) report the estimates when $ATmsg$ is the AT variable. The AT proxy is positively affected by $RealizedSpread$ (average sum of lagged coefficients: 0.0673^{***}) and highly affected by its lagged values (0.8359^{***}). This evidences that AT activity is positively affected by increases in spreads, corroborating with Hendershott e Riordan (2013), who show that AT activity is lower when spreads are narrow. The effect of lagged $ATmsg$ on $RealizedSpread$ is positive and significant (0.4139^{***}), showing that the effect of AT on spreads is not only contemporaneous, but also lagged. Most of the literature explores the contemporaneous effect of these variables. $RealizedSpread$ presents positive autocorrelation as the sum of its lagged coefficients (0.5595^{***}).

Columns (3) and (4) of Table 24 present the results when $ATtrades$ is the AT proxy. Lagged $RealizedSpread$ also positively affects $ATtrades$, even though the average sum of lagged coefficients is smaller (0.0201^{**}) than $ATmsg$. The coefficients of the effect of AT on liquidity are also smaller, with a sum of lagged coefficients of 0.0729^{***}. We show evidence of the bidirectional relationship between AT and $RealizedSpread$, confirming that AT activity is attracted by higher spreads (HENDERSHOTT et al., 2011).

Figure 8 depicts the coefficients estimated for Equations 3.1.1 and 3.1.2 for each stock in our sample. Panel a) reports the sum of lagged coefficients of $ATmsg$ on $RealizedSpread$ (γ_p coefficients of Equation 3.1.1). Only one asset has a negative coefficient and only three assets do not present significance for the Granger Causality tests at the 5% level. This confirms the evidence in Table 24 of lagged $ATmsg$ affecting $RealizedSpread$. Panel b) exhibits the sum of δ_p coefficients of Equation 3.1.2. The effect of lagged values of $RealizedSpread$ on $ATmsg$ is positive for all 26 stocks within significance in the Granger Causality tests. Panel c) presents the sum of γ_p coefficients when $ATtrades$ is the AT proxy. Although the majority of stocks present positive coefficients (23 out of 26), 9 stocks (around 35% of the sample) lack significance in the Granger Causality tests. Panel d) reports the effect of lagged $RealizedSpread$ on $ATtrades$. As

Table 24 – Summary of VAR estimates - AT and RealizedSpread

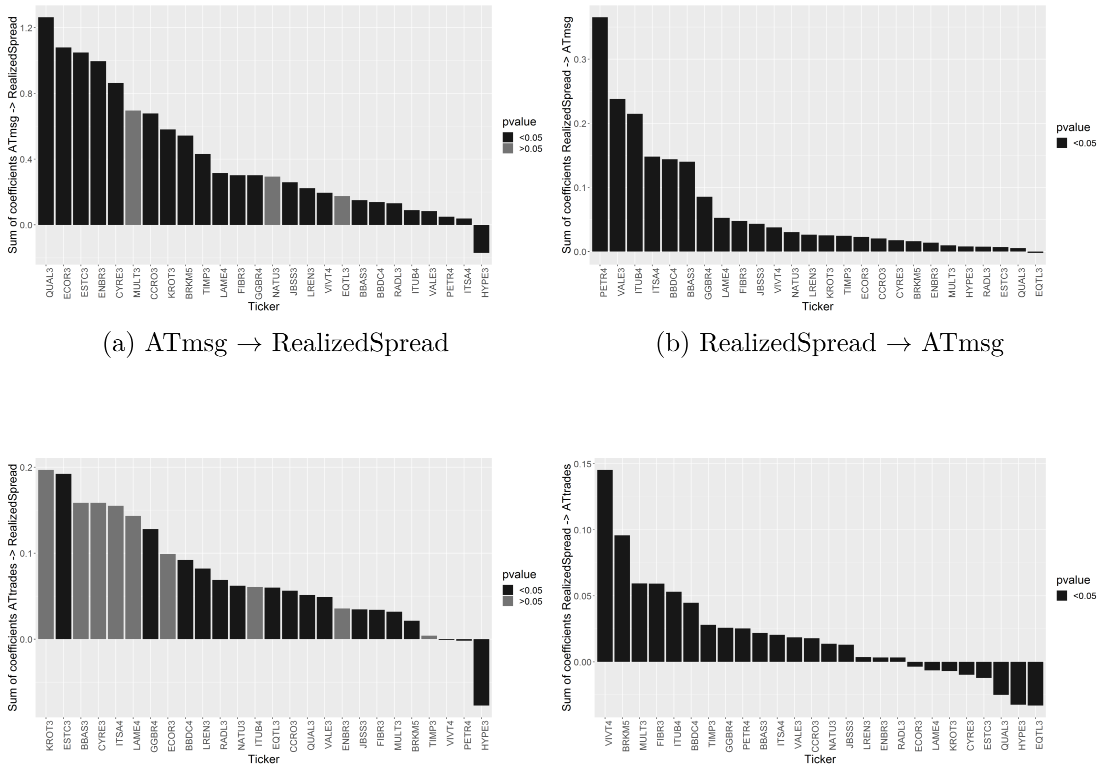
The table reports the average results for vector autoregression estimates in the model: $Res\_liq^k_{itd} = \alpha_0 + \sum_{p=1}^{P} \alpha_p Res\_liq^k_{it-pd} + \sum_{p=1}^{P} \gamma_p \widetilde{AT}_{it-pd} + \zeta X + v_{itd}$ (equation 3.1.1) and $\widetilde{AT}_{itd} = \beta_0 + \sum_{p=1}^{P} \beta_p \widetilde{AT}_{it-pd} + \sum_{p=1}^{P} \delta_p Res\_liq^j_{it-pd} + \zeta X + e_{itd}$ (equation 3.1.2). Res_liq^k is the adjusted *RealizedSpread* from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 (*ATmsg* and *ATtrades*). X is a set of control varbiables including *MRet* as the average return of stocks available, *Invprice* as the inverse of last traded price, *Vol* as the squared return and *MVol* as the average *Vol*. The number of p lags is set to 20 and the first 20 observations for each stock i for each day d are set to zero. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	ATmsg	RealizedSpread	ATtrades	RealizedSpread
	(1)	(2)	(3)	(4)
$\sum AT_{t-p}$	0.8359***	0.4139***	0.9234***	0.0729***
	(0.0159)	(0.0741)	(0.0061)	(0.0129)
$\sum RealizedSpread_{t-p}$	0.0673***	0.5595***	0.0201**	0.5489***
	(0.0176)	(0.0082)	(0.0077)	(0.0082)
$MRet$	0.0057***	-0.0023**	0.0125***	-0.0021**
	(0.0015)	(0.001)	(0.0018)	(0.001)
$Invprice$	0.1065**	-0.0608*	-0.0941**	0.1005***
	(0.0418)	(0.0325)	(0.0395)	(0.0303)
Vol	-0.1191***	0.3351***	-0.0017	0.3325***
	(0.0283)	(0.0134)	(0.0125)	(0.0133)
$MVol$	-0.0169**	0.0411***	-0.0035	0.0403***
	(0.0075)	(0.0019)	(0.0041)	(0.0019)
Average adj.R^2	0.7463	0.2471	0.6912	0.2462
Average nObs	90276	90276	90276	90276

in Panel b), all coefficients are significant in causality tests, although 8 stocks have a negative sum of coefficients.

Table 25 reports the average results for Equations 3.1.1 and 3.1.2 when *Amihud* is the liquidity variable. Columns (1) and (2) report the estimates when *ATmsg* is the AT variable. The AT proxy is negatively affected by *Amihud* (average sum of lagged coefficients: -0.0051), although the result lacks significance. The effect of lagged *ATmsg* on *Amihud* is positive and significant (0.4534***), evidencing the effect of AT on price impact in a dynamic setting. *Amihud* presents positive autocorrelation as the sum of its lagged coefficients is 0.6191***. Columns (3) and (4) of Table 25 present the results when *ATtrades* is the AT proxy. Lagged *Amihud* does not affect *ATtrades* in a significant manner (-0.002). Also, time series estimations show no significant effects of *ATtrades* on *Amihud*, as the average sum of lagged coefficients is small and not significant (0.0031).

Figure 8 – VAR estimates - AT and $RealizedSpread$



(a) ATmsg → RealizedSpread (b) RealizedSpread → ATmsg

(c) ATtrades → RealizedSpread (d) RealizedSpread → ATtrades

The figure reports the sum of coefficients estimated through Equations 3.1.1 and 3.1.2 per asset. Panel a) reports the effect of lagged $ATmsg$ on $RealizedSpreads$ and Panel b) reports the effect of lagged $RealizedSpreads$ on $ATmsg$. Panel c) reports the effect of lagged $ATtrades$ on $RealizedSpreads$ and Panel d) reports the effect of lagged $RealizedSpreads$ on $ATtrades$. The sample includes 26 stocks listed at the B3 exchange from 2017-04-03 to 2018-07-30.

Figure 9 details the estimations of Equations 3.1.1 and 3.1.2 for each stock in our sample. Panel a) reports the sum of lagged coefficients of $ATmsg$ on $Amihud$ (γ_p coefficients of Equation 3.1.1). Similar to Panel a) of Figure 8, only one asset has a negative coefficient. However, the effects of γ_p lack significance in causality tests for 8 out of 26 assets (near 30%). Panel b) exhibits the sum of δ_p coefficients of Equation 3.1.2. The effect of lagged values of $Amihud$ on $ATmsg$ is positive for 22 stocks, within significance in most of the Granger Causality tests. Curiously, one asset has presented a large negative coefficient, suggesting that the effects studied in this paper may not be unique for each asset. Panel c) presents the sum of γ_p coefficients when $ATtrades$ is the AT proxy. 14 stocks present positive values, although significance on the causality tests is much smaller (only 6 stocks present positive coefficients and significant values at the 5% level). Therefore, there is non conclusive evidence of the effect of $ATtrades$ on $Amihud$. Panel d) reports the effect of lagged $Amihud$ on $ATtrades$. P-values are significant for 22 stocks in our sample, suggesting that liquidity plays a role in AT activity.

Table 25 – Summary of VAR estimates - AT and Amihud

The table reports the average results for vector autoregression estimates in the model: $Res\_liq^k_{itd} = \alpha_0 + \sum_{p=1}^{P} \alpha_p Res\_liq^k_{it-pd} + \sum_{p=1}^{P} \gamma_p \widetilde{AT}_{it-pd} + \zeta X + v_{itd}$ (equation 3.1.1) and $\widetilde{AT}_{itd} = \beta_0 + \sum_{p=1}^{P} \beta_p \widetilde{AT}_{it-pd} + \sum_{p=1}^{P} \delta_p Res\_liq^j_{it-pd} + \zeta X + e_{itd}$ (equation 3.1.2). Res_liq^k is the adjusted *Amihud* from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 ($ATmsg$ and $ATtrades$). X is a set of control varbiables including $MRet$ as the average return of stocks available, $Invprice$ as the inverse of last traded price, Vol as the squared return and $MVol$ as the average Vol. The number of p lags is set to 20 and the first 20 observations for each stock i for each day d are set to zero. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

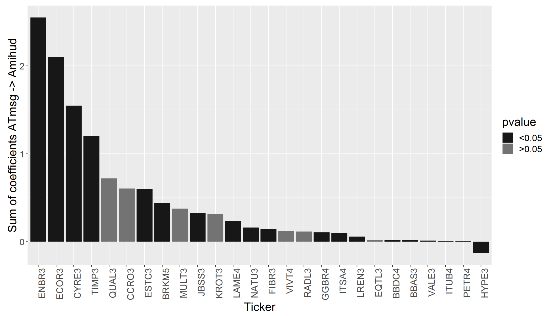
	ATmsg	Amihud	ATtrades	Amihud
$\sum AT_{t-p}$	0.8202***	0.4534***	0.9247***	0.0031
	(0.0194)	(0.1323)	(0.006)	(0.0085)
$\sum Amihud_{t-p}$	-0.0051	0.6191***	-0.002	0.6227***
	(0.0264)	(0.0097)	(0.0064)	(0.01)
$MRet$	0.0059***	0.0066***	0.0126***	0.0067***
	(0.0015)	(0.001)	(0.0018)	(0.001)
$Invprice$	0.122***	0.004	-0.0901**	0.0734***
	(0.0425)	(0.0188)	(0.0402)	(0.0217)
Vol	-0.1115***	0.1631***	0.0028	0.1609***
	(0.026)	(0.0259)	(0.0117)	(0.0255)
$MVol$	-0.0136*	-0.0346***	-0.0025	-0.0351***
	(0.0067)	(0.0064)	(0.0039)	(0.0066)
Average adj R^2	0.7367	0.2471	0.6896	0.1219
Average nObs	90276	90276	90276	90276

3.2 Commonality in liquidity and AT

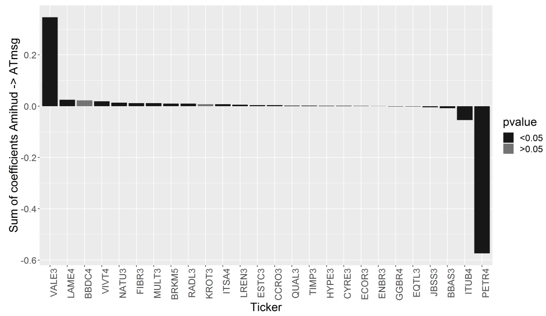
Although individual liquidity of assets is relevant for investors, one can notice this variable is correlated between assets and with the market as a whole. Chordia et al. (2000) were the first to document commonality in liquidity (hereafter, CIL), evidencing that liquidity in stocks comove with market liquidity. A wide literature shows that CIL demands a return premium, as stocks with higher sensibility to market liquidity compensate this risk with higher returns (ACHARYA; PEDERSEN, 2005; ANTHONISZ; PUTNIŅŠ, 2016; KAROLYI et al., 2012). As AT/HFT tend to rely on momentum and market-making strategies, it may be expected that these agents present correlated trading activity, therefore inducing CIL. Boehmer et al. (2015) show evidence of at least three main strategies that HFTs engage in the Canadian market, and Benos et al. (2017) also provide evidence of correlated HFT activity in the U.K. market. Chaboud et al. (2014) report correlated AT strategies in the FX market.

Moriyasu et al. (2018) find evidence of AT activity increasing CIL in the Tokyo Stock

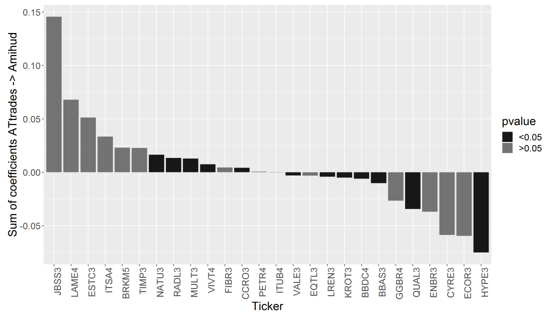
Figure 9 – VAR estimates - AT and *Amihud*



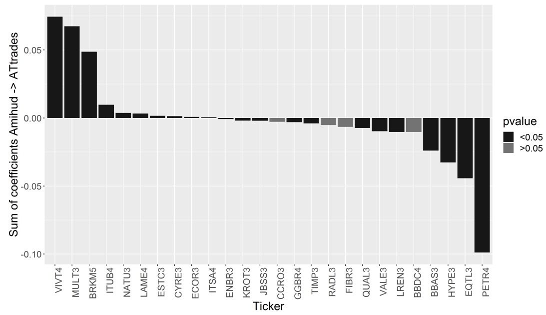
(a) ATmsg → Amihud



(b) Amihud → ATmsg



(c) ATtrades → Amihud



(d) Amihud → ATtrades

The figure reports the sum of coefficients estimated through Equations 3.1.1 and 3.1.2 per asset. Panel a) reports the effect of lagged $ATmsg$ on $Amihud$ and Panel b) reports the effect of lagged $Amihud$ on $ATmsg$. Panel c) reports the effect of lagged $ATtrades$ on $Amihud$ and Panel d) reports the effect of lagged $Amihud$ on $ATtrades$. The sample includes 26 stocks listed at the B3 exchange from 2017-04-03 to 2018-07-30.

Exchange, in which the effect is higher in market drops. In the European market, Malceniece et al. (2019) report that HFT amplifies CIL and commonality in returns. Anagnostidis e Fontaine (2018) and Jain et al. (2016) also provide evidence of the effects of AT/HFT on systemic sources of risk. As commonality is a source of systemic risk, AT/HFT activity may have a direct impact on a firm's cost of capital. Although research on AT/HFT is extensive in developed markets, little attention has been given to emerging ones, especially in the relationship between AT/HFT activity and CIL[36]. In this section, we test the effect of AT on CIL in the Brazilian equities market.

Our approach is similar to that of Moriyasu et al. (2018) and Malceniece et al. (2019). We estimate CIL based on a regression of residual liquidity on market liquidity. We include lead and lag observations for market liquidity in order to capture any adjustments in CIL (CHORDIA et al., 2000). Equation 3.2.1 details the estimation:

[36] Lee (2015) studies the effect of HFT on liquidity in the South Korean futures market. Jawed e Chakrabarti (2018) study the entry of HFTs on the Indian market. Boehmer et al. (2018) provide worldwide evidence of the effect of AT on market quality (including emerging markets).

$$\Delta Res\_liq_{itd}^{k} = \alpha_0 + \sum_{j=-1}^{1} \alpha_j \Delta \overline{Res\_liq}_{-it+jd}^{k} + \epsilon_{itd}, \tag{3.2.1}$$

where ΔRes_liq^k is the 1-minute variation of residual liquidity from Equation 2.1.5 for the k liquidity measure, $\Delta \overline{Res\_liq}_{-it+jd}^{k}$ is the variation of average residual liquidity excluding stock i. We run the above regression for and retrieve the R-squared for each stock and each d day in our sample. As a measure of CIL, we calculate the logit transformation as in Karolyi et al. (2012) and Hameed et al. (2010):

$$Commo_{id}^{k} = ln\Big(\frac{R^2}{1-R^2}\Big). \tag{3.2.2}$$

Therefore, $Commo_{id}^{k}$ is the transformed R-squared of Equation 3.2.1 for stock i at day d. This measure is set as a dependent variable in a fixed-effects panel regression against our proxies for AT activity and a set of control variables:

$$Commo_{id}^{k} = \alpha_0 + \alpha_1 \widetilde{AT}_{id} + \alpha_2 Vol_{id} + \alpha_3 InvPrice_{id} + \alpha_4 MRet_{-id} + \alpha_5 MVol_{-id} + \epsilon_{id}. \tag{3.2.3}$$

Our procedure differs from the one used in Moriyasu et al. (2018) and Malceniece et al. (2019), where the estimates of Equation 3.2.2 are calculated for each month using daily observations. We believe calculating commonality using high-frequency data may present a more reliable picture of the relationship between liquidity and AT, especially as the subject of study (AT) is highly time-sensitive. Even though high-frequency data may present microstructure noise, monthly estimates may severely bias our results.

Table 26 presents the results from Equation 3.2.3. Columns (1) and (2) report the estimates when $Commo^{RSpread}$ is the dependent variable. The impact of both AT proxies on CIL is positive and statistically significant (coefficients 0.136*** and 0.274***, respectively). When $Commo^{Amihud}$ is the dependent variable (Columns (3) and (4)), only the coefficient of $ATmsg$ is statistically significant (0.174***). Therefore, the effect of AT is more pronounced within $RealizedSpread$. The main channel through which AT should affect CIL is correlated strategies, as trading patterns may be replicated over several assets (BENOS et al., 2017; BROGAARD et al., 2014).

Control variables such as $MRet$ have a negative effect on CIL when $RealizedSpread$ is the liquidity measure, confirming that commonality should be higher in down markets (ANTHONISZ; PUTNIŅŠ, 2016), although the effect is not statistically significant for $Commo^{Amihud}$. $MVol$ has positive and significant coefficients for all estimations, implying that higher volatility entails higher CIL, which is also expected as risk aversion may cause flights to quality (BRUNNERMEIER; PEDERSEN, 2009; ACHARYA; PEDERSEN, 2005).

Table 26 – Effect of algorithmic trading on commonality in liquidity

The table reports the effect of two AT proxies on intraday commonality in liquidity (CIL). For each trading day we run the following OLS regression: $\Delta Res\_liq^k_{itd} = \alpha_0 + \sum_{j=-1}^{1} \alpha_j \Delta \overline{Res\_liq}^k_{-it+jd} + \epsilon_{itd}$, where Res_liq^k is one of the two adjusted liquidity measures, $\Delta\overline{Res\_liq}_{-i}$ is the average liquidity of stocks available excluding stock i. Contemporaneous, lead and lag relations are included. For each day and each stock, we retrieve the R-squared. Our measure of CIL is the logit transformed R-squared in the form of: $Commo^k_{id} = ln\left(\frac{R^2}{1-R^2}\right)$. We estimate the following equation in order to assess the effect of AT on CIL: $Commo^k_{id} = \alpha_0 + \alpha_1 \widetilde{AT}_{id} + \alpha_2 Vol_{id} + \alpha_3 InvPrice_{id} + \alpha_4 MRet_{-id} + \alpha_5 MVol_{-id} + \epsilon_{id}$. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 ($ATmsg$ and $ATtrades$). $MRet$ is the average return of stocks available, Invprice is the inverse of last traded price, Vol is the squared return and MVol is the average Vol. All variables are standardized before estimations. Newey e West (1986) standard errors are reported in parenthesis. Individual fixed effects are included. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	Dependent variable:			
	$Commo^{RSpread}$		$Commo^{Amihud}$	
	(1)	(2)	(3)	(4)
$ATmsg$	0.136***		0.174***	
	(0.034)		(0.040)	
$ATtrades$		0.274***		0.027
		(0.037)		(0.037)
$MRet$	−0.047***	−0.049***	−0.015	−0.018
	(0.018)	(0.018)	(0.017)	(0.017)
$Invprice$	−0.245***	−0.087**	−0.111**	−0.035
	(0.039)	(0.038)	(0.044)	(0.044)
Vol	−0.0001	0.006	0.027**	0.028**
	(0.013)	(0.013)	(0.012)	(0.012)
$MVol$	0.309***	0.289***	0.105***	0.103***
	(0.022)	(0.022)	(0.014)	(0.014)
Observations	8,294	8,294	8,294	8,294
R^2	0.089	0.095	0.011	0.009
Adjusted R^2	0.085	0.091	0.007	0.005
F Statistic (df = 5; 8263)	160.483***	172.543***	18.415***	14.936***

3.3 Robustness checks

As robustness checks, we estimate previous equations using different time aggregations and specifications. Table 27 reports the results for equation 3.0.1 when *RealizedSpread* is the liquidity proxy and variables are aggregated at the 5-minute, 15-minute and daily intervals (1-minute averages for each day). All main results from Table 23 hold, including the effect of *ATmsg* and *ATtrades* on *RealizedSpread*. Coefficients range from 0.065*** to 0.199*** for *ATmsg* and from 0.123*** to 0.168*** for *ATtrades*. Table 28 reports the estimates for *Amihud* as the liquidity variable. Both signal and significance of the effects of *ATmsg* and *ATtrades* on *Amihud* are robust to different time frames.

Table 27 – Effect of algorithmic trading on RealizedSpread - Multiple time frames

The table reports the effect of two proxies of AT on *RealizedSpread*. The estimated regression model is: $Res\_liq^k_{itd} = \alpha_0 + \alpha_1 \widetilde{AT}_{itd} + \alpha_2 MRet_{-itd} + \alpha_3 InvPrice_{itd} + \alpha_4 Vol_{itd} + \alpha_5 MVolat_{-it} + \epsilon_{it}$, where Res_liq^k is one of the k adjusted liquidity proxies (*RealizedSpread* and *Amihud*) from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 (*ATmsg* and *ATtrades*). *MRet* is the average return of stocks available, Invprice is the inverse of last traded price, Vol is the squared return and MVol is the average Vol. All variables are standardized before estimations. Columns (1) and (2), (3) and (4), (5) and (6) reports the results for 5-minute, 15-minute and daily time aggregation, respectively. Newey e West (1986) standard errors are reported in parenthesis. Individual fixed effects are included. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	5 min		15 min		Daily	
	(1)	(2)	(3)	(4)	(5)	(6)
ATmsg	0.065***		0.059***		0.199***	
	(0.005)		(0.006)		(0.017)	
ATtrades		0.123***		0.108***		0.168***
		(0.005)		(0.008)		(0.015)
MRet	−0.010***	−0.010***	−0.008***	−0.008***	−0.020***	−0.016***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.005)	(0.005)
Invprice	0.090***	0.152***	0.086***	0.143***	−0.154***	−0.014
	(0.005)	(0.005)	(0.008)	(0.008)	(0.018)	(0.018)
Vol	0.421***	0.414***	0.407***	0.403***	0.962***	0.959***
	(0.006)	(0.001)	(0.002)	(0.002)	(0.007)	(0.007)
MVol	0.073***	0.068***	0.090***	0.086***	0.162***	0.150***
	(0.003)	(0.001)	(0.002)	(0.002)	(0.006)	(0.006)
Observations	493,876	493,876	176,241	176,241	8,294	8,294
R^2	0.194	0.195	0.191	0.192	0.817	0.817
Adjusted R^2	0.194	0.195	0.191	0.191	0.816	0.816
F Statistic	23,784.800*** (df = 5; 493845)	23,888.560*** (df = 5; 493845)	8,321.984*** (df = 5; 176210)	8,348.623*** (df = 5; 176210)	7,375.319*** (df = 5; 8263)	7,355.299*** (df = 5; 8263)

Table 29 summarizes the results of VAR Equations 3.1.1 and 3.1.2 using 5-minute, 15-minute and daily intervals when *RealizedSpread* is the liquidity variable. The number of p lags are set to 4, 2 and 1, respectively. Coefficients on control variables were omitted for the sake of brevity. The main results from Table 24 hold, as both *ATmsg* and *ATtrades* have a positive effect on *RealizedSpread*. Table 30 reports similar results when analyzing the effect of *ATmsg* on *Amihud*, although the effect of *ATtrades* on *Amihud* lacks significance.

In order to check robustness for the effect of AT on CIL, we consider the sum of $\Delta\overline{Res\_liq}^k$

Table 28 – Effect of algorithmic trading on Amihud - Multiple time frames

The table reports the effect of two proxies of AT on *Amihud*. The estimated regression model is: $Res\_liq^k_{itd} = \alpha_0 + \alpha_1 \widetilde{AT}_{itd} + \alpha_2 MRet_{-itd} + \alpha_3 InvPrice_{itd} + \alpha_4 Vol_{itd} + \alpha_5 MVolat_{-it} + \epsilon_{it}$, where Res_liq^k is one of the k adjusted liquidity proxies (*RealizedSpread* and *Amihud*) from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 (*ATmsg* and *ATtrades*). *MRet* is the average return of stocks available, Invprice is the inverse of last traded price, Vol is the squared return and MVol is the average Vol. All variables are standardized before estimations. Columns (1) and (2), (3) and (4), (5) and (6) reports the results for 5-minute, 15-minute and daily time aggregation, respectively. Newey e West (1986) standard errors are reported in parenthesis. Individual fixed effects are included. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	5 min		15 min		Daily	
	(1)	(2)	(3)	(4)	(5)	(6)
ATmsg	0.110***		0.137***		0.128***	
	(0.003)		(0.007)		(0.037)	
ATtrades		0.065***		0.062***		0.134***
		(0.005)		(0.008)		(0.034)
MRet	0.013***	0.013***	0.017***	0.017***	−0.008	−0.004
	(0.002)	(0.001)	(0.002)	(0.002)	(0.010)	(0.010)
Invprice	0.267***	0.330***	0.228***	0.302***	1.007***	1.107***
	(0.007)	(0.006)	(0.009)	(0.009)	(0.039)	(0.039)
Vol	0.209***	0.200***	0.276***	0.268***	−0.057***	−0.059***
	(0.005)	(0.001)	(0.002)	(0.002)	(0.016)	(0.016)
MVol	−0.029***	−0.032***	−0.022***	−0.024***	0.017	0.007
	(0.002)	(0.001)	(0.002)	(0.002)	(0.013)	(0.013)
Observations	420,477	420,477	159,880	159,880	8,294	8,294
R^2	0.054	0.052	0.085	0.083	0.094	0.095
Adjusted R^2	0.054	0.052	0.085	0.083	0.091	0.091
F Statistic	4,766.404*** (df = 5; 420446)	4,658.875*** (df = 5; 420446)	2,986.954*** (df = 5; 159849)	2,910.996*** (df = 5; 159849)	172.151*** (df = 5; 8263)	173.038*** (df = 5; 8263)

coefficients as a dependent variable on Equation 3.2.1 instead of its R-squared[37]. Table 31 summarizes the results. When *RealizedSpread* is defined as the liquidity variable, the main results from Table 26 hold, as both *ATmsg* and *ATtrades* have positive and statistically significant coefficients over $Commo^{RSpread}$. When *Amihud* is used (Columns (3) and (4)), the coefficients lack significance. Thus, our results reflect a more robust effect of AT on *RealizedSpread* rather than in the *Amihud* measures.

[37] We have estimated the mentioned equation using only the contemporaneous coefficient ($t = 1$) as a proxy for CIL. The results remain qualitatively unchanged and are available upon request.

Table 29 – Summary of VAR estimates - AT and RealizedSpread - Multiple time frames

The table reports the average results for vector autoregression estimates in the model: $Res\_liq^k_{itd} = \alpha_0 + \sum_{p=1}^{P} \alpha_p Res\_liq^k_{it-pd} + \sum_{p=1}^{P} \gamma_p \widetilde{AT}_{it-pd} + \zeta X + v_{itd}$ (equation 3.1.1) and $\widetilde{AT}_{itd} = \beta_0 + \sum_{p=1}^{P} \beta_p \widetilde{AT}_{it-pd} + \sum_{p=1}^{P} \delta_p Res\_liq^j_{it-pd} + \zeta X + e_{itd}$ (equation 3.1.2). Res_liq^k is the adjusted *RealizedSpread* from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 ($ATmsg$ and $ATtrades$). X is a set of control varbiables including $MRet$ as the average return of stocks available, $Invprice$ as the inverse of last traded price, Vol as the squared return and $MVol$ as the average Vol. Panels A,B and C reports time aggregation of 5 minutes, 15 minutes and daily observations. The number of p lags is set to 4,2 and 1 for Panels A,B and C, respectively. The first p observations for each stock i for each day d are set to zero with the exception of Panel C. Control variables coefficients were omitted for the sake of brevity. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	ATmsg	RealizedSpread	ATtrades	RealizedSpread
	(1)	(2)	(3)	(4)
Panel A: 5-minute time frame				
$\sum AT_{t-p}$	0.7368***	0.4152***	0.8336***	0.3336***
	(0.0231)	(0.0859)	(0.0215)	(0.0735)
$\sum RealizedSpread_{t-p}$	0.0403***	0.3736***	-0.0135***	0.355***
	(0.0095)	(0.0095)	(0.0031)	(0.0101)
Average.adj.R^2	0.7011	0.2663	0.8473	0.2665
Average.nObs	18991	18991	18991	18991
Panel B: 15-minute time frame				
$\sum AT_{t-p}$	0.6701***	0.2357***	0.7806***	0.294***
	(0.0246)	(0.0625)	(0.0251)	(0.0661)
$\sum RealizedSpread_{t-p}$	0.0288***	0.2722***	-0.0083***	0.2599***
	(0.0069)	(0.0096)	(0.0022)	(0.01)
Average.adj.R^2	0.6838	0.2543	0.8231	0.2551
Average.nObs	6776	6776	6776	6776
Panel C: Daily time frame				
$\sum AT_{t-p}$	0.7002***	0.4915***	0.6583***	0.3332***
	(0.0323)	(0.1665)	(0.0322)	(0.0804)
$\sum RealizedSpread_{t-p}$	0.0143***	0.1212***	-0.0061**	0.1103***
	(0.003)	(0.0064)	(0.0026)	(0.0064)
Average.adj.R^2	0.8101	0.847	0.7703	0.8479
Average.nObs	318	318	318	318

Table 30 – Summary of VAR estimates - AT and Amihud - Multiple time frames

The table reports the average results for vector autoregression estimates in the model: $Res\_liq^k_{itd} = \alpha_0 + \sum_{p=1}^{P} \alpha_p Res\_liq^k_{it-pd} + \sum_{p=1}^{P} \gamma_p \widetilde{AT}_{it-pd} + \zeta X + v_{itd}$ (equation 3.1.1) and $\widetilde{AT}_{itd} = \beta_0 + \sum_{p=1}^{P} \beta_p \widetilde{AT}_{it-pd} + \sum_{p=1}^{P} \delta_p Res\_liq^j_{it-pd} + \zeta X + e_{itd}$ (equation 3.1.2). Res_liq^k is the adjusted *Amihud* from the residuals of equation 2.1.5. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 ($ATmsg$ and $ATtrades$). X is a set of control varbiables including $MRet$ as the average return of stocks available, $Invprice$ as the inverse of last traded price, Vol as the squared return and $MVol$ as the average Vol. Panels A,B and C reports time aggregation of 5 minutes, 15 minutes and daily observations. The number of p lags is set to 4,2 and 1 for Panels A,B and C, respectively. The first p observations for each stock i for each day d are set to zero with the exception of Panel C. Control variables coefficients were omitted for the sake of brevity. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	ATmsg	Amihud	ATtrades	Amihud
	(1)	(2)	(3)	(4)
Panel A: 5-minute time frame				
$\sum AT_{t-p}$	0.7154***	1.0245***	0.8129***	-0.1233
	(0.0275)	(0.3461)	(0.0238)	(0.0999)
$\sum Amihud_{t-p}$	0.0927**	0.4142***	-0.0088***	0.4139***
	(0.039)	(0.0101)	(0.0025)	(0.0104)
Average.adj.R^2	0.6933	0.2663	0.8439	0.1942
Average.nObs	18991	18991	18991	18991
Panel B: 15-minute time frame				
$\sum AT_{t-p}$	0.6538***	0.8787***	0.7723***	0.0517
	(0.0272)	(0.272)	(0.0258)	(0.0472)
$\sum Amihud_{t-p}$	0.1307**	0.3043***	-0.0078**	0.299***
	(0.0619)	(0.0099)	(0.0038)	(0.0102)
Average.adj.R^2	0.6785	0.2543	0.8216	0.2319
Average.nObs	6776	6776	6776	6776
Panel C: Daily time frame				
$\sum AT_{t-p}$	0.6972***	0.7871*	0.6472***	0.3359**
	(0.0325)	(0.4139)	(0.0321)	(0.1321)
$\sum Amihud_{t-p}$	-0.0128	0.3289***	-0.0018	0.3384***
	(0.0135)	(0.0233)	(0.0068)	(0.0238)
Average.adj.R^2	0.8039	0.847	0.7697	0.314
Average.nObs	318	318	318	318

Table 31 – Effect of algorithmic trading on commonality in liquidity - Different CIL measure

The table reports the effect of two proxies of AT on intraday commonality in liquidity (CIL). For each trading day we run the following OLS regression: $\Delta Res\_liq^k_{itd} = \alpha_0 + \sum_{j=-1}^{1} \alpha_j \Delta \overline{Res\_liq}^k_{-it+jd} + \epsilon_{itd}$, where Res_liq^k is one of the two adjusted liquidity measures, $\overline{Res\_liq}_{-i}$ is the average liquidity of stocks available excluding stock i. Contemporaneous, lead and lag relations are included. For each day and each stock, we sum $\overline{Res\_liq}^k$ coefficients and use them as measures of commonality in liquidity. We estimate the following equation in order to assess the effect of AT on CIL: $Commo^k_{id} = \alpha_0 + \alpha_1 \widetilde{AT}_{id} + \alpha_2 Vol_{id} + \alpha_3 InvPrice_{id} + \alpha_4 MRet_{-id} + \alpha_5 MVolat_{-id} + \epsilon_{id}$. $\widetilde{AT}$ is one of the two proxies for algorithmic trading estimated via equation 2.1.6 ($ATmsg$ and $ATtrades$). $MRet$ is the average return of stocks available, Invprice is the inverse of last traded price, Vol is the squared return and MVol is the average Vol. All variables are standardized before estimations. Newey e West (1986) standard errors are reported in parenthesis. Individual fixed effects are included. Sample period spans from 2017-04-03 to 2018-07-30 for 26 selected stocks from the Brazilian equities market. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level, respectively.

	Dependent variable:			
	$Commo^{RSpread}$		$Commo^{Amihud}$	
	(1)	(2)	(3)	(4)
$ATmsg$	0.081***		0.006	
	(0.011)		(0.010)	
$ATtrades$		0.087***		0.003
		(0.012)		(0.015)
$MRet$	0.054***	0.053***	−0.019**	−0.019**
	(0.008)	(0.009)	(0.007)	(0.007)
$Invprice$	−0.107***	−0.043***	0.118***	0.122***
	(0.015)	(0.014)	(0.026)	(0.025)
Vol	−0.0005	0.002	0.013**	0.013**
	(0.004)	(0.004)	(0.006)	(0.006)
$MVol$	0.062***	0.056***	0.050***	0.050***
	(0.007)	(0.007)	(0.006)	(0.006)
Observations	8,294	8,294	8,294	8,294
R^2	0.079	0.080	0.015	0.015
Adjusted R^2	0.075	0.077	0.011	0.011
F Statistic (df = 5; 8263)	140.863***	144.382***	24.354***	24.337***

4 Concluding remarks

Although the financial literature dedicates effort to understanding the relationship of AT and financial markets, little attention has been given to emerging markets. We provide the first evidence of this relationship for the Brazilian stock market by showing that AT negatively affects two measures of liquidity and increases commonality in liquidity.

Using the starting date of a new data center for the Brazilian stock exchange in 2017 as an instrument to control for the endogeneity between AT and liquidity, we show evidence that 1-minute AT activity increases spreads and a price-impact proxy. Both panel and vector autoregression estimates confirm our results using distinct time aggregation methods. Our results are stronger for the effect of AT on realized spreads. This is contrary to many of the studies on developed markets, which report AT as being beneficial to financial markets. Our sample excludes stocks with designated market makers, which may evidence the effect of ATs acting as voluntary market makers. We also evidence AT increasing intraday commonality in liquidity. As algorithmic trading is based on patterns, we expect that trading strategies and trend following algorithms should be correlated, inducing higher commonality.

As the net effect of AT on financial markets may never be consensual, our research has implications for the Brazilian case that could be extended to other emerging (and developed) markets. Policymakers should be aware of the positive and negative effects of AT in order to better oversee AT activity in financial markets. Further research may analyze the effect of AT on other market variables such as volatility, risk and returns.

Bibliography

ABDI, F.; RANALDO, A. A simple estimation of bid-ask spreads from daily close, high, and low prices. *The Review of Financial Studies*, Oxford University Press, v. 30, n. 12, p. 4437–4480, 2017. Cited 2 times on pages 9 and 33.

ACERBI, C.; SCANDOLO, G. Liquidity risk theory and coherent measures of risk. *Quantitative Finance*, v. 8, n. 7, p. 681–692, 2008. Cited on page 59.

ACHARYA, V. V.; PEDERSEN, L. H. Asset pricing with liquidity risk. *Journal of financial Economics*, Elsevier, v. 77, n. 2, p. 375–410, 2005. Cited 9 times on pages 17, 20, 24, 35, 40, 70, 79, 91, and 93.

ADRIAN, T.; SHIN, H. S. Liquidity and leverage. *Journal of financial intermediation*, Elsevier, v. 19, n. 3, p. 418–437, 2010. Cited on page 36.

ADRIAN, T.; SHIN, H. S. Procyclical leverage and value-at-risk. *The Review of Financial Studies*, Oxford University Press, v. 27, n. 2, p. 373–403, 2014. Cited on page 36.

AHN, H.-J.; CAI, J.; HAMAO, Y.; MELVIN, M. Little guys, liquidity, and the informational efficiency of price: Evidence from the tokyo stock exchange on the effects of small investor participation. *Pacific-Basin Finance Journal*, Elsevier, v. 29, p. 163–181, 2014. Cited 3 times on pages 15, 16, and 29.

AMIHUD, Y. Iliquidity and stock returns: cross-section ant time series effects. *Journal of Financial Markets*, v. 5, p. 31–56, 2002. Cited 13 times on pages 17, 20, 23, 33, 34, 38, 40, 41, 44, 47, 75, 80, and 81.

AMIHUD, Y.; HAMEED, A.; KANG, W.; ZHANG, H. The illiquidity premium: International evidence. *Journal of Financial Economics*, Elsevier, v. 117, n. 2, p. 350–368, 2015. Cited 3 times on pages 8, 9, and 33.

AMIHUD, Y.; MENDELSON, H. Asset pricing and the bid-ask spread. *Journal of Financial Economics*, Elsevier, v. 17, n. 2, p. 223–249, 1986. Cited 2 times on pages 9 and 20.

AMIHUD, Y.; MENDELSON, H.; PEDERSEN, L. *Market Liquidity Asset Pricing, Risk, and Crises*. [S.l.]: Cambridge University Press, 2012. Cited on page 33.

AMIHUD, Y.; MENDELSON, H.; UNO, J. Number of shareholders and stock prices: Evidence from japan. *The Journal of Finance*, Wiley Online Library, v. 54, n. 3, p. 1169–1184, 1999. Cited 2 times on pages 16 and 29.

ANAGNOSTIDIS, P.; FONTAINE, P. C. Liquidity provision, commonality and high frequency trading. *EUROFIDAI Working Paper*, 2018. Cited 2 times on pages 74 and 92.

ANG, A.; HODRICK, R. J.; XING, Y.; ZHANG, X. The cross-section of volatility and expected returns. *The Journal of Finance*, Wiley Online Library, v. 61, n. 1, p. 259–299, 2006. Cited on page 35.

ANGELIDIS, T.; BENOS, A. Liquidity adjusted value-at-risk based on the components of the bid-ask spread. *Applied Financial Economics*, Taylor & Francis, v. 16, n. 11, p. 835–851, 2006. Cited on page 59.

ANTHONISZ, S. A.; PUTNIŅŠ, T. J. Asset pricing with downside liquidity risks. *Management Science*, INFORMS, v. 63, n. 8, p. 2549–2572, 2016. Cited 8 times on pages 9, 33, 35, 39, 44, 70, 91, and 93.

ARTZNER, P.; DELBAEN, F.; EBER, J.; HEATH, D. Coherent measures of risk. *Mathematical Finance*, v. 9, n. 3, p. 203–228, 1999. Cited on page 59.

ASTORINO, E.; CHAGUE, F.; GIOVANNETTI, B.; SILVA, M. E. Variance premium and implied volatility in a low-liquidity option market. *Brazilian Review of Economics*, 2017. Cited 3 times on pages 18, 20, and 25.

BANGIA, A.; DIEBOLD, F. X.; SCHUERMANN, T.; STROUGHAIR, J. D. Modeling liquidity risk, with implications for traditional market risk measurement and management. In: *Risk management: The state of the art*. [S.l.]: Springer, 2001. p. 3–13. Cited on page 59.

BAO, J.; PAN, J.; WANG, J. The illiquidity of corporate bonds. *The Journal of Finance*, Wiley Online Library, v. 66, n. 3, p. 911–946, 2011. Cited on page 35.

BARINOV, A. Turnover: liquidity or uncertainty? *Management Science*, INFORMS, v. 60, n. 10, p. 2478–2495, 2014. Cited on page 37.

BARRAS, L.; MALKHOZOV, A. Does variance risk have two prices? evidence from the equity and option markets. *Journal of Financial Economics*, Elsevier, v. 121, n. 1, p. 79–92, 2016. Cited on page 47.

BATTALIO, R.; CORWIN, S. A.; JENNINGS, R. Unrecognized odd lot liquidity supply: A hidden trading cost for high priced stocks. *The Journal of Trading*, v. 12, n. 1, p. 35–41, 2017. Cited on page 15.

BEKAERT, G.; ENGSTROM, E. Asset return dynamics under habits and bad environment–good environment fundamentals. *Journal of Political Economy*, University of Chicago Press Chicago, IL, v. 125, n. 3, p. 713–760, 2017. Cited on page 43.

BEKAERT, G.; HARVEY, C. R. Research in emerging markets finance: looking to the future. *Emerging Markets Review*, Elsevier, v. 3, n. 4, p. 429–448, 2002. Cited on page 16.

BEKAERT, G.; HOEROVA, M. The vix, the variance premium and stock market volatility. *Journal of Econometrics*, Elsevier, v. 183, n. 2, p. 181–192, 2014. Cited 6 times on pages 10, 35, 43, 49, 51, and 70.

BELKHIR, M.; SAAD, M.; SAMET, A. Stock extreme illiquidity and the cost of capital. *Journal of Banking & Finance*, Elsevier, 2018. Cited 3 times on pages 9, 33, and 39.

BEN-REPHAEL, A.; KADAN, O.; WOHL, A. The diminishing liquidity premium. *Journal of Financial and Quantitative Analysis*, Cambridge Univ Press, v. 50, n. 1-2, p. 197–229, 2015. Cited 4 times on pages 8, 9, 23, and 33.

BENOS, E.; BRUGLER, J.; HJALMARSSON, E.; ZIKES, F. Interactions among high-frequency traders. *Journal of Financial and Quantitative Analysis*, Cambridge University Press, v. 52, n. 4, p. 1375–1402, 2017. Cited 3 times on pages 84, 91, and 93.

BENOS, E.; SAGADE, S. Price discovery and the cross-section of high-frequency trading. *Journal of Financial Markets*, Elsevier, v. 30, p. 54–77, 2016. Cited 2 times on pages 10 and 74.

BIAIS, B.; FOUCAULT, T.; MOINAS, S. Equilibrium fast trading. *Journal of Financial Economics*, Elsevier, v. 116, n. 2, p. 292–313, 2015. Cited on page 75.

BIS. Basel iii: Finalising post-crisis reforms. *Bank for International Settlements*, 2017. Cited on page 60.

BLACK, J. R.; STOCK, D.; YADAV, P. K. The pricing of different dimensions of liquidity: Evidence from government guaranteed bonds. *Journal of Banking & Finance*, Elsevier, v. 71, p. 119–132, 2016. Cited on page 33.

BOEHMER, E.; FONG, K. Y.; WU, J. J. Algorithmic trading and market quality : international evidence. *Working Paper*, 2018. Cited 3 times on pages 75, 80, and 92.

BOEHMER, E.; LI, D.; SAAR, G. Correlated high-frequency trading. *Manuscript, Cornell University, Ithaca, NY*, 2015. Cited 2 times on pages 84 and 91.

BOLLERSLEV, T.; TAUCHEN, G.; ZHOU, H. Expected stock returns and variance risk premia. *The Review of Financial Studies*, Oxford University Press, v. 22, n. 11, p. 4463–4492, 2009. Cited on page 43.

BRENNAN, M.; HUH, S.-W.; SUBRAHMANYAM, A. An analysis of the amihud illiquidity premium. *The Review of Asset Pricing Studies*, Oxford University Press, v. 3, n. 1, p. 133–176, 2013. Cited on page 38.

BROGAARD, J. High frequency trading and its impact on market quality. *Northwestern University Kellogg School of Management Working Paper*, v. 66, 2010. Cited 2 times on pages 11 and 76.

BROGAARD, J.; GARRIOTT, C. High-frequency trading competition. *Journal of Financial and Quantitative Analysis*, Cambridge University Press, v. 54, n. 4, p. 1469–1497, 2019. Cited on page 74.

BROGAARD, J.; HENDERSHOTT, T.; RIORDAN, R. High-frequency trading and price discovery. *The Review of Financial Studies*, Oxford University Press, v. 27, n. 8, p. 2267–2306, 2014. Cited 5 times on pages 10, 22, 33, 75, and 93.

BRUNNERMEIER, M. K.; PEDERSEN, L. H. Market liquidity and funding liquidity. *The Review of Financial Studies*, Society for Financial Studies, v. 22, n. 6, p. 2201–2238, 2009. Cited 14 times on pages 12, 16, 23, 34, 35, 36, 44, 45, 52, 63, 67, 70, 76, and 93.

BUDISH, E.; CRAMTON, P.; SHIM, J. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, MIT Press, v. 130, n. 4, p. 1547–1621, 2015. Cited on page 75.

CARR, P.; WU, L. Variance risk premiums. *The Review of Financial Studies*, v. 22, n. 3, p. 1311–1341, 2008. Cited 2 times on pages 18 and 55.

CARTEA, Á.; PAYNE, R.; PENALVA, J.; TAPIA, M. Ultra-fast activity and intraday market quality. *Journal of Banking & Finance*, Elsevier, v. 99, p. 157–181, 2019. Cited 2 times on pages 33 and 75.

CHABOUD, A. P.; CHIQUOINE, B.; HJALMARSSON, E.; VEGA, C. Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, Wiley Online Library, v. 69, n. 5, p. 2045–2084, 2014. Cited 4 times on pages 11, 76, 87, and 91.

CHACKO, G.; DAS, S.; FAN, R. An index-based measure of liquidity. *Journal of Banking & Finance*, Elsevier, v. 68, p. 162–178, 2016. Cited 6 times on pages 8, 9, 33, 43, 52, and 60.

CHEN, Z.; LU, A. A market-based funding liquidity measure. *Working Paper*, 2017. Cited on page 60.

CHORDIA, T.; ROLL, R.; SUBRAHMANYAM, A. Commonality in liquidity. *Journal of financial economics*, Elsevier, v. 56, n. 1, p. 3–28, 2000. Cited 5 times on pages 8, 33, 35, 91, and 92.

CHORDIA, T.; ROLL, R.; SUBRAHMANYAM, A. Market liquidity and trading activity. *The Journal of Finance*, v. 56, n. 2, p. 501–530, 2001. Cited 3 times on pages 16, 23, and 52.

CHORDIA, T.; ROLL, R.; SUBRAHMANYAM, A. Order imbalance, liquidity, and market returns. *Journal of Financial economics*, v. 65, n. 1, p. 111–130, 2002. Cited on page 52.

CHORDIA, T.; SARKAR, A.; SUBRAHMANYAM, A. An empirical analysis of stock and bond market liquidity. *Review of Financial Studies*, v. 18, n. 1, p. 85–129, 2005. Cited 4 times on pages 33, 44, 45, and 62.

CHUNG, K. H.; CHUWONGANANT, C. Uncertainty, market structure, and liquidity. *Journal of Financial Economics*, v. 113, n. 3, p. 476–499, 2014. Cited 3 times on pages 34, 44, and 52.

CHUNG, K. H.; CHUWONGANANT, C. Market volatility and stock returns: The role of liquidity providers. *Journal of Financial Markets*, Elsevier, v. 37, p. 17–34, 2018. Cited on page 34.

COMERTON-FORDE, C.; HENDERSHOTT, T.; JONES, C. M.; MOULTON, P. C.; SEASHOLES, M. S. Time variation in liquidity: The role of market-maker inventories and revenues. *The Journal of Finance*, Wiley Online Library, v. 65, n. 1, p. 295–331, 2010. Cited on page 36.

CONRAD, J.; WAHAL, S.; XIANG, J. High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics*, Elsevier, v. 116, n. 2, p. 271–291, 2015. Cited on page 74.

CORWIN, S. A.; SCHULTZ, P. A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance*, Wiley Online Library, v. 67, n. 2, p. 719–760, 2012. Cited 6 times on pages 18, 20, 25, 34, 37, and 47.

DAVIS, R. L.; ROSEMAN, B. S.; NESS, B. F. V.; NESS, R. V. 1-share orders and trades. *Journal of Banking & Finance*, Elsevier, v. 75, p. 109–117, 2017. Cited 5 times on pages 9, 15, 16, 21, and 28.

DIONNE, G.; PACURAR, M.; ZHOU, X. Liquidity-adjusted intraday value at risk modeling and risk management: an application to data from deutsche börse. *Journal of Banking & Finance*, v. 59, p. 202–219, 2015. Cited on page 59.

DYL, E. A.; MABERLY, E. D. Odd-lot transactions around the turn of the year and the january effect. *Journal of Financial and Quantitative Analysis*, Cambridge Univ Press, v. 27, n. 04, p. 591–604, 1992. Cited on page 15.

EASLEY, D.; O'HARA, M. Price, trade size, and information in securities markets. *Journal of Financial economics*, Elsevier, v. 19, n. 1, p. 69–90, 1987. Cited on page 8.

EVANS, M. D.; HNATKOVSKA, V. V. International capital flows, returns and world financial integration. *Journal of International Economics*, Elsevier, v. 92, n. 1, p. 14–33, 2014. Cited on page 74.

FAMA, E. F. The behavior of stock-market prices. *The journal of Business*, JSTOR, v. 38, n. 1, p. 34–105, 1965. Cited on page 8.

FAMA, E. F.; FRENCH, K. R. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, v. 33, n. 1, p. 3–56, 1993. Cited on page 67.

FONG, K.; HOLDEN, C.; TOBEK, O. Are volatility over volume liquidity proxies useful for global or us research? 2018. Cited 5 times on pages 9, 33, 34, 39, and 40.

FONG, K. Y.; HOLDEN, C. W.; TRZCINKA, C. A. What are the best liquidity proxies for global research? *Review of Finance*, v. 21, n. 4, p. 1355–1401, 2017. Cited 2 times on pages 10 and 33.

FOUCAULT, T.; KOZHAN, R.; THAM, W. W. Toxic arbitrage. *The Review of Financial Studies*, Oxford University Press, v. 30, n. 4, p. 1053–1094, 2017. Cited 2 times on pages 75 and 78.

FOUCAULT, T.; PAGANO, M.; RÖELL, A. *Market liquidity: theory, evidence, and policy.* [S.l.]: Oxford University Press, 2013. Cited 2 times on pages 8 and 33.

FOUCAULT, T.; SRAER, D.; THESMAR, D. J. Individual investors and volatility. *The Journal of Finance*, Wiley Online Library, v. 66, n. 4, p. 1369–1406, 2011. Cited on page 15.

GABRIELSEN, A.; MARZO, M.; ZAGAGLIA, P. Measuring market liquidity: An introductory survey. *Arvix Working Paper*, 2011. Disponível em: <http://arxiv.org/abs/1112.6169v1>. Cited on page 33.

GARMAN, M. B.; KLASS, M. J. On the estimation of security price volatilities from historical data. *Journal of business*, p. 67–78, 1980. Cited on page 38.

GLOSTEN, L. R.; MILGROM, P. R. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, Elsevier, v. 14, n. 1, p. 71–100, 1985. Cited 2 times on pages 8 and 10.

GORMLEY, T. A.; MATSA, D. A. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies*, Soc Financial Studies, v. 27, n. 2, p. 617–661, 2014. Cited on page 19.

GOYENKO, R. Y.; HOLDEN, C. W.; TRZCINKA, C. A. Do liquidity measures measure liquidity? *Journal of Financial Economics*, v. 92, n. 2, p. 153 – 181, 2009. Cited 5 times on pages 10, 33, 44, 70, and 80.

GOZLUKLU, A. E.; PEROTTI, P.; RINDI, B.; FREDELLA, R. Lot size constraints and market quality: evidence from the borsa italiana. *Financial Management*, Wiley Online Library, v. 44, n. 4, p. 905–945, 2015. Cited 3 times on pages 15, 18, and 29.

GRANGER, C. W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 424–438, 1969. Cited 3 times on pages 52, 59, and 88.

GRECHUK, B.; MOLYBOHA, A.; ZABARANKIN, M. Maximum entropy principle with general deviation measures. *Mathematics of Operations Research*, v. 34, n. 2, p. 445–467, 2009. Cited on page 39.

GRILLINI, S.; OZKAN, A.; SHARMA, A.; JANABI, M. A. A. Pricing of time-varying illiquidity within the eurozone: Evidence using a markov switching liquidity-adjusted capital asset pricing model. *International Review of Financial Analysis*, Elsevier, v. 64, p. 145–158, 2019. Cited 2 times on pages 9 and 33.

HAGSTRÖMER, B.; NORDEN, L. The diversity of high-frequency traders. *Journal of Financial Markets*, Elsevier, v. 16, n. 4, p. 741–770, 2013. Cited 3 times on pages 12, 74, and 76.

HAMEED, A.; KANG, W.; VISWANATHAN, S. Stock market declines and liquidity. *The Journal of Finance*, Wiley Online Library, v. 65, n. 1, p. 257–293, 2010. Cited 5 times on pages 36, 39, 84, 86, and 93.

HASBROUCK, J. *Empirical market microstructure*. [S.l.]: Oxford University Press New York, 2007. v. 1. Cited on page 33.

HASBROUCK, J.; SAAR, G. Low-latency trading. *Journal of Financial Markets*, Elsevier, v. 16, n. 4, p. 646–679, 2013. Cited 2 times on pages 74 and 78.

HAUSER, S.; LAUTERBACH, B. The impact of minimum trading units on stock value and price volatility. *Journal of Financial and Quantitative Analysis*, Cambridge Univ Press, v. 38, n. 03, p. 575–589, 2003. Cited 3 times on pages 16, 18, and 29.

HAUSMAN, J. A. Specification tests in econometrics. *Econometrica: Journal of the econometric society*, JSTOR, p. 1251–1271, 1978. Cited on page 19.

HENDERSHOTT, T.; JONES, C. M.; MENKVELD, A. J. Does algorithmic trading improve liquidity? *The Journal of Finance*, Wiley Online Library, v. 66, n. 1, p. 1–33, 2011. Cited 9 times on pages 10, 11, 18, 33, 74, 75, 79, 80, and 88.

HENDERSHOTT, T.; RIORDAN, R. Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, Cambridge University Press, v. 48, n. 4, p. 1001–1024, 2013. Cited 2 times on pages 33 and 88.

HIRSCHEY, N. Do high-frequency traders anticipate buying and selling pressure? *Working Paper*, 2017. Cited 3 times on pages 10, 75, and 78.

HOFFMANN, P. A dynamic limit order market with fast and slow traders. *Journal of Financial Economics*, Elsevier, v. 113, n. 1, p. 156–169, 2014. Cited on page 75.

HOLDEN, C. W.; JACOBSEN, S.; SUBRAHMANYAM, A. et al. The empirical analysis of liquidity. *Foundations and Trends in Finance*, Now Publishers, Inc., v. 8, n. 4, p. 263–365, 2014. Cited 2 times on pages 33 and 74.

HULL, M.; MCGROARTY, F. Do emerging markets become more efficient as they develop? long memory persistence in equity indices. *Emerging Markets Review*, Elsevier, v. 18, p. 45–61, 2014. Cited on page 75.

ISAKA, N. Long-run effects of minimum trading unit reductions on stock prices. *International Review of Finance*, Wiley Online Library, v. 14, n. 1, p. 75–103, 2014. Cited 2 times on pages 16 and 29.

JAIN, P. K.; JAIN, P.; MCINISH, T. H. Does high-frequency trading increase systemic risk? *Journal of Financial Markets*, Elsevier, v. 31, p. 1–24, 2016. Cited 2 times on pages 76 and 92.

JAME, R. Liquidity provision and the cross section of hedge fund returns. *Management Science*, INFORMS, v. 64, n. 7, p. 2973–3468, 2017. Cited on page 60.

JAWED, M. S.; CHAKRABARTI, P. Role of algorithmic and co-location trading on the speed of information adjustments: Evidence from india. *Emerging Markets Finance and Trade*, Taylor & Francis, v. 54, n. 9, p. 2021–2039, 2018. Cited 2 times on pages 75 and 92.

JOHNSON, H. Odd lot trades: The behavior, characteristics, and information content, over time. *Financial Review*, Wiley Online Library, v. 49, n. 4, p. 669–684, 2014. Cited on page 15.

JOHNSON, H.; NESS, B. F. V.; NESS, R. A. V. Are all odd-lots the same? odd-lot transactions by order submission and trader type. *Journal of Banking & Finance*, Elsevier, v. 79, p. 1–11, 2017. Cited 3 times on pages 9, 15, and 21.

JONES, C. M. What do we know about high-frequency trading? *Columbia Business School Research Paper*, n. 13–11, 2013. Cited 3 times on pages 11, 28, and 74.

JORION, P. et al. *Financial risk manager handbook*. [S.l.: s.n.], 2007. v. 406. Cited on page 58.

JOVANOVIC, B.; MENKVELD, A. J. Middlemen in limit order markets. *Available at SSRN 1624329*, 2016. Cited on page 74.

KAMARA, A.; LOU, X.; SADKA, R. The divergence of liquidity commonality in the cross-section of stocks. *Journal of Financial Economics*, v. 89, n. 3, p. 444–466, 2008. Cited 3 times on pages 17, 35, and 40.

KAROLYI, G. A.; LEE, K.-H.; DIJK, M. A. V. Understanding commonality in liquidity around the world. *Journal of Financial Economics*, v. 105, n. 1, p. 82–112, 2012. Cited 4 times on pages 35, 44, 91, and 93.

KEARNEY, C. Emerging markets research: trends, issues and future directions. *Emerging Markets Review*, Elsevier, v. 13, n. 2, p. 159–183, 2012. Cited on page 16.

KEMPF, A.; MAYSTON, D. L.; GEHDE-TRAPP, M.; YADAV, P. K. Resiliency: A dynamic view of liquidity. *Working Paper*, 2015. Cited on page 33.

KIRILENKO, A.; KYLE, A. S.; SAMADI, M.; TUZUN, T. The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, Wiley Online Library, v. 72, n. 3, p. 967–998, 2017. Cited on page 74.

KLEIN, O.; SONG, S. Multimarket high-frequency trading and commonality in liquidity. *Available at SSRN 2984887*, 2017. Cited on page 76.

KOCH, A.; RUENZI, S.; STARKS, L. Commonality in liquidity: a demand-side explanation. *The Review of Financial Studies*, Society for Financial Studies, v. 29, n. 8, p. 1943–1974, 2016. Cited on page 35.

KORAJCZYK, R. A.; SADKA, R. Pricing the commonality across alternative measures of liquidity. *Journal of Financial Economics*, v. 87, n. 1, p. 45–72, 2008. Cited 2 times on pages 58 and 62.

KYLE, A. S. Continuous auctions and insider trading. *Econometrica*, JSTOR, p. 1315–1335, 1985. Cited on page 8.

LAKONISHOK, J.; MABERLY, E. The weekend effect: Trading patterns of individual and institutional investors. *The Journal of Finance*, Wiley Online Library, v. 45, n. 1, p. 231–243, 1990. Cited on page 15.

LEE, E. J. High frequency trading in the korean index futures market. *Journal of Futures Markets*, Wiley Online Library, v. 35, n. 1, p. 31–51, 2015. Cited 3 times on pages 75, 87, and 92.

LESMOND, D. A. Liquidity of emerging markets. *Journal of Financial Economics*, Elsevier, v. 77, n. 2, p. 411–452, 2005. Cited on page 75.

LINTON, O.; MAHMOODZADEH, S. Implications of high-frequency trading for security markets. *Annual Review of Economics*, Annual Reviews, v. 10, p. 237–259, 2018. Cited on page 74.

LONGSTAFF, F. A.; PAN, J.; PEDERSEN, L. H.; SINGLETON, K. J. How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*, v. 3, n. 2, p. 75–103, 2011. Cited on page 35.

MALCENIECE, L.; MALCENIEKS, K.; PUTNIŅŠ, T. J. High frequency trading and comovement in financial markets. *Journal of Financial Economics*, Elsevier, 2019. Cited 7 times on pages 11, 75, 76, 79, 80, 92, and 93.

MALKIEL, B. G.; FAMA, E. F. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, Wiley Online Library, v. 25, n. 2, p. 383–417, 1970. Cited on page 8.

MANCINI, L.; RANALDO, A.; WRAMPELMEYER, J. Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. *The Journal of Finance*, Wiley Online Library, v. 68, n. 5, p. 1805–1841, 2013. Cited on page 58.

MENKVELD, A. J. High frequency trading and the new market makers. *Journal of Financial Markets*, Elsevier, v. 16, n. 4, p. 712–740, 2013. Cited 5 times on pages 10, 12, 22, 74, and 76.

MENKVELD, A. J. The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics*, Annual Reviews, v. 8, p. 1–24, 2016. Cited 2 times on pages 28 and 74.

MERTON, R. C. A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, v. 42, n. 3, p. 483–510, 1987. Cited on page 29.

MESTEL, R.; MURG, M.; THEISSEN, E. Algorithmic trading and liquidity: Long term evidence from austria. *Finance Research Letters*, Elsevier, v. 26, p. 198–203, 2018. Cited 5 times on pages 11, 76, 80, 84, and 86.

MORCK, R.; YEUNG, B.; YU, W. The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of financial economics*, Elsevier, v. 58, n. 1-2, p. 215–260, 2000. Cited 2 times on pages 11 and 76.

MORIYASU, H.; WEE, M.; YU, J. The role of algorithmic trading in stock liquidity and commonality in electronic limit order markets. *Pacific-Basin Finance Journal*, Elsevier, v. 49, p. 103–128, 2018. Cited 8 times on pages 11, 74, 76, 80, 84, 91, 92, and 93.

MÜLLER, F. M.; RIGHI, M. B. Numerical comparison of multivariate models to forecasting risk measures. *Risk Management*, Springer, v. 20, n. 1, p. 29–50, 2018. Cited on page 59.

NAGEL, S. Evaporating liquidity. *Review of Financial Studies*, v. 25, n. 7, p. 2005–2039, 2012. Cited 5 times on pages 16, 23, 34, 36, and 52.

NEWEY, W. K.; WEST, K. D. *A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1986. Cited 8 times on pages 18, 25, 85, 87, 94, 95, 96, and 99.

NIETO, B. Bid–ask spread estimator from high and low daily prices: Practical implementation for corporate bonds. *Journal of Empirical Finance*, Elsevier, 2018. Cited 2 times on pages 9 and 33.

O'HARA, M. High frequency market microstructure. *Journal of Financial Economics*, Elsevier, v. 116, n. 2, p. 257–270, 2015. Cited on page 74.

O'HARA, M.; YAO, C.; YE, M. What's not there: Odd lots and market data. *The Journal of Finance*, Wiley Online Library, v. 69, n. 5, p. 2199–2236, 2014. Cited 5 times on pages 9, 15, 19, 20, and 28.

PAN, J.; SINGLETON, K. J. Default and recovery implicit in the term structure of sovereign cds spreads. *The Journal of Finance*, Wiley Online Library, v. 63, n. 5, p. 2345–2384, 2008. Cited on page 35.

PASTOR, L.; STAMBAUGH, R. F. Liquidity risk and expected stock returns. *Journal of Political Economy*, v. 111, n. 3, p. 642–685, 2003. Cited 3 times on pages 40, 47, and 79.

PERLIN, M. The effects of the introduction of market makers in the brazilian equity market. *Brazilian Review of Finance*, Brazilian Finance Society, v. 11, n. 2, 2013. Cited on page 78.

PERLIN, M. S.; RAMOS, H. P. Gethfdata: A r package for downloading and aggregating high frequency trading data from bovespa. *Brazilian Review of Finance*, v. 14, n. 3, p. 443–478, 2017. Cited 2 times on pages 17 and 78.

RAMOS, H. P.; PERLIN, M. S.; RIGHI, M. B. Mispricing in the odd lots market in brazil. *The North American Journal of Economics and Finance*, Elsevier, v. 42, p. 618–628, 2017. Cited on page 78.

RIGHI, M. B.; CERETTA, P. S. Shortfall deviation risk: An alternative for risk measurement. *Journal of Risk*, v. 19, n. 2, p. 81–116, 2016. Cited on page 59.

ROCKAFELLAR, R.; URYASEV, S. The fundamental risk quadrangle in risk management, optimization and statistical estimation. *Surveys in Operations Research and Management Science*, v. 18, n. 1-2, p. 33–53, 2013. Cited on page 39.

ROCKAFELLAR, R.; URYASEV, S.; ZABARANKIN, M. Generalized deviations in risk analysis. *Finance and Stochastics*, v. 10, p. 51–74, 2006. Cited 2 times on pages 34 and 39.

ROLL, R. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, Wiley Online Library, v. 39, n. 4, p. 1127–1139, 1984. Cited 3 times on pages 34, 38, and 47.

ROSEMAN, B. S.; NESS, B. F. V.; NESS, R. A. V. Odd-lot trading in us equities. *Working Paper*, 2016. Cited 2 times on pages 15 and 19.

ROSSI, E.; MAGISTRIS, P. S. D. Long memory and tail dependence in trading volume and volatility. *Journal of Empirical Finance*, Elsevier, v. 22, p. 94–112, 2013. Cited 2 times on pages 47 and 53.

RUBIA, A.; SANCHIS-MARCO, L. On downside risk predictability through liquidity and trading activity: A dynamic quantile approach. *International Journal of Forecasting*, Elsevier, v. 29, n. 1, p. 202–219, 2013. Cited 2 times on pages 36 and 62.

RUENZI, S.; UNGEHEUER, M.; WEIGERT, F. Extreme downside liquidity risk. *Working Paper*, 2016. Cited 2 times on pages 36 and 70.

SANVICENTE, A. Determinants of transactions costs in the brazilian stock market. *Brazilian Review of Finance*, v. 10, n. 2, p. 179–196, 2012. ISSN 1984-5146. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/rbfin/article/view/3536>. Cited on page 9.

SCHESTAG, R.; SCHUSTER, P.; UHRIG-HOMBURG, M. Measuring liquidity in bond markets. *Review of Financial Studies*, p. hhv132, 2016. Cited 2 times on pages 10 and 33.

SEC. Concept release on equity market structure. *Concept Release 34-61358*, File S7-02-10, p. 3594–3614, 2010. Cited on page 78.

SEC. Findings regarding the market events of may 6, 2010. *Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues*, 2010. Cited on page 11.

SHILLER, R. J. *Finance and the good society.* [S.l.]: Princeton University Press, 2013. Cited on page 8.

SKJELTORP, J. A.; SOJLI, E.; THAM, W. W. Trading on algos. *Working Paper*, 2015. Cited on page 74.

SOPRANO, A. *Liquidity management: a funding risk handbook.* [S.l.]: John Wiley & Sons, 2015. Cited on page 59.

STOLL, H. R. The pricing of security dealer services: An empirical study of nasdaq stocks. *The Journal of Finance*, Wiley Online Library, v. 33, n. 4, p. 1153–1172, 1978. Cited on page 34.

STOLL, H. R. Presidential address: friction. *The Journal of Finance*, Wiley Online Library, v. 55, n. 4, p. 1479–1514, 2000. Cited on page 34.

SYAMALA, S. R.; WADHWA, K.; GOYAL, A. Determinants of commonality in liquidity: Evidence from an order-driven emerging market. *The North American Journal of Economics and Finance*, Elsevier, v. 42, p. 38–52, 2017. Cited on page 36.

UPSON, J.; NESS, R. A. V. Multiple markets, algorithmic trading, and market liquidity. *Journal of Financial Markets*, Elsevier, v. 32, p. 49–68, 2017. Cited on page 78.

VICTOR, F. G.; PERLIN, M. S.; MASTELLA, M. Commonalities in liquidity – evidence and intraday patterns in the brazilian market. *Brazilian Review of Finance*, Brazilian Finance Society, v. 11, n. 3, 2013. Cited on page 78.

VIRGILIO, G. P. M. High-frequency trading: a literature review. *Financial Markets and Portfolio Management*, Springer, p. 1–26, 2019. Cited on page 74.

WAGALATH, L.; ZUBELLI, J. P. A liquidation risk adjustment for value at risk and expected shortfall. *International Journal of Theoretical and Applied Finance*, v. 21, n. 2, 2017. Cited on page 59.

WEISS, G. N.; SUPPER, H. Forecasting liquidity-adjusted intraday value-at-risk with vine copulas. *Journal of Banking & Finance*, Elsevier, v. 37, n. 9, p. 3334–3350, 2013. Cited on page 59.

WU, Y. Asset pricing with extreme liquidity risk. *Journal of Empirical Finance*, Elsevier, 2019. Cited 2 times on pages 9 and 33.

APPENDIX A – Corwin and Schultz (2012) Spread Estimator

$$Spread_{id} = \frac{2(e^{\alpha} - 1)}{1 - e^{\alpha}} \tag{A.0.1}$$

$$\alpha = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}} \tag{A.0.2}$$

$$\beta = \sum_{j=0}^{1} \left(\log \left(\frac{H_{id+j}}{L_{id+j}} \right) \right)^2 \tag{A.0.3}$$

$$\gamma = \left(\log \left(\frac{\max\left(H_{id}, H_{id+1}\right)}{\min\left(L_{id}, L_{id+1}\right)} \right) \right), \tag{A.0.4}$$

where H_{id} (L_{id}) is the highest (lowest) price traded for asset i on day d.