Can Sentiments of Social Media Participants reflect by Financial Market Liquidity

PhD Dissertation

Candidate Name: Jawad Saleemi

Universitat Politècnica de València
46022, València
Spain

Supervised by:

Prof. Francisco Guijarro, & Prof. Ismael Moya-Clemente

09/12/2023

Acknowledgement

I am keenly indebted to my supervisors, Prof. Francisco Guijarro and Prof. Ismael Moya-Clemente, for their unwavering assistance and counselling throughout the doctoral programme. Undoubtedly, their prowess and stoicism have deeply contributed a pivotal role in the execution of this dissertation. I am appreciative of the economic and social sciences department, universitat Politècnica de València, for all potential resources to conduct the PhD research. I also extend this gratitude to the Academic Commission of Doctoral Programme (CAPD).

Abstract

This doctoral dissertation falls in the research area of economic and social sciences department, and focuses on the behavioral perspective of market liquidity. The time-varying liquidity and its related issues are one of the dominant concerns in the market microstructure literature. The critical role of market liquidity in executing the transactions or determining the yield on investment is raising concerns for both academics and those who engage in the trading. There is thus need to unveil the potential issues, that may impact the financial market liquidity.

This dissertation seeks to understand market liquidity and its related issues in the light of investors' behavior. The behavioral perspective of liquidity is examined using microblogging-opinionated information. The escalation of behavioral finance literature also comprises the authenticity of microblogging data in both modeling and predicting various concerns associated with the efficient functioning of financial markets. However, previous research in the behavioral finance domain might have ignored a few potential implications of microblogging-opinionated information on market liquidity at the market and firm levels. Therefore, the dissertation aims to be the first empirical attempt in this area of research. The thesis is carried out as a compendium of scientific papers, whose memory includes several research articles published in the indexed journals.

The first article provides insights into relationship between microblogging content and liquidity-facilitating cost. During trading periods, this study suggested that investors' mood was less influential in affecting the time-varying liquidity and its providing cost. However, the incoming information on a given day was more influential for following trading sessions. The sentiments built on a two-day basis were associated with the liquidity-facilitating cost. The second article covers the dimensions of market liquidity using microblogging opinions. This research revealed that investor sentiments in environments of pessimism had more authoritative power on liquidity dimensions including the trading costs, transaction immediacy, price dispersion and trading volume. Finally, the third research paper explores the systematic sentiment risk for liquidity in relation to the microblogging data. This study depicted that the bank index liquidity was exposed to the systematic sentiment and liquidity risks, but non-financial firm index liquidity was only exposed to a systematic liquidity risk.

The emotion-driven market participants on microblogging platform may not only influence the time-varying market liquidity and its dimensions, but they may also expose to the systematic risk for liquidity withing a broader market. Thus, liquidity and its related aspects are suggested to be priced against the adverse selection issues in the market. Additionally, the measurement of incoming information on microblogging platform may better assist the liquidity providers in the construction of portfolio.

Resumen

Esta tesis doctoral se enmarca en el área de investigación del Departamento de Economía y Ciencias Sociales, y se centra en la perspectiva conductual de la liquidez del mercado. La liquidez que varía en el tiempo y sus problemas relacionados son una de las preocupaciones dominantes en la literatura de microestructura del mercado. El papel crítico de la liquidez del mercado en la ejecución de transacciones o la determinación del rendimiento de la inversión genera inquietudes tanto para académicos como para aquellos que participan en el mercado. Por lo tanto, es necesario desvelar los problemas potenciales que pueden afectar la liquidez del mercado financiero.

Esta tesis busca entender la liquidez del mercado y sus problemas relacionados a la luz del comportamiento de los inversores. La perspectiva conductual de la liquidez se examina utilizando información orientada a opiniones en microblogs. La creciente literatura de finanzas conductuales también incluye la autenticidad de los datos de microblogs tanto en la modelización como en la predicción de diversas preocupaciones asociadas con el funcionamiento eficiente de los mercados financieros. Sin embargo, la investigación previa en el ámbito de las finanzas conductuales podría haber pasado por alto algunas implicaciones potenciales de la información orientada a opiniones en microblogs sobre la liquidez del mercado a nivel de mercado y de empresa. Por lo tanto, la tesis pretende ser una aplicación empírica en esta área de investigación. La tesis se lleva a cabo como un compendio de artículos científicos, cuya memoria incluye varios artículos de investigación publicados en revistas indexadas.

El primer artículo proporciona información sobre la relación entre el contenido de microblogs y el coste de facilitación de la liquidez. Durante los períodos de negociación, este estudio sugirió que el estado de ánimo de los inversionistas tenía menos influencia en afectar la liquidez que varía en el tiempo y su coste de facilitación. Sin embargo, la información entrante en un día dado fue más influyente para las sesiones de negociación siguientes. Los sentimientos construidos sobre una base de dos días estaban asociados con el costo de facilitación de la liquidez. El segundo artículo aborda las dimensiones de la liquidez del mercado utilizando opiniones de microblogs. Esta investigación reveló que los sentimientos de los inversores en entornos de pesimismo tenían más poder autoritario sobre las dimensiones de la liquidez, incluidos los costes de negociación, la inmediatez de la transacción, la dispersión de precios y el volumen de negociación. Finalmente, el tercer artículo de investigación explora el riesgo sistemático de sentimiento para la liquidez en relación con los datos de microblogs. Este estudio mostró que la liquidez del índice bancario estaba expuesta al riesgo sistemático de sentimiento y liquidez, pero la liquidez del índice de empresas no financieras solo estaba expuesta a un riesgo sistemático de liquidez.

Los participantes del mercado impulsados por los sentimientos observados en la plataforma de microblogging pueden no solo influir en la liquidez del mercado, que varía en el tiempo y sus dimensiones, sino que también pueden exponerse al riesgo sistemático para la liquidez dentro de un mercado más amplio. Por lo tanto, se sugiere que la liquidez y sus aspectos relacionados se valoren frente a los problemas de selección adversa en el mercado. Además, la medición de la información entrante en la plataforma de microblogging puede ayudar mejor a los proveedores de liquidez en la construcción de carteras.

Resum

Aquesta tesi doctoral s'emmarca en l'àrea d'investigació del Departament d'Economia I Ciències Socials, i es centra en la perspectiva conductual de la liquiditat del mercat. La liquiditat que varia en el temps i els seus problemes relacionats són una de les preocupacions dominants en la literatura de microestructura del mercat. El paper crític de la liquiditat del mercat en l'execució de transaccions o la determinació del rendiment de la inversió genera inquietuds tant per a acadèmics com per a aquells que participen en el mercat. Per tant, és necessari desvetlar els problemes potencials que poden afectar la liquiditat del mercat financer.

Aquesta tesi busca entendre la liquiditat del mercat i els seus problemes relacionats a la llum del comportament dels inversors. La perspectiva conductual de la liquiditat s'examina utilitzant informació orientada a opinions en microblogs. La creixent literature de finances conductuals també inclou l'autenticitat de les dades de microblogs tant en la modelització com en la predicció de diverses preocupacions associades amb el funcionament eficient dels mercats financers. No obstant això, la recerca prèvia en l'àmbit de les finances conductuals podria haver passat per alt algunes implicacions potencials de la informació orientada a opinions en microblogs sobre la liquiditat del mercat a nivell de mercat i d'empresa. Per tant, la tesi pretén ser una aplicació empírica en aquesta àrea d'investigació. La tesi es duu a terme com a compendi d'articles científics, la memòria de la qual inclou diversos articles de recerca publicats en revistes indexades.

El primer article proporciona informació sobre la relació entre el contingut de microblogs i el cost de facilitació de la liquiditat. Durant els períodes de negociació, aquest estudi va suggerir que l'estat d'ànim dels inversors tenia menys influència en afectar la liquiditat que varia en el temps i el seu cost de facilitació. No obstant això, la informació entrant en un dia donat era més influent per a les sessions de negociació següents. Els sentiments construïts sobre una base de dos dies estaven associats amb el cost de facilitació de la liquiditat. El segon article aborda les dimensions de la liquiditat del mercat utilitzant opinions de microblogs. Aquesta recerca va revelar que els sentiments dels inversors en entorns de pessimisme tenien més poder autoritari sobre les dimensions de la liquiditat, inclosos els costos de negociació, la immediatesa de la transacció, la dispersió de preus i el volum de negociació. Finalment, el tercer article de recerca explora el risc sistemàtic de sentiment per a la liquiditat en relació amb les dades de microblogs. Aquest estudi va mostrar que la liquiditat de l'índex bancari estava exposada al risc sistemàtic de sentiment i liquiditat, però la liquiditat de l'índex d'empreses no financeres només estava exposada a un risc sistemàtic de liquiditat.

Els participants del mercat impulsats pels sentiments observats a la plataforma de microblogging poden no només influir en la liquiditat del mercat, que varia en el temps i les seves dimensions, sinó que també poden exposar-se al risc sistemàtic per a la liquiditat dins d'un mercat més ampli. Per tant, es suggereix que la liquiditat i els seus aspectes relacionats es valoren davant dels problemes de selecció adversa en el mercat. A més, la mesura de la informació entrant a la plataforma de microblogging pot ajudar millor els proveïdors de liquiditat en la construcció de carteres.

Table of Contents

Tables

Figures

Chapter 1

1.1 Introduction

The escalation of behavior finance literature is ascribable to the authenticity of various information sources in both modeling and predicting financial assets. Among a wide range of opinion providers, it undoubtedly matters to unfold the authoritative role of most valuable information source on investors' sentiments. Sentiment debate falls in the area of natural language processing, that aids us to understand individual opinions using the binary emotion attributes (optimistic vs pessimistic), or multi-level quantitative results.

Sentiment analysis through text mining is a multidisciplinary discussion, but it can generally explain into the speculative behavior of information users on fundamental asset's value. The speculative behavior of investors is denoted into rational or irrational sentiments (Saleemi, 2023). Investors' behavior in the financial market can reasonably reflect asset price levels (Brown and Cliff, 2005). Therefore, the quantification of investor's opinions should be addressed in the asset pricing model.

A reasonable stream in the behavioral domain indicates the influence of opinionated-content on various variables associated with the efficient functioning of financial markets (Oliveira et al., 2017). Recently, the measurement of microblogging-opinionated information has also become a prevalent research topic in both modeling and predicting financial markets (Smailović et al, 2013; Sprenger et al., 2014; Poria et al., 2017; Li et al., 2018; Bank et al., 2019). Microblog networking may not only a cost-effective approach by eliminating geographical barriers, but it can spread the information in real-time basis compared to traditional opinion providers (Oliveira et al., 2017).

The authenticity of microblogging-opinionated information is addressed in mitigating the information asymmetry (Prokofieva, 2015), and alleviating the negative market reactions (Mazboudi and Khalil, 2017). The asset's value for investment concern determines its execution in the market (Cervelló-Royo and Guijarro, 2020), and opinionated-rumors regarding earning expectations can also impact the transaction execution (Chen et al., 2011; Zhang et al., 2022).

Existing literature links the microblogging-opinionated content to the behavioral perspective of prices, returns, volatility or trading quantity (Groß-Klußmann and Hautsch, 2011; Oliveira et al., 2017). However, previous research might have ignored the potential implications of microblogging data on market liquidity at both market and firm levels. Liquidity and its related dimensions are one of the dominant strands in the market microstructure literature.

Market liquidity can immediately determine traders' movement, as it is a crucial attribute of securities (Amihud and Mendelson, 1991). Liquidity is perceived to be a highly volatile risk in the market, which means that it can evaporate within minutes. Securities sensitive to information can trigger the liquidity risk (Glosten and Milgrom, 1985; Gorton and Metrick, 2010), and then the liquidity providers secure transaction against the risk of an informed trader (Saleemi, 2020). This risk is often perceived as a priced factor (Amihud et al., 2015).

Market liquidity is discussed in a multidimensional perspective, and there is currently no unified approach for its estimation in the market (Goyenko et al., 2009; Abdi and Ranaldo, 2017). Over time, several methods focusing either on cost-based liquidity or price impact volume-based liquidity have been introduced in the asset pricing literature. The bid-ask spread is a dominant strand of the cost-based liquidity measurement (Gregoriou, 2013), which captures the transaction speed at a possible trading cost (Roll, 1984; Corwin and Schultz, 2012). Another stream in the field estimates liquidity in light of the asset's price dispersion and its trading volume (Amihud, 2002). Trading cost, trading speed, price dispersion, and trading volume are crucial determinants of market liquidity (Le and Gregoriou, 2020).

The bid-ask spread not only indicates the speed of transaction in the market, but it also captures almost all costs associated with the transaction execution (Sarr and Lybek, 2002). These costs can compensate specialists against the inventory holding risk, informed counterparty, and order processing. Their impact in the market is suggested to be time-varying (DeGennaro and Robotti, 2007). Specialists enable continues trading by facilitating the immediacy of trade execution in the market. This activity can relate to the risk of holding inventory against the future price uncertainty, and then it compensates liquidity providers by imposing a cost on the counterparty, i.e., a spread.

With respect to the concept of information effects, the asymmetric information often relates to the spread. In the adverse selection phenomenon, there is a potential risk of loss for the uninformed trader. Therefore, the information-sensitive assets can be illiquid (Gorton and Metrick, 2010). In this debate, the market specialists tend to reduce their risk exposure, and then incline the spread as compensation against the informed transaction (Easley and O´Hara, 2004; Saleemi, 2020). The spread is closely linked with trading quantity due to the asymmetric information effects in the market (Le and Gregoriou, 2020). A small trading quantity declines the size of spread, which successively injects liquidity to the market and ameliorates price accuracy (Sarkissian, 2016).

1.2 Research Significance

A far-reaching role of market liquidity across the global financial markets has been under discussion since the economic crisis of 2007-2009. This crisis has undoubtedly raised fundamental concerns on liquidity risk, and its management. Due to the drastic repercussions of liquidity evaporation in funding and securities markets during a potential crisis, market liquidity has become a serious concern for both academics and those who engage in the trading.

In literature of market microstructure, market liquidity is addressed in a multidimensional concept over time (Goyenko et al., 2009), but most importantly, the adverse selection problem in the market is largely priced into the market liquidity (Saleemi, 2022). The provider of liquidity would be interested in the fundamental value of an investment, and as a result, avoids the risk of losing from informed counterparty.

Securities sensitive to information causes of adverse selection issues in the market, and therefore, the financial market can be perceived as illiquid (Gorton and Metrick, 2010). The adverse selection in the market can encourage informed optimistic traders to accept the financial position at a higher price, while informed pessimistic inventory-holder would redeem its position at a lower price. In this context, the asymmetry of information can be a risk factor for liquidity providers, and therefore, it is better to be priced into liquidity.

Financial news or surveys may be prominent conventional-sources of information to understand the stock market behavior. The recent research is also investigating microblogging-opinionated content for use in both predicting and modeling investors' behavior. Microblogging-opinionated information is reasonably applied to understand several factors associated with the efficient functioning of financial markets. However, there is still room to unfold the authoritative role of microblogging content on market liquidity and its related issues in the market. As there is no previous literature on how the microblogging-based investor sentiments may affect the behavior of liquidity providers in real-time, this PhD research aims to be the first empirical attempt to study this phenomenon.

Microblog networking can be a time-effective approach in terms of eliminating geographical barriers, and it is open to exchange information on financial markets and certain assets as a whole. The effect of this information in real-time can be relevant concern to understanding liquidity providers' movements, as well as their decision-making process. It is therefore essential to examine the far-reaching effects of microblogging-opinionated information on liquidity and its associated aspects in the market. In asset pricing studies, the novelty of this research is not only limited to methodological augmentation, but also extents investigations at both market and firm levels. This may help us to understand the systematic liquidity risk and its management in light of microblogging-opinionated content.

This research addresses the market liquidity as a behavior perspective of microblogging-based investor sentiments. The aim of this work is accomplished by exploring various research questions:

1. Is there a significant pattern between microblogging data and liquidity-facilitating cost?
2. What role may microblogging-opinionated data play in the estimation of liquidity dimensions?
3. Is there a systematic sentiment risk for liquidity due to the microblogging-based informed transaction?

Thesis, as demonstrated in Table 1, is produced as a compendium of articles, whose memory addresses the research questions in comprehensive manner. Article 1 investigates whether a positive or negative bias sentiment impacts the size of liquidity-providing cost in the financial market. This work helps us to understand the behavior of liquidity providers using the microblogging-opinionated information. Article 2 examines the liquidity dimensions as a behavior phenomenon of investor sentiments. In this context, the microblogging data is analyzed as a priced factor in various dimensions of liquidity. Finally, Article 3 provides insights into the systematic sentiment risk for liquidity within a broader market. This work seeks to unveil whether the liquidity of individual assets is exposed to the systematic sentiment risk.

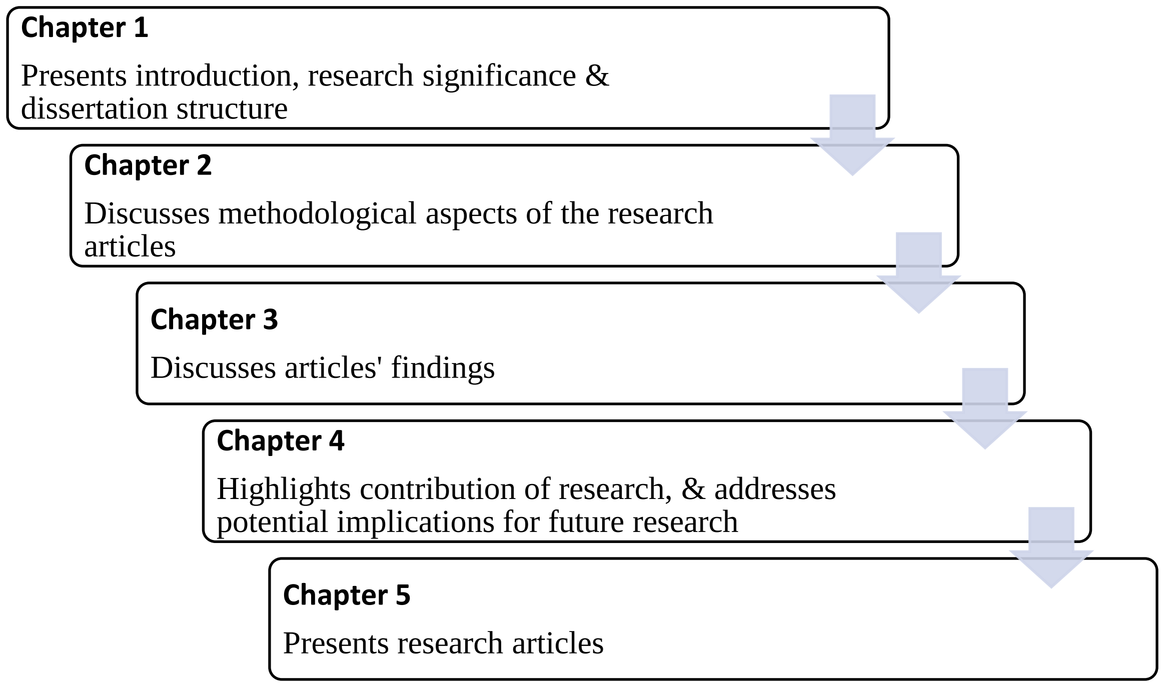
Table 1. Outline of the dissertation.

Articles	**Research Questions**	**Contributions**
Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index **DOI**: 10.3390/su11247048 **Journal**: Sustainability **Status**: Open-accessed publication	Is there a significant pattern between microblogging data and liquidity-facilitating cost?	I contributed to all aspects, including conceptualization, methodology, software, resources, data curation, writing-original draft preparation, and writing-review and editing.
Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data **DOI**: 10.3390/jrfm14090394 **Journal**: Journal of Risk & Financial Management **Status**: Open-accessed publication	What role may microblogging-opinionated data play in the estimation of liquidity dimensions?	I contributed to all aspects, including conceptualization, methodology, software, resources, data curation, writing-original draft preparation, and writing-review and editing.
Investor Sentiments and Liquidity Pricing: Applying the Microblogging Content to the Systematic Risk **DOI**: **Journal**: **Status**:	Is there a systematic sentiment risk for liquidity due to the microblogging-based informed transaction?	I contributed to all aspects, including conceptualization, methodology, software, resources, data curation, writing-original draft preparation, and writing-review and editing.

1.3 Structure of the Dissertation

This dissertation, as shown in Figure 1, comprises five chapters. The concept of the dissertation is introduced and structured in the first chapter. This chapter addresses the significance of dissertation along with the adopted research questions. The methodological aspect of research is presented in the second chapter. This chapter highlights the data curation and methods, that are applied in the corresponding research article. The third chapter deals with the results of dissertation. The fourth chapter addresses the significant contribution of this dissertation in the area of research, and underlines the potential implications for future research. Finally, the last chapter of this dissertation presents the scientific work, whose memory includes the articles published in the indexed journals. The published articles are presented in their original format.

Figure 1. Structure of the research.



Chapter 2

Material and Methods

This chapter, as briefly addressed in Table 2, deals with the methodological aspect of each article, that is produced within the domain of dissertation. Some of key concerns related to the methodological approach in each article are concisely addressed in this section.

The first article examines the influential role of microblogging-opinionated data on liquidity, as well as on its facilitating cost. Based on the theoretical foundation of liquidity, the study applied various liquidity measures including (a) high-low difference, (b) Glosten-Milgrom spread, (c) quoted spread, and (d) effective spread. Microblogging-opinionated information is collected using the libraries, i.e., "ROAuth", "rtweet", on R programming language, and then processed through "TM library", "Syuzhet library", "Lubridate library". This process helped to transform the data into a positive or negative value. Such quantification guides the importance of each text with positive scores indicating optimistic sentiment and negative scores indicating pessimistic sentiment on a given trading day. To understand the authoritative role of microblogging data on financial market liquidity and its facilitating cost, the variables are examined using the ordinary least squared technique, multiple linear regression, lagged analysis, autocorrelation analysis, Kolmogorov-Smirnov test, and Breusch-Pagan test.

The second article unveils the authoritative role of microblogging-based investor sentiments on liquidity dimensions including (1) trading speed, (2) trading cost, (3) dispersion of price, and (4) trading quantity. The cost-based market liquidity (CBML) model was applied to estimate the trading speed and its associated cost, whereas the price impact volume-based liquidity technique, i.e., Martin Liquidity Index (MLI), captured the liquidity in light of price dispersion compared to trading quantity. The sentiment indicators, i.e., Negative Ratio (NR) and Positive Ratio (PR), were constructed using the structured form of microblogging data. This process included the cleaning of data through the "TM library", and then converting the data in either a bullish sentiment score or bearish sentiment score. As the number of quantified data on a single day is very large, the sentiment scores for given trading day is aggregated for construction of sentiment indicators. This process was performed through the "Syuzhet", and "Lubridate" libraries on R programming software. To unveil the impact of microblogging-based investor sentiments on liquidity dimensions, the variables were explored using the multiple linear regression, vector autoregression, Jarque-Bera test, Autoregressive conditional heteroscedastic test, and forecast error variance decomposition test.

The third article investigates the systematic risk by exploring the potential of microblogging-opinionated content in determining the cost-based market liquidity for individual assets and their respective markets. In this context, an index of banks and non-financial firms were built using the weighted market capitalization technique. Microblogging text was first cleaned using the "tm library ". This process comprised removing punctuation, stop words, trailing spaces, and converting the text to lowercase. The structured data was classified into bullish or bearish sentiments. Given the large volume of data for trading day, the bullish sentiments were aggregated for sentiment analysis. The same aggregation technique was applied for construction of bearish sentiment indicator. To gain insights into the systematic sentiment risk for liquidity within a broader market, the variables are examined using the multiple linear regression, vector error correction model, weighted market capitalization technique, impulse response analysis, Augmented Dickey-Fuller test, and Johansen trace cointegration test.

Table 2. Summary of methodological aspect.

Articles	**Material & Methods**
Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index	**Sentiment tools**: TM library, ROAuth library, rtweet library, Syuzhet library, Lubridate library. **Liquidity proxies**: S, GMS, QS, ES. **Methods**: OLS, MLR, Lagged analysis, Autocorrelation, KS test, BP test. **Software**: R programming language
Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data	**Sentiment tools:** TM library, ROAuth library, rtweet library, Syuzhet library, Lubridate library. **Liquidity proxies**: CBML, PIVBL. **Methods**: MLR, VAR, JB test, ARCH test, FEVD test. **Software**: R programming language
Investor Sentiments and Liquidity Pricing: Applying the Microblogging Content to the Systematic Risk	**Sentiment tools:** TM library, ROAuth library, rtweet library, Syuzhet library, Lubridate library. **Liquidity proxy**: CBML. **Methods**: MLR, VECM, WMCT, IRA, ADF test, JTC test. **Software**: R programming language

Notes: Text mining: TM; Spread: S; Glosten-Milgrom Spread: GMS; Quoted spread: QS; Effective spread: ES; Cost-based market liquidity: CBML; Price impact volume-based liquidity: PIVBL; ordinary least squared technique: OLS; Multiple linear regression: MLR; Kolmogorov-Smirnov: KS; Breusch-Pagan: BP; Vector autoregression: VAR; Jarque-Bera: JB; Autoregressive conditional heteroscedastic: ARCH; Forecast error variance decomposition: FEVD; Vector Error Correction Model: VECM; Weighted market capitalization technique: WMCT; Impulse response analysis: IRA; Augmented Dickey-Fuller: ADF; Johansen trace cointegration: JTC.

Chapter 3

Findings of the Dissertation

This chapter deals with the articles' findings. As final chapter comprises the presentation of research articles in original published-format, results of the dissertation are concisely discussed in this section.

In the first article, the investors' mood extracted from microblogging data and several liquidity proxies were applied to investigate the relationship between liquidity-facilitating cost and investors' opinions. The findings, based on analysis of the market, i.e., S&P500 Index, suggested that investors' mood had little authoritative power to estimate liquidity or its facilitating cost during trading periods. This guides us to understand, that any changes in investors' perceptions are very weaky associated with the time-varying variations in liquidity and its related cost. In other words, a much lower proportion of changes in investors' perceptions influences the variations in liquidity and its facilitating cost. These findings discount the relation between investors' mood and liquidity-facilitating cost on a daily basis. This means, that investors' sentiments constructed from microblogging data were not intensively associated with the liquidity or its facilitating cost on a given trading day. The opinionated-information on social media may guide investors, not necessarily on the same trading day, but also in the following trading sessions. In this debate, the concept of two-day moving average helped us to understand whether the liquidity or its related cost responds to the investors' mood over a period of time. on a two-day basis, the liquidity or its associated cost was significantly explained by investors' mood. This analysis provided a new insight into the association between these variables.

In the second article, the sentiment indicators constructed from microblogging content and measures of liquidity dimensions were applied to examine the association between liquidity dimensions and investor sentiments. The multivariate analysis was applied to disentangle various aspects involved in this area of research. The results, based on analysis of the market, i.e., Australian Securities Exchange (ASX), found that the pessimistic investor sentiments had higher influence on liquidity dimensions. However, there was no association between microblogging-based optimistic sentiments and liquidity dimensions. During trading sessions, the pessimistic investor sentiments led to a higher spread. The higher spread indicated an unwillingness of liquidity suppliers to accept the financial holding without imposing a higher cost on counterparty. Therefore, a higher trading cost influenced the speed of transactions and shrinks the market liquidity. Additionally, a higher dispersion of asset's price relative to its trading quantity was significantly explained by the pessimistic investor sentiments during the trading sessions. This means, that investors acquired a smaller number of trades and therefore, the liquidity declined in the pessimistic periods. Other than the same trading session, the study also analyzed whether the liquidity dimensions on a given trading day was linked to the past time series of investor sentiments. The previous sentiments' series were not associated with the trading speed and its related cost on a given trading day. Similarly, the dispersion of asset price and its trading quantity was not significantly explained by past series of microblogging-based optimistic sentiments. However, the price impact volume-based liquidity was significantly explained by lag_{t-1} of the pessimistic sentiments.

In the third article, the microblogging-based informed transaction was checked as a systematic sentiment risk for liquidity, not only for individual stocks, but also their respective markets. The analysis was performed on Financial Times Stock Exchange 100 Index (FTSE), where two sub-indices were built: one for banks and another for non-financial firms (NFF). Sub-indices in the market were constructed using the capitalization weighted average approach. An incline in the pessimistic opinions leads to a wider trading cost of the market index. This means, that the liquidity suppliers were more hesitant to accept financial position without imposing higher costs on seller in environments of pessimism. Therefore, negative opinions were priced into the overall market liquidity during trading periods. However, market index liquidity was not significantly explained by the microblogging-based positive opinions. The following experiment explores the presence of systematic sentiment and liquidity risks at the individual stock level. In this context, the assets are divided into financial and non-financial sectors, and then corresponding indices are built. The behavioral analysis of financial sector indicated, that the systematic sentiment risk appeared to be priced in the bank index liquidity. The financial index liquidity was also positively associated with the liquidity of its corresponding market index. This implies, that the bank index liquidity was exposed to the systematic liquidity risk. The behavioral analysis of non-financial sector revealed, that the systematic sentiment risk was not priced in the NFF index liquidity. Meantime, the non-financial index liquidity was positively explained by liquidity of its corresponding market index. Therefore, the NFF index liquidity was exposed to the systematic liquidity risk. Additionally, changes in the trading cost on a given day was examined as function of past changes in sentiment series. The analysis of market index unveiled, that changes in market index liquidity on a given trading day were not associated with past changes in the sentiment indicators, either in the short or long run. Similarly, changes in NFF index liquidity on a given trading day were not correlated with past changes in sentiment series, either in the short or long run. However, changes in bank index liquidity were not significantly explained by changes in previous sentiment series, except for lag_{t-4} of bearish sentiments. This guides, that the bank index liquidity was exposed to the pessimistic sentiments in the long run.

Chapter 4

Conclusion

This chapter concludes the dissertation with highlighting contribution of the research, as well as potential implications for future research.

This doctoral dissertation seeks to unveil the authoritative role of microblogging-based investor sentiments on the financial market liquidity and its related aspects, that previous studies might have ignored. In this context, the study addresses a few potential research questions including (i) Is there a significant pattern between microblogging data and liquidity-facilitating cost?, (ii) What role may microblogging-opinionated data play in the estimation of liquidity dimensions?, and (iii) Is there a systematic sentiment risk for liquidity due to the microblogging-based informed transaction?.

Firstly, the work investigated the pattern between microblogging-opinionated content and liquidity-facilitating cost. The findings revealed, that investors' mood had little authoritative power to estimate the liquidity or its facilitating cost during trading periods. However, investors' mood on social media had more authoritative power to impact the liquidity and its related cost on a two-day basis. These results helped us to understand, that the microblogging-opinionated data may attract the investment interest, not necessarily on the same trading day, but also in the following trading sessions.

Secondly, the work extended into the understanding of relationship between microblogging-based investor sentiments and liquidity dimensions. The analysis unveiled, that investor sentiments in the pessimistic periods had higher impact on liquidity and its dimensions. During trading sessions, the microblogging-based pessimistic sentiments led to higher trading costs, lower liquidity, larger price dispersion, and lower trading volume. A higher trading cost in environments of pessimism impacted the speed of transaction execution, and then declined liquidity in the market. A larger price dispersion relative to the trading quantity indicated, that investors acquired a small amount of trades in environments of pessimism, and thereby, the liquidity declined in the market. During trading sessions, these findings guided us to price the dimensions of liquidity against the investor sentiments extracted from microblogging data.

Finally, the systematic sentiment risk for liquidity was investigated in light of microblogging-based informed transaction, not only for individual assets, but also their respective markets. The analysis reported, that the market index liquidity was priced pessimistically during trading periods. Meantime, systematic sentiment and liquidity risks were associated with the bank index liquidity. Nevertheless, systematic sentiment risk was not noted in the non-financial firm index liquidity, but a systematic risk for liquidity was observed in the non-financial firm index. Additionally, bank index liquidity was exposed to the pessimistic sentiments in the long run. Short-run relevance was depicted in the commonality of liquidity between firm index and market index. The systematic sentiment risk should be priced in liquidity within a broader market. This may help to manage the liquidity risk and its related issues in a comprehensive manner.

In this dissertation, the concept of market liquidity is potentially covered in relation to the transparency of information on assets' value over time. The higher the risk of adverse selection issue in market, more a transaction is perceived as illiquid. As accuracy of information about asset's fundamental value can be a genuine concern in the investment, the facilitation of liquidity seems a priced risk factor against the informed trader.

The findings considerably suggest cavernous applications of microblogging behavioral perspective in the estimation of market liquidity and its related aspects. Microblogging opinions' authenticity is not only applicable in the estimation of time-varying liquidity, but it is also a crucial key in the measurement of liquidity dimensions including the trading costs, transaction immediacy, price dispersion and trading volume. Additionally, the microblogging-opinionated data matters to be addressed as a systematic sentiment risk for liquidity due to the respective markets for individual assets.

Beside conventional sources of behavioral perspective, the assessment of microblogging opinions can also assist the liquidity suppliers in the liquidity risk management and portfolio construction. The investment interest on microblogging platform can eliminate the adverse selection issues in the market, and as a result, incoming information on asset's value becomes more transparent. More the incoming information on microblogging platform is addressed by users of financial liquidity, more the investment can be perceived as liquid.

Based on findings of this doctoral dissertation, other researchers are encouraged to make additional efforts by including new markets for behavioral analysis of market liquidity in relation to the microblogging data. Although this dissertation fills a potential gap in the behavioral finance literature, the geographical dataset may be concluding in a limited sense that the microblogging-opinionated information had influence on liquidity and its related issues in the market. In other markets, there is thus a considerable need to explore the behavioral perspective of liquidity and its related aspects using the microblogging-opinionated information. The future research may also investigate the relationship dynamics between microblogging data and financial market liquidity, particularly in light of various events including the global pandemic uncertainty, sovereign default in Sri Lanka, Russia-Ukraine war, or a historic economic and political unrest in Pakistan.

References

Abdi, F., & Ranaldo, A. (2017). A simple estimation of bid-ask spreads from daily close, high, and low prices. *The Review of Financial Studies, 30*(12), 4437–4480. doi:10.1093/rfs/hhx084

Amihud, Y. (2002). Illiquidity and stock returns cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56. doi:10.1016/S1386-4181(01)00024-6

Amihud, Y., & Mendelson, H. (1991). Liquidity, maturity, and the yields on U.S. treasury securities. *The Journal of Finance, 46*(4), 1411–1425. doi:10.1111/j.1540-6261.1991.tb04623.x

Amihud, Y., Hameed, A., Kang, W., & Zhang, H. (2015). The Illiquidity Premium: International Evidence. *Journal of Financial Economics, 117*(2), 350–368. doi:10.1016/j.jfineco.2015.04.005

Bank, S., Yazar, E. E., & Sivri, U. (2019). Can social media marketing lead to abnormal portfolio returns? *European Research on Management and Business Economics, 25*, 54-62. doi:10.1016/j.iedeen.2019.04.006

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business, 78*(2), 405-440. doi:10.1086/427633

Cervelló-Royo, R., & Guijarro, F. (2020). Forecasting stock market trend: a comparison of machine learning algorithms. *Finance, Markets and Valuation, 6*(1), 37–49. doi:10.46503/NLUF8557

Chen, H., De, P., Hu, Y., & Hwang, B. H. (2011). Sentiment revealed in social media and its effect on the stock market. *IEEE Statistical Signal Processing Workshop (SSP)*, 25–28. doi:10.1109/SSP.2011.5967675

Corwin, S. A., & Schultz, P. (2012). A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *The Journal of Finance, 67*(2), 719-760. doi:10.1111/j.1540-6261.2012.01729.x

Degennaro, R. P., & Robotti, C. (2007). Financial Market Frictions. *Economic Review, 92*(3), 1-16.

Easley, D., & O'Hara, M. (2004). Information and the cost of capital. *The Journal of Finance, 59*(4), 1553-1583. doi:10.1111/j.1540-6261.2004.00672.x

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*, 71–100. doi:10.1016/0304-405X(85)90044-3

Gorton, G., & Metrick, A. (2010). Haircuts. *Federal Reserve Bank St Louis Review, 92*(6), 507–520. doi:10.20955/r.92.507-20

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*(2), 153–181. doi:10.1016/j.jfineco.2008.06.002

Gregoriou, A. (2013). Earnings announcements and the components of the bid-ask spread: evidence from the London stock exchange. *Journal of Economic Studies, 40*(2), 112–126. doi:10.1108/01443581311283646

Groß-Klußmann, A., & Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Empirical Finance, 18*(2), 321-340. doi:10.1016/j.jempfin.2010.11.009

Le, H., & Gregoriou, A. (2020). How do you capture liquidity? A review of the literature on Low-frequency stock liquidity. *Journal of Economic Surveys, 34*(5), 1170-1186. doi:10.1111/joes.12385

Li, Q., Chen, Y., Wang, J., Chen, Y., & Chen, H. (2018). Web media and stock markets: A survey and future directions from a big data perspective. *IEEE Transactions on Knowledge and Data Engineering, 30*(2), 381–399. doi:10.1109/TKDE.2017.2763144

Mazboudi, M., & Khalil, S. (2017). The attenuation effect of social media: Evidence from acquisitions by large firms. *Journal of Financial Stability, 28*(C), 115-124. doi:10.1016/j.jfs.2016.11.010

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: using twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications, 73*, 125-144. doi:10.1016/j.eswa.2016.12.036

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion, 37*, 98–125. doi:10.1016/j.inffus.2017.02.003

Prokofieva, M. (2015). Twitter-based dissemination of corporate disclosure and the intervening effects of firms' visibility: Evidence from Australian-listed companies. *Journal of Information Systems, 29*(2), 107-136. doi:10.2308/isys-50994

Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance, 39*(4), 1127-1139. doi:10.1111/j.1540-6261.1984.tb03897.x

Saleemi, J. (2020). An estimation of cost-based market liquidity from daily high, low and close prices. *Finance, Markets and Valuation, 6*(2), 1-11. doi:10.46503/VUTL1758

Saleemi, J. (2022). Asymmetric information modelling in the realized spread: A new simple estimation of the informed realized spread. *Finance, Markets and Valuation, 8*(1), 1–12. doi:10.46503/JQYH3943

Saleemi, J. (2023). Microblogging Perceptive and Pricing Liquidity: Exploring Asymmetric Information as a Risk Determinant of Liquidity in the Pandemic Environments. *Economic Analysis Letters, 2*(1), 1–9. doi:10.58567/eal02010001

Sarkissian, J. (2016). Option pricing under quantum theory of securities price formation. *SSRN Electronic Journal*. doi:10.2139/ssrn.2848014

Sarr, A., & Lybek, T. (2002). Measuring liquidity in financial markets. *International Monetary Fund, 2*, 1–64. doi:10.5089/9781451875577.001

Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2013). Predictive sentiment analysis of Tweets: a stock market application. *In Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*, 77-88. doi:10.1007/978-3-642-39146-0_8

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: the information content of stock microblogs. *European Financial Management, 20*(5), 926-957. doi:10.1111/j.1468-036X.2013.12007.x

Zhang, H., Chen, Y., Rong, W., & Wang, J. (2022). Effect of social media rumors on stock market volatility: A case of data mining in China. *Frontiers in Physics, 10*, 987799. doi:10.3389/fphy.2022.987799

Chapter 5

Research Articles

This chapter deals with the research articles, that are produced within the domain of doctoral dissertation. The presentation of research articles is included in the respective published-format, and arranged as,

1) Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index.

2) Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data.

3) Investor Sentiments and Liquidity Pricing: Applying the Microblogging Content to the Systematic Risk.

Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index

Abstract: Microblogging services can enrich the information investors use to make financial decisions on the stock markets. As liquidity has immediate consequences for a trader's movements, this risk is an attractive area of interest for both academics and those who participate in the financial markets. This paper focuses on market liquidity and studies the impact on liquidity and trading costs of the popular Twitter microblogging service. Sentiment analysis extracted from Twitter and different popular liquidity measures were gathered to analyze the relationship between liquidity and investors' opinions. The results, based on the analysis of the S&P 500 Index, found that the investors' mood had little influence on the spread of the index.

Keywords: social media; opinion mining; financial market liquidity; sentiment analysis; trading costs

1. Introduction

Sentiment analysis is a field of natural language processing (NLP) that aids in understanding and extracting different opinions on a given subject. As 2.5 quintillion bytes of data are generated daily all over the world by the participants of social media, sentiment analysis tools can be used to make sense of the data. While opinion mining categorizes opinions into positive or negative, sentiment analysis is a field of interest for both academics and practitioners, since the constantly expanding social networks enable the exchange of information and opinions on products, services or any other subject. Sentiment analysis not only transforms such unstructured public information into structured data, but also makes it possible to apply the data to various areas, such as customer feedback, product/service reviews, net promoter scoring, stakeholder relations, marketing, financial market predictions, or almost any other field.

Social media are cost-effective and easily available networks that exchange information for both private and business purposes while eradicating geographical barriers. They are a broad source of co-creation values in which the participants contribute to evaluating and refining conceptualizations [1].

Social media can be divided into six types [1]: Social networking sites, blogs and micro blogs, collaborative projects, virtual game worlds, content communities, and virtual communities. This categorization has contributed a great deal to expanding the literature of social media into diverse notable fields, such as polling estimation [2,3], tourism [4], medicine and healthcare [5,6], collaborative learning [7], social participation [8], sport [9], communication [10], organizing [11], recruiting/selection decisions [12], crisis event analysis [13–15], public-spending review [16], and stock market predictions related to returns, prices, volatility, and trading volume [17–19].

To the authors' knowledge, however, the impact of social media on financial market liquidity and thus on trading costs has not been explored. Market liquidity facilitates the efficient and stable functioning of financial markets [20]. It is a multi-dimensional concept, generally referring to the immediacy of the execution of a trade with a limited price impact and low transaction costs [21].

While various of its aspects can be studied, this study investigates whether the Twitter microblogging social network influences liquidity.

The financial crisis of 2007–2009 crucially underlined the significance of liquidity on the functioning of the global financial markets [20]. The severe trading losses led the major players in the financial system to reassess their risk profiles and business models, which was considered a major step towards the implementation of rigorous regulatory reforms throughout the financial sector. Policy makers emphasized the need to constrain the banks' riskier business lines, such as investment banking and trading [22]. There has been a measurable reduction in banks' trading capacity: Bank holdings of trading assets have dropped by more than 40% between 2008 and 2015. However, concerns are growing on their willingness and ability to take risks as market makers or whether they would withdraw abruptly in cases of liquidity distress.

Financial market liquidity can be suddenly reduced for several reasons. Firstly, it depends in part on the transparency of information on security values, which vary over time. Secondly, the number of liquidity providers and their access to capital is a significant determinant of market liquidity. Thirdly, increased uncertainty about market liquidity makes the provision of liquidity riskier and increases the compensation that liquidity providers demand, i.e., the trading cost (bid–ask spread) increases.

The existing literature argues that asymmetric information is one of the factors in determining the liquidity and trading cost [23–26]. Social media have constructed a diversified structure of social networks, in which the participants, irrespective of their professional background, are open to exchanging information on financial markets and certain securities as a whole. It is therefore essential to investigate the impact of such information and opinions on financial market liquidity.

The aim of this paper is thus to analyze the impact of social media on financial market liquidity. This analysis can have potential implications for both academics and investors in terms of quantifying social media-based sentiments towards financial market liquidity and trading costs.

The rest of the paper is structured as follows. A brief survey of the literature is included in Section 2. A description of the data collected and different liquidity measures is given in Section 3. The research results are presented and discussed in Section 4 and the paper ends with the main conclusions highlighted in Section 5.

2. Review of the Literature

Market participants evidently generate spaces within products, services and firms by reflecting opinions and concerns [27], which further develops the field of communications [28]. Social media have not only contributed to revolutionizing reciprocity between stakeholders and businesses, but have also changed the approach of business-related content with regard to production, distribution and consumption [27]. Firms have cost-effective opportunities to communicate and build a strong relationship with stakeholders through social networking [29] without distance and geographical barriers [30].

Firms' engagement in social media can build up direct relationships with customers, expand business by identifying new opportunities, create their product-related communities, collect opinions and concerns and generally improve gaps [31]. Due to their non-transactional nature, social media are well suited to collecting feedback from a very large potential audience, initiating two-way communications and developing relationships with customers through interaction [1] and are therefore a cost-effective source of targeting an immense audience and gathering large volumes of feedback.

Twitter, as a microblogging social network, is one of the most popular communication sources through which participants can interact globally with messages known as 'tweets'. Besides facilitating private communications between the participants, Twitter also distributes information on professional contexts, for example financial market-related groups, such as StockTwits, Financial Times, Market-Watch, etc. That prompt investors to share their opinions on investment in financial markets and certain securities [32]. There are around 313 million active users of Twitter who interact with tweets in more than 40 languages. The influential role of Twitter has been revealed in various fields such as election results and political debates [33,34], academic communications [35], brand reputations [36], stock volatility, returns and volumes [18], and portfolio returns [19].

Financial market analysis is one of the most attractive areas in the literature on market microstructures—it is concerned with the details of how exchanges occur in the market by means of various theories. The chartist theory suggests that patterns and trends of its past behaviour tend to recur in the future and provides future asset prices based on historical data. The random walk theory argues that asset prices evolve randomly and cannot be predicted from historical patterns and stock market trends, so that prices are identical independent variables.

Besides such assumptions, variations in asset prices can be influenced by the media [37–41]. Public opinions through social media can significantly influence the investment decision-making process [42] and have an impact on the financial market. The social media have been extensively examined by researchers in order to determine the state of the financial market at both the indicator and firm levels [17,18,43–45]. These studies use sentiment analysis tools to extract opinions and information in terms of binary sentiment results (positive or bullish vs negative or bearish) or multi-level sentiment results. Some recent works are focused on the relationship between social media and market behaviour. A good example is [46], who analyze how social media impact financial markets, which is different if we compare it with traditional sources of investor attention such as newspapers, analyst coverage, earnings announcements, and business news wires. They show that "increases in Twitter activity are associated with positive abnormal returns and, when occurring in conjunction with traditional information supply events, increase the diffusion of information to investors". Authors also introduce an interesting difference between the supply of information (Tweeter activity) and the consumption of information (retweet activity), and show that the consumption of information increases the magnitude of the price impact. In a similar way, reference [47] study whether social media can provide new insights on market panics and manias that are not already captured by traditional data. They show that highly abnormal social media sentiment—as measured by Twitter and StockTwits messages—is preceded by very strong momentum and followed by mean-reverting return. Authors design a strategy based on this mean-reverting effect which outperforms a benchmark mean-reversion strategy that does not use social-media data. Other areas where social media are related with financial markets include IPO performance [48], information asymmetry [49], market manipulation [50] and communication of financial information [51].

As previously stated, investors' sentiments may be reflected by the financial markets [17,37,40]. Positive sentiments cause asset prices and returns to rise, while negative sentiments may reduce them. It is therefore of interest to apply sentiment analysis tools to investors' sentiments, which not only show investors' emotions based on their perceptions but also investigate their impact on financial market forecasting.

The existing literature often uses Pearson correlation coefficients [32,45,52] and beta coefficients of linear regression models [17,44,53] to examine the relationship between the financial market and investors' sentiments. Most of the existing literature indicates the positive relationship between social media and the financial market [17,32,44,52,54], but limited to certain aspects such as prices, returns, volatility or trading volume [17–19,32,37,42–45,52,53].

Liquidity is a time-varying risk factor [55]. The risk arises in situations in which a share cannot be traded quickly enough to prevent or minimize a loss. The liquidity risk, in general, is considered the centre of any financial crisis [20]. Liquidity tends to be highly volatile, which means that it can vanish within minutes. It has become an important issue for traders and can even cause a systemic risk. Due to the severe consequences of an evaporation of liquidity in securities and funding markets during a potential financial crisis, the systemic liquidity risk should be closely monitored [20].

Market participants who seek to make an immediate trade would possibly trade at the best available price, i.e., the bid price if buying or the asking price if selling. The bid–ask spread has gained huge interest among market participants due to the fact that it is a significant measure of trading costs and thus a proxy for financial market liquidity [20,56–59]. The size of the spread reflects an asset's liquidity, i.e., the ease and cost of trading an asset.

The literature concentrates on three factors—adverse selection costs [23–26], inventory holding costs [56,60–62] and order processing costs [63]—in order to determine the bid–ask spread. Reference [57] developed a three-way decomposition model by combining the spread components—order processing, adverse selection and inventor holding costs—and disclosed the significance of these components in estimating the true spread.

When securities become information-sensitive, the financial markets are not perceived as liquid [64]. This causes an adverse selection problem: informed optimistic investors would buy an asset even at a higher ask-price, while pessimistic sellers have an incentive to sell at a lower bid-price. Traders with private information on the fundamental value of securities would consider the price effect of their trades, and market makers are likely to protect themselves against informed traders, so that reduced liquidity produces a wider bid–ask spread.

In addition to private information on the fundamental value of an asset, the literature also illuminates the significance of private information on order flows [65,66]. For example, if a trading desk foresees that a hedge fund will liquidate a huge position which will likely depress prices, then the trading desk will sell early while the price is high and buy back later at a lower price. Informed buyers have incentives in large trades, which increase dealers' potential losses so that dealers would widen the spread.

The bid–ask spread is a compensation for dealers who offer immediacy while accepting the risk of holding an inventory. Dealers are risk aversion agents that facilitate liquidity in the market while optimizing their own security portfolios. In fact, all the buyers are not present in the market at all times, so this gap between buyer and seller is bridged by market makers, who may buy a security in anticipation of being able to resell it to the buyer. However, market makers take into account the risk of price fluctuations in the meantime and would be compensated for this risk in terms of imposing a cost on the seller, i.e., a higher spread. Additionally, the spread compensates dealers who offer immediacy by bearing some of the fixed costs. Consistent with the empirical literature, reference [67] showed that the bid–ask spread is a positive function of the price level and return variance, a negative function of measures of market activity, depth, and continuity, and negatively correlated with the degree of competition. The illiquidity premium was documented for the equity market in [56], while [68] measured the effective bid–ask spread by using the first-order serial covariance of price changes. Later, reference [25] developed a technique for estimating a model that decomposed the bid–ask spread into two components, one due to asymmetric information and one due to inventory costs, specialist monopoly power, and clearing costs.

3. Data Sampling and Methodology

This paper studies the impact of microblogging data (tweets) on the market liquidity of the S&P500 Index, and as a result, on the transaction cost associated with trading, that is the bid-ask spread. This not only captures the trading cost, but is also a true measure of actual market liquidity, which can be measured in various ways. Based on the theoretical foundation of market liquidity, we apply various liquidity measures including (1) high–low difference, (2) spread derived by [25], (3) quoted spread, and (4) effective spread.

The liquidity measures applied in this study are standard and have previously been examined in different aspects of the asset pricing literature, although all liquidity measures and indices are in fact the proxies for illiquidity [69]. Investors take the significance of financial market liquidity into consideration at the time of decision-making because it is a great indicator of the efficiency of financial markets. The literature on the market microstructure proposes and constructs the bid–ask spread in several ways. The bid–ask spread simply is defined as the difference between the seller's asking price for an asset and the bid price offered by the buyer. The high or ask price refers to the highest price during the trading day, whereas the low or bid price is defined as the lowest price during the same day. A spread can be computed by using the daily high and low prices, which is given by:

$$S = High_t - Low_t \tag{1}$$

An alternative liquidity measure of daily high and low prices was derived by [25], who considers that any transaction discloses something about a trader's private information. The bid–ask spread was modelled in the following manner:

$$GMS = Ask_t - Bid_t \tag{2}$$

The expected value of the security conditional on a trade at: $Ask_t = v^H\pi + \bar{v}(1-\pi)$, and where bid price is assumed by: $Bid_t = v^L\pi + \bar{v}(1-\pi)$. v^H and v^L are high and low possible values, respectively, for an asset with equal probability. An informed optimistic trader is present with probability π. Assuming risk neutrality, uninformed pessimistic traders value the security at $\bar{v} = \frac{(v^H+v^L)}{2}$. The model assumes that the spread would be greater in case of a higher probability of trading with an informed trader.

The quoted spread (QS) and effective spread (ES) are the most common measures of market liquidity and significantly explain the spread context. The quoted spread is simply the difference between the ask (high) quote and the bid (low) quote at a given time in the market, divided by the average of the two quotes (mean of $High_t$ and Low_t):

$$QS = \frac{High_t - Low_t}{M_t} \tag{3}$$

By taking into consideration the hidden orders, order internalization by market-makers, the effective spread is considered a leading measure of financial market liquidity, which is defined as the absolute value of the difference between the trade price, P_t, and the midpoint of the quotes, $M_t = \left(\frac{High_t + Low_t}{2}\right)$, divided by the mean of $High_t$ and Low_t:

$$ES = \frac{2|P_t - M_t|}{M_t} \tag{4}$$

The data used in this research was obtained from the Center for Research in Security Prices, which contains daily observations of high, low, and closing prices. The dataset therefore has both time and individual dimensions. To understand the impact of microblogging data on financial market liquidity and trading costs, the study took into consideration tweets, consisting of 23,008 observations, related to the S&P500 Index and collected daily during the period 3 July 2019–1 October 2019. The analysis was carried out on R programming software, in which the machine learning strategy and a linear regression model are applied to disentangle the various aspects involved. The machine learning strategy was used to extract the aggregated sentiments, while the regression model used sentiments as the independent variable and the abovementioned liquidity measures as the dependent variables. Hence, we used four regression models to analyze the link between investors' mood and liquidity of the S&P Index as shown in Equation (5):

$$Liquidity_t = \alpha + \beta_1 Sentiments_t + \epsilon_t \tag{5}$$

where Liquidity_t corresponds to each liquidity measure of Equations (1)–(4) in t, Sentiments_t represents the extracted sentiment for that period, and ϵ_t is the error term. The regression model was performed by using the ordinary least squared technique (OLS), and no control variables were included in the process.

In order to construct the 'Sentiments' variable from unstructured data (tweets), we executed some pre-processing tasks using the software R. At the first stage, R served to clean each tweet by removing punctuation and stop words, converting words into lower case, striping any leading or trailing spaces and for privacy reasons, setting all participants' addresses into '@user'. At the final stage, R categorizes each structured tweet into a numerical positive or negative value between −5 and 5 that defines the importance of a tweet with positive scores indicating positive sentiment and negative scores indicating negative sentiment on a given day. For example, the most positive sentiment got a score of 4.75 from the tweet "Watch it! An amazing truth. How the brain works. If you like it Share it Please Awesome information SP&500 #amazing #unbeliev", and the most negative sentiment was valued at −3.9 from the tweet "In a pathetic attempt to avoid panic selling S&P; a stock market crash, the Trump administration wasted $130 billion of tax". Figure 1 shows the distribution of each tweet sentiments throughout the analysed period. The Box-plot distribution indicates both positive and negative sentiments, in which most of the market participants can be viewed as bullish.

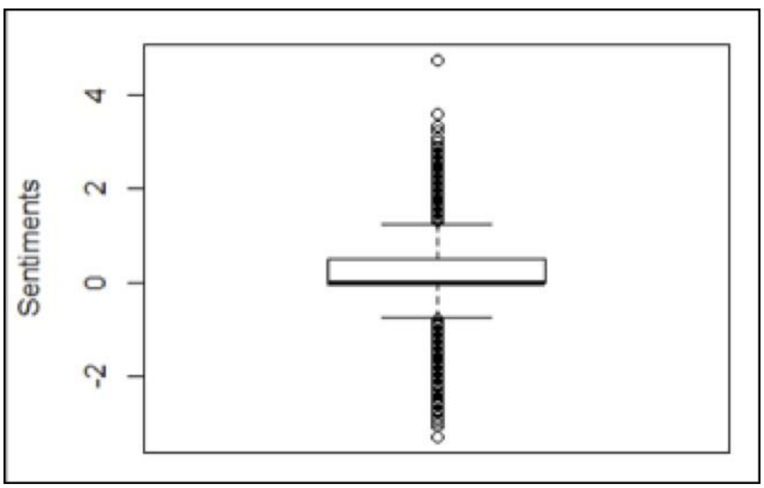


Figure 1. Bullish vs bearish.

4. Research Findings and Discussion

The descriptive statistics of the variables, liquidity measures and aggregated sentiments, for the data sample are presented in Table 1, which shows significant differences among the applied bid–ask spreads. As can be seen in Table 1, the variables are positively skewed, which indicates the right-skewed distributions of variables with values to the right of their mean, whereas higher variable kurtosis represents the possibility of extreme values.

Table 1. Descriptive statistics of variables for the data sample.

	S	*GMS*	*QS*	*ES*	*Sentiments*
Min	9.030	4.520	0.002997	0.000040	−6.950
Median	22.77	11.39	0.007764	0.004452	44.70
Mean	27.69	13.84	0.009410	0.005359	42.32
Max	92.04	46.02	0.031950	0.023520	101.6
Std. Dev.	17.21	8.600	0.005982	0.004918	21.53
Skewness	1.729	1.729	1.766	1.382	0.125
Kurtosis	5.966	5.966	6.143	5.173	3.396

The liquidity and sentiment measurements are shown in the graph in Figures 2 and 3, respectively. The time-series plots presented in Figure 2 clearly disclose differences between the computed liquidity.

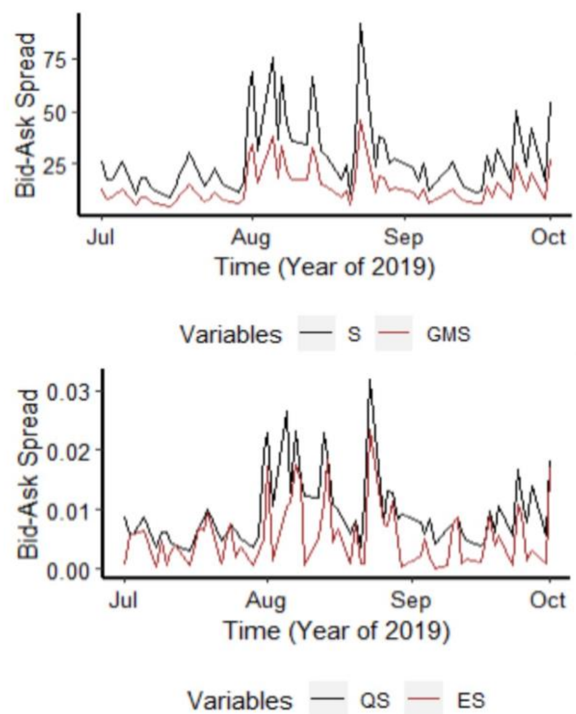


Figure 2. Time-variations in financial market liquidity.

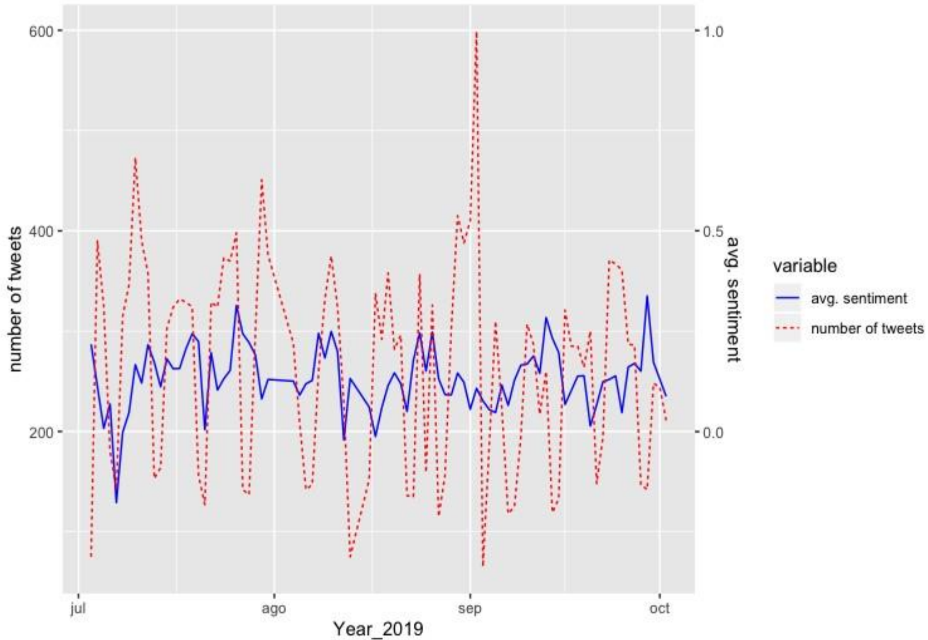


Figure 3. Time-variations in sentiments average and the number of tweets per day.

We have performed an autocorrelation analysis to study whether the time series are linearly related to lagged versions of themselves. Figure 4 shows the autocorrelation plot for each liquidity measure. We observe that values are within 95% confidence interval (represented by the dashed blue line) for lags > 1, but lag = 1 falls outside this confidence interval. So that, we can conclude that all liquidity measures are serially correlated, and this must be considered in the regression models by including the lagged computation of the liquidity measure (Equation (6)).

$$Liquidity_t = \alpha + \beta_1 Sentiments_t + \beta_2 Liquidity_{t-1} + \epsilon_t \tag{6}$$

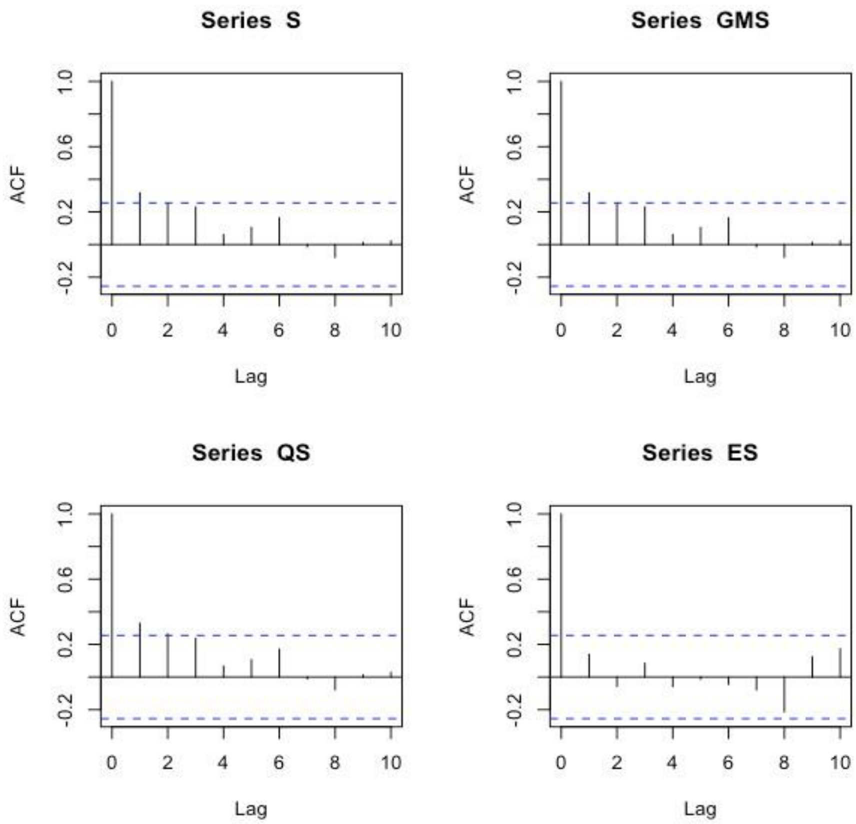


Figure 4. Autocorrelation plot for liquidity measures.

The correlation coefficients between liquidity measures, shown in Table 2, are highly correlated. This implies that the applied liquidity measures significantly respond to any variations in market liquidity over time. Figure 3 shows the evolution of the sentiment average regarding the S&P500 index along with the daily number of tweets collected. It has been observed that the sentiments are not constant and change over time, while something similar happens to the number of tweets. However, no clear relationship can be found between the average sentiment and the number of tweets. We analysed the relationship between these time-variations in sentiments and liquidity by means of a regression analysis.

Table 2. Linear correlation coefficients for the liquidity measures.

	S	***GMS***	***QS***	***ES***
S	1.000	1.000	0.999	0.790
GMS	1.000	1.000	0.999	0.790
QS	0.999	0.999	1.000	0.791
ES	0.790	0.790	0.791	1.000

Table 3 shows the coefficients obtained through a linear regression where sentiments average, “Sent”, is the independent variable and each liquidity measure is considered as the dependent variable. The lagged liquidity measure has been also included to control for autocorrelation (Equation (6)) .

The regression analyses notably reveals that the coefficients associated to sentiments, β_1, are positive in each dataset but not significant. This indicates that any changes in investors' perceptions are very weakly correlated to variations in financial market liquidity and trading costs. R-Squared, p-values, and F-statistics in Table 3 further explain that a much lower proportion of the changes in market participants' perceptions affect the time-varying liquidity and trading cost for each dataset. These results were obtained after controlling for autocorrelation. In addition, we have checked the residuals of all regression models. The Kolmogorov–Smirnov test shows that residuals are not normally distributed except for the ES liquidity measure (p-value = 0.421). Heteroscedasticity was also checked. In this case, the Breusch–Pagan test showed that residuals were homoscedastic for all regressions.

Table 3. Regression Analysis.

		Estimate	**p-Value**
S (1)	*Intercept*	18.676	0.001 **
	Sentiments	0.005	0.960
	S_{t-1}	0.329	0.015 *
GMS (2)	*Intercept*	9.338	0.001 **
	Sentiments	0.002	0.960
	GMS_{t-1}	0.329	0.015 *
QS (3)	*Intercept*	0.006	0.002 **
	Sentiments	0.000	0.976
	QS_{t-1}	0.342	0.011 *
ES (4)	*Intercept*	0.003	0.042 *
	Sentiments	0.000	0.212
	ES_{t-1}	0.115	0.419

(1) Adjusted R-squared: 0.073, F-statistic: 3.232, p-value: 0.047; (2) Adjusted R-squared: 0.073, F statistic: 3.232, p-value: 0.047; (3) Adjusted R-squared: 0.082, F-statistic: 3.539, p-value: 0.036; (4) Adjusted R-squared: 0.014, F-statistic: 1.411, p-value: 0.252; Signif. codes: '***' < 0.001; '**' < 0.01; '*' < 0.05.

These results only discount the relation between investors' sentiments and financial liquidity on a daily basis. In other words, it seems that the liquidity on a given trading day is not related to the investors' sentiments collected from tweets. The following experiment was constructed on a two-day basis: we computed the two-day moving average for each considered liquidity measure and the sentiment score to analyse whether the relation between the computed liquidity measures and sentiments was not only observed at one point in time, but was constructed over a period of time. Investors may need extra time to analyse social media's mood and then use this information, not necessarily on the same day, but also during the following trading sessions.

Table 4 gives the regression statistics for each dataset applied in the study. The regression estimates are again positive in each dataset, but their relationship reflects that changes in investors' mood are very weakly effective in determining the size of the bid–ask spread. In all regressions the improve in the R^2 comes from the inclusion of the lagged liquidity measure. However, the coefficient of sentiments with the ES liquidity measure is statistically significant even after controlling for autocorrelation, which implies that investors' mood and the liquidity of the S&P500 Index are related on a two-day basis. This gives a new insight into the relation between these two variables, which will be analysed in a further study with a larger database and different financial assets to those considered here. Residuals of all regression models have been also checked. The Kolmogorov–Smirnov and Breusch–Pagan tests show that residuals are normally distributed and homoscedastic in all cases.

Table 4. Regression analysis based on two-days moving average of liquidity.

		Estimate	***p*-Value**
S (1)	*Intercept*	6.119	0.114
	Sentiments	0.044	0.479
	S_{t-1}	0.715	0.000 ***
GMS (2)	*Intercept*	3.059	0.114
	Sentiments	0.022	0.479
	GMS_{t-1}	0.715	0.000 ***
QS (3)	*Intercept*	0.002	0.129
	Sentiments	0.000	0.484
	QS_{t-1}	0.723	0.000 ***
ES (4)	*Intercept*	0.001	0.404
	Sentiments	0.000	0.047 *
	ES_{t-1}	0.515	0.000 ***

(1) Adjusted R-squared: 0.505, F-statistic: 29.576, p-value: 0.000; (2) Adjusted R-squared: 0.505, F-statistic: 29.576, p-value: 0.000; (3) Adjusted R-squared: 0.516, F-statistic: 30.892, p-value: 0.000; (4) Adjusted R-squared: 0313, F-statistic: 13.793, p-value: 0.000; Signif. codes: '***' < 0.001; '**' < 0.01; '*' < 0.05.

5. Conclusions

In this study we analysed the impact of investors' mood on market liquidity and on the costs associated with trading. We performed a sentiment analysis of tweets related to the S&P500 Index and considered four different measures of liquidity. On a daily basis, we found that even though the regression estimates are positive, they are not statistically significant. However, if a two-day moving average is computed on all the variables concerned, the results are slightly improved. The investors' mood was found to be positive and significantly related to the effective spread of the liquidity measure.

Our findings should encourage other researchers to make additional efforts to study a larger dataset by widening the analysed period and including new assets. In the present study we only investigated the S&P500 Index, which is limited to concluding in a broader sense that Twitter, as a source of information, has little influence on any changes that occur in the size of spread and time-varying liquidity. There is thus a great need for future research in this area to study the relationship between microblogging data and market liquidity at the sector and firm levels. This would undoubtedly help us to understand the significance of microblogging data on financial market liquidity and trading costs in a broader sense.

References

1. Kaplan, A.M.; Haenlein, M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus. Horiz.* **2010**, *53*, 59–68. [CrossRef]
2. O'connor, P. Managing a hotel's image on TripAdvisor. *J. Hosp. Mark. Manag.* **2010**, *19*, 754–772. [CrossRef]
3. Ceron, A.; Curini, L.; Iacus, S.M. Using sentiment analysis to monitor electoral campaigns: Method matters—Evidence from the United States and Italy. *Soc. Sci. Comput. Rev.* **2015**, *33*, 3–20. [CrossRef]
4. Zeng, B.; Gerritsen, R. What do we know about social media in tourism? A review. *Tour. Manag. Perspect.* **2014**, *10*, 27–36. [CrossRef]
5. Grajales, F.J., III; Sheps, S.; Ho, K.; Novak-Lauscher, H.; Eysenbach, G. Social media: A review and tutorial of applications in medicine and health care. *J. Med. Internet Res.* **2014**, *16*, e13. [CrossRef]
6. Adams, S.A.; Van Veghel, D.; Dekker, L. Developing a research agenda on ethical issues related to using social media in healthcare: Lessons from the first Dutch Twitter heart operation. *Camb. Q. Healthc. Ethics* **2015**, *24*, 293–302. [CrossRef]
7. Zhang, X.; Wang, W.; de Pablos, P.O.; Tang, J.; Yan, X. Mapping development of social media research through different disciplines: Collaborative learning in management and computer science. *Comput. Hum. Behav.* **2015**, *51*, 1142–1153. [CrossRef]
8. Boulianne, S. Social media use and participation: A meta-analysis of current research. *Inf. Commun. Soc.* **2015**, *18*, 524–538. [CrossRef]
9. Filo, K.; Lock, D.; Karg, A. Sport and social media research: A review. *Sport Manag. Rev.* **2015**, *18*, 166–181. [CrossRef]
10. McFarland, L.A.; Ployhart, R.E. Social media: A contextual framework to guide research and practice. *J. Appl. Psychol.* **2015**, *100*, 1653. [CrossRef]
11. Leonardi, P.M.; Vaast, E. Social media and their affordances for organizing: A review and agenda for research. *Acad. Manag. Ann.* **2017**, *11*, 150–188. [CrossRef]
12. Roth, P.L.; Bobko, P.; Van Iddekinge, C.H.; Thatcher, J.B. Social media in employee-selection-related decisions: A research agenda for uncharted territory. *J. Manag.* **2016**, *42*, 269–298. [CrossRef]
13. Pope, D.; Griffith, J. An Analysis of Online Twitter Sentiment Surrounding the European Refugee Crisis. In Proceedings of the International Joint Conference on Knowledge

Discovery, Knowledge Engineering and Knowledge Management, Porto, Portugal, 12–14 November 2016; pp. 299–306.
14. Shaikh, S.; Feldman, L.B.; Barach, E.; Marzouki, Y. Tweet sentiment analysis with pronoun choice reveals online community dynamics in response to crisis events. In *Advances in Cross-cultural Decision Making*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 345–356.
15. Öztürk, N.; Ayvaz, S. Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis. *Telemat. Inform.* **2018**, *35*, 136–147. [CrossRef]
16. Agostino, D.; Arena, M.; Catalano, G.; Erbacci, A. Public engagement through social media: The spending review experience. *Public Money Manag.* **2017**, *37*, 55–62. [CrossRef]
17. Bollen, J.; Mao, H.; Zeng, X. Twitter mood predicts the stock market. *J. Comput. Sci.* **2011**, *2*, 1–8. [CrossRef]
18. Oliveira, N.; Cortez, P.; Areal, N. The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Syst. Appl.* **2017**, *73*, 125–144. [CrossRef]
19. Bank, S.; Yazar, E.E.; Sivri, U. Can social media marketing lead to abnormal portfolio returns? *Eur. Res. Manag. Bus. Econ.* **2019**, *25*, 54–62. [CrossRef]
20. Saleemi, J. An Empirical Analysis of Cost-Based Market Liquidity Measures for US & Norwegian Banks. Master's Thesis, Universitetet i Nordland, Bodø, Norway, 2014.
21. PricewaterhouseCoopers. *Global Financial Markets Liquidity Study*; Technical Report, PwC; 2015. Available online: https://www.pwc.se/sv/pdf-reports/global-financial-markets-liquidity-study.pdf (accessed on 1 November 2019).
22. Kunitsyna, N.; Britchenko, I.; Kunitsyn, I. Reputational risks, value of losses and financial sustainability of commercial banks. *Entrep. Sustain. Issues* **2018**, *5*, 943–955. [CrossRef]
23. Akerlof, G.A. The market for lemons: Quality uncertainty and the market mechanism. *Q. J. Econ.* **1970**, *84*, 488–500. [CrossRef]
24. Bagehot, W. The only game in town. *Financ. Anal. J.* **1971**, *27*, 12–14. [CrossRef]
25. Glosten, L.R.; Milgrom, P.R. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *J. Financ. Econ.* **1985**, *14*, 71–100. [CrossRef]
26. Easley, D.; O'hara, M. Price, trade size, and information in securities markets. *J. Financ. Econ.* **1987**, *19*, 69–90. [CrossRef]
27. Oviedo-García, M.; Muñoz-Expósito, M.; Castellanos-Verdugo, M.; Sancho-Mejías, M. Metric proposal for customer engagement in Facebook. *J. Res. Interact. Mark.* **2014**, *8*, 327–344. [CrossRef]
28. Li, Z.; Li, C. Tweet or "re-tweet"? An experiment of message strategy and interactivity on Twitter. *Internet Res.* **2014**, *24*, 648–667. [CrossRef]
29. Kelly, L.; Kerr, G.; Drennan, J. Avoidance of advertising in social networking sites: The teenage perspective.

J. Interact. Advert. **2010**, *10*, 16–27. [CrossRef]
30. Sawhney, M.; Verona, G.; Prandelli, E. Collaborating to create: The Internet as a platform for customer engagement in product innovation. *J. Interact. Mark.* **2005**, *19*, 4–17. [CrossRef]
31. Michaelidou, N.; Siamagka, N.T.; Christodoulides, G. Usage, barriers and measurement of social media marketing: An exploratory investigation of small and medium B2B brands. *Ind. Mark. Manag.* **2011**, *40*, 1153–1159. [CrossRef]
32. Sprenger, T.O.; Tumasjan, A.; Sandner, P.G.; Welpe, I.M. Tweets and trades: The information content of stock microblogs. *Eur. Financ. Manag.* **2014**, *20*, 926–957. [CrossRef]
33. Larsson, A.O.; Moe, H. Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media Soc.* **2012**, *14*, 729–747. [CrossRef]
34. Hong, S.; Kim, S.H. Political polarization on twitter: Implications for the use of social media in digital governments. *Gov. Inf. Q.* **2016**, *33*, 777–782. [CrossRef]
35. Holmberg, K.; Thelwall, M. Disciplinary differences in Twitter scholarly communication. *Scientometrics* **2014**, *101*, 1027–1042. [CrossRef]
36. Vidya, N.A.; Fanany, M.I.; Budi, I. Twitter sentiment to analyze net brand reputation of mobile phone providers. *Procedia Comput. Sci.* **2015**, *72*, 519–526. [CrossRef]
37. Tetlock, P.C. Giving content to investor sentiment: The role of media in the stock market. *J. Financ.* **2007**, *62*, 1139–1168. [CrossRef]
38. Fang, L.; Peress, J. Media coverage and the cross-section of stock returns. *J. Financ.* **2009**, *64*, 2023–2052. [CrossRef]
39. Luo, X.; Zhang, J.; Duan, W. Social media and firm equity value. *Inf. Syst. Res.* **2013**, *24*, 146–163. [CrossRef]
40. Chen, H.; De, P.; Hu, Y.J.; Hwang, B.H. Wisdom of crowds: The value of stock opinions transmitted through social media. *Rev. Financ. Stud.* **2014**, *27*, 1367–1403. [CrossRef]
41. Li, Q.; Chen, Y.; Wang, J.; Chen, Y.; Chen, H. Web media and stock markets: A survey and future directions from a big data perspective. *IEEE Trans. Knowl. Data Eng.* **2017**, *30*, 381–399. [CrossRef]
42. Nofsinger, J.R. Social mood and financial economics. *J. Behav. Financ.* **2005**, *6*, 144–160. [CrossRef]
43. Zhao, S.; Tong, Y.; Liu, X.; Tan, S. Correlating Twitter with the stock market through non-Gaussian SVAR. In Proceedings of the 2016 Eighth International Conference on Advanced Computational Intelligence (ICACI), Chiang Mai, Thailand, 14–16 February 2016; pp. 257–264.
44. Sul, H.; Dennis, A.R.; Yuan, L.I. Trading on twitter: The financial information content of emotion in social media. In Proceedings of the 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 January 2014; pp. 806–815.

45. Ruan, Y.; Alfantoukh, L.; Durresi, A. Exploring stock market using twitter trust network. In Proceedings of the 2015 IEEE 29th International Conference on Advanced Information Networking and Applications, Gwangju, Korea, 25–27 March 2015; pp. 428–433.
46. Rakowski, D.A.; Shirley, S.; Stark, J. TwitTer Activity, Investor Attention, and the Diffusion Of Information. 2018. Available online: https://ssrn.com/abstract=3010915 (accessed on 1 November 2019).
47. Agrawal, S.; Azar, P.D.; Lo, A.W.; Singh, T. Momentum, Mean-Reversion, and Social Media: Evidence from StockTwits and Twitter. *J. Portf. Manag.* **2018**, *44*, 85–95. [CrossRef]
48. Liew, J.K.S.; Wang, G.Z. Twitter sentiment and IPO performance: A cross-sectional examination. *J. Portf. Manag.* **2016**, *42*, 129–135. [CrossRef]
49. Blankespoor, E.; Miller, G.S.; White, H.D. The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *Account. Rev.* **2013**, *89*, 79–112. [CrossRef]
50. Renault, T. Market Manipulation and Suspicious Stock Recommendations on Social Media. 2017. Available online: https://ssrn.com/abstract=3010850 (accessed on 1 November 2019).
51. Al Guindy, M. Is Corporate Tweeting Informative or Is It Just Hype? Evidence from the SEC Social Media Regulation. 2016. Available online: https://ssrn.com/abstract=2824668 (accessed on 1 November 2019).
52. Ranco, G.; Aleksovski, D.; Caldarelli, G.; Grcˇar, M.; Mozeticˇ, I. The effects of Twitter sentiment on stock price returns. *PLoS ONE* **2015**, *10*, e0138441. [CrossRef] [PubMed]
53. Smailovic´, J.; Grcˇar, M.; Lavracˇ, N.; Žnidaršicˇ, M. Predictive sentiment analysis of tweets: A stock market application. In *International Workshop on Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 77–88.
54. Zheludev, I.; Smith, R.; Aste, T. When can social media lead financial markets? *Sci. Rep.* **2014**, *4*, 4213. [CrossRef] [PubMed]
55. Hasbrouck, J.; Seppi, D.J. Common factors in prices, order flows, and liquidity. *J. Financ. Econ.* **2001**, *59*, 383–411. [CrossRef]
56. Amihud, Y.; Mendelson, H. Asset pricing and the bid–ask spread. *J. Financ. Econ.* **1986**, *17*, 223–249. [CrossRef]
57. Huang, R.D.; Stoll, H.R. The components of the bid–ask spread: A general approach. *Rev. Financ. Stud.* **1997**,
10, 995–1034. [CrossRef]
58. Corwin, S.A.; Schultz, P. A simple way to estimate bid–ask spreads from daily high and low prices. *J. Financ.* **2012**, *67*, 719–760. [CrossRef]
59. Mancini, L.; Ranaldo, A.; Wrampelmeyer, J. Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. *J. Financ.* **2013**, *68*, 1805–1841. [CrossRef]
60. Demsetz, H. The cost of transacting. *Q. J. Econ.* **1968**, *82*, 33–53. [CrossRef]

61. Ho, T.; Stoll, H.R. Optimal dealer pricing under transactions and return uncertainty. *J. Financ. Econ.* **1981**, *9*, 47–73. [CrossRef]
62. Ho, T.S.; Stoll, H.R. The dynamics of dealer markets under competition. *J. Financ.* **1983**, *38*, 1053–1074. [CrossRef]
63. Brock, W.A.; Kleidon, A.W. Periodic market closure and trading volume: A model of intraday bids and asks.
J. Econ. Dyn. Control **1992**, *16*, 451–489. [CrossRef]
64. Gorton, G.; Metrick, A. Haircuts. *Fed. Reserve Bank St. Louis Rev.* **2009**, *92*, 507–519.
65. Madrigal, V. Non-fundamental speculation. *J. Financ.* **1996**, *51*, 553–578. [CrossRef]
66. Brunnermeier, M.K.; Pedersen, L.H. Predatory trading. *J. Financ.* **2005**, *60*, 1825–1863. [CrossRef]
67. Copeland, T.E.; Galai, D. Information effects on the bid–ask spread. *J. Financ.* **1983**, *38*, 1457–1469. [CrossRef]
68. Roll, R. A simple implicit measure of the effective bid–ask spread in an efficient market. *J. Financ.* **1984**, *39*, 1127–1139. [CrossRef]
69. Bedowska-Sójka, B.; Echaust, K. Commonality in Liquidity Indices: The Emerging European Stock Markets. *Systems* **2019**, *7*, 24. [CrossRef]

Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data

Abstract: Market liquidity has an immediate impact on the execution of transactions in financial markets. Informed counterparty risk is often priced into market liquidity. This study investigates whether microblogging data, as a non-financial information tool, is priced along with market liquidity dimensions. The analysis is based on the Australian Securities Exchange (ASX), and from the results, we conclude that microblogging content in pessimistic periods has a higher impact on liquidity and its dimensions. On a daily basis, pessimistic investor sentiments lead to higher trading costs, illiquidity, a larger price dispersion and a lower trading volume.

Keywords: microblogging data; data mining; investor sentiments; asset pricing; market liquidity; liquidity dimensions

1. Introduction

This work investigates whether microblogging data, as a source of information, can explain liquidity dimensions.

In the behavioral finance literature, emotion-driven market participants with stochastic predictions are gaining a considerable amount of interest. Recent research often quantified surveys, message boards (e.g., ragingbull.com, accessed on 19 August 2021) or financial news to construct sentiment indicators for modeling stock market behavior. Researchers are also exploring microblogging data for use in both modeling and predicting stock market behavior (Zhang et al. 2011). Moreover, microblogging sentiment indicators may be more economically meaningful than traditional sources of financial information (Oliveira et al. 2017).

The participation of companies in microblogging platforms can contribute to the development of valuable knowledge among investors (Prokofieva 2015), and increase the opportunity for significant returns (Bank et al. 2019). Market liquidity is often reported to be priced into asset returns (Saleemi 2020). An abundance of studies can be found that examine microblogging data for financial market prediction. However, there is still room to explore the impact of microblogging content on various liquidity dimensions.

The novelty of our work lies in the methodological contribution compared to related works. This study links the different dimensions of market liquidity with sentiment analysis using content from the popular social media platform Twitter.

In contrast to previous studies, investor sentiment tools are applied to uncover their role in the liquidity dimensions of microblogging content. This research fills a gap in the behavioral finance literature, and helps us to understand the impact on informed counterparty liquidity in a broader sense.

Liquidity, or its risk, is an active area of research as it imposes immediate consequences on the financial transaction (Guijarro et al. 2019). Market liquidity can be explained by its dimensions, which include transaction execution cost, trading quantity, immediacy of transaction execution and asset price dispersion (Le and Gregoriou 2020). Trading is considered illiquid (Gorton and Metrick 2010), and it is assumed that asymmetric information risk should be priced into liquidity

(Saleemi 2020). Microblogging platforms allow market participants to exchange financial information on a real-time basis. To our knowledge, this is the first paper to study whether microblogging content, as an indicator of investor sentiments, is priced in the various dimensions of liquidity.

As microblogging content is gaining considerable attention in the behavioral finance literature, the aim of this research is therefore to explore whether liquidity dimensions can be significantly explained by microblogging sentiment indicators. As there is no previous literature on how investor sentiments may affect the different dimensions of liquidity, we do not hypothesize what the sign of the relationship between the two variables should be. Our paper aims to be the first empirical approach to the study of this problem. The results may have potential implications for both researchers and traders in terms of quantifying microblogging content-based sentiments with regards to market liquidity dimensions.

The rest of the paper is structured as follows. The literature is reviewed in Section 2. The procedure used to build the model and the data set is explained in Section 3. Section 4 discusses the findings of the research. Finally, Section 5 highlights the main results of the research.

2. Literature Review

The proliferation of behavior finance literature is attributable to the authoritative role of various sources of information on investor sentiments. Among the diversified structure of social networks, it may be of great interest to identify the most valued opinion providers. Microblogging platforms, in particular Twitter, allow participants to exchange potential content about financial markets on a real-time basis (Oliveira et al. 2017). Investor sentiment can be linked to systematic risk (Lee et al. 2002).

Investor sentiment determines asset price levels and therefore needs to be taken into account in the asset pricing model (Brown and Cliff 2005). Aggregate opinion has a significant impact on financial assets, the valuations of which are extremely subjective and difficult to arbitrage (Baker and Wurgler 2006). Moreover, financial assets without media coverage earn higher returns (Fang and Peress 2009), while monetary policy decisions in bear market periods have a greater impact on financial assets (Kurov 2010).

Incoming news significantly influences stock returns, volatility and trading volumes (Groß-Klußmann and Hautsch 2011). Microblogging content has some predictive power on returns, market-adjusted returns and future directional stock price movements (Oh and Sheng 2011). Twitter is a potential indicator of how the financial market will behave the next day (Zhang et al. 2011), while investor sentiments extracted from Twitter comments can predict asset price movements a few days in advance (Smailovic´ et al. 2013).

Microblogging content can have greater effects on stock market performance than conventional media (Yu et al. 2013). Media investment interest plays a crucial role in reducing the information asymmetry, which in turn can stabilize the market, protect investors and improve corporate governance (Wei et al. 2014). In that sense, microblogging data can be a reliable source of stock-related news (Sprenger et al. 2014).

According to Walker (2016), the media can drive market behavior. In addition, companies' activity on Twitter can reduce the expected negative reactions in the market (Mazboudi and Khalil 2017). According to Li et al. (2018), users' attention to Twitter can better reflect stock trends. Aggregate opinion on Twitter is relevant for predicting a company's forthcoming quarterly earnings (Bartov et al. 2018), although Twitter content is less effective in determining market liquidity and trading cost (Guijarro et al. 2019).

Market liquidity and its related issues comprise one of the dominant strands of the asset pricing literature. With respect to the concept of information effects, the informed trader drives market liquidity (Glosten and Milgrom 1985). Immediacy, tightness, depth, breadth and resilience are the five key characteristics of a liquid market, according to Sarr and Lybek (2002). Market liquidity can be determined by trading cost, trading quantity, trading speed and price dispersion (Le and Gregoriou 2020). It follows then that informed trading risk must be priced in the liquidity (Saleemi 2020).

Liquidity is considered as a time-varying risk factor (Hasbrouck and Seppi 2001), as well as a crucial attribute of capital assets (Amihud and Mendelson 1991). The financial asset whose return is more sensitive to liquidity shocks has a higher expected return (Le and Gregoriou 2020). More recently, it has been found that returns are very sensitive to liquidity shocks in environments of high uncertainty, such as the current COVID-19 crisis (Saleemi 2021).

Market frictions are the costs associated with the execution of a transaction, which directly affect liquidity. Their impact has been shown to be time-varying (DeGennaro and Robotti 2007). Transaction costs can be divided into two major elements: the explicit cost and the implicit cost. The explicit cost is identifiable before the transaction takes place. However, the implicit cost is less observable and represents a large fraction of the total cost of the transaction. The bid–ask spread is a key point for the quantification of transaction costs, as it captures almost all the costs associated with the execution of the transaction (Sarr and Lybek 2002).

Since the late 1960s, the bid–ask spread has been extensively investigated in the asset pricing literature (Gregoriou 2013). Market-makers enable continuous trading by matching buy and sell orders. Liquidity providers facilitate the immediacy of trade execution by accepting the risk of holding inventory. Investors tend to reduce their risk exposure to future price uncertainty. In this context, liquidity providers impose a cost on the seller, i.e., a higher spread. The higher the volatility of asset prices, the higher the spread will be set by liquidity providers (Ho and Stoll 1981).

Another stream in the field links asymmetric information to the size of the spread. Information-sensitive stocks are illiquid. In the case of informed trading, there is a potential risk of loss for the uninformed party. Therefore, liquidity providers tend to increase the spread as compensation for this potential loss (Easley and O'Hara 2004). Another component of the spread is the order processing cost (Roll 1984). In case the order processing cost is higher, liquidity suppliers will buy an asset at the lowest bid price with the expectation of reselling it at the highest ask price.

Another interesting result is that the bid–ask spread is closely related to trading volume. The higher the cost of trading, the lower the amount of trading (Easley and O'Hara 1992). A small spread translates into a larger amount of trading, as the number of active trading participants causes the spread to become narrower. There are also causal effects on the spread of the amount of trading. A small trading volume reduces the size of the spread, which in turn adds liquidity to the market and improves price accuracy (Sarkissian 2016). According to Le and Gregoriou (2020), there is a strong relationship between higher trading volume and higher spread due to asymmetric information effects.

3. Materials and Methods

Our paper investigates whether informed trading based on microblogging content influences liquidity dimensions. To do so, we extracted investor sentiments from the popular social network Twitter, collecting different measures of each liquidity dimension in order to investigate the relationship between microblogging content and liquidity dimensions. Studies of asset pricing introduce several measures that capture one or more dimensions of market liquidity. This paper focuses on a small number of proxies for each dimension of liquidity, namely the bid–ask spread and liquidity based on the volume of price impact.

Depending on the frequency of the data, liquidity indicators are modeled in two ways: high-frequency data and low-frequency data. High-frequency measures estimate liquidity and its dimensions from intraday financial transactions. In contrast, the construction of low-frequency proxies is based on the daily characteristics of a security, such as the opening, high, low and closing prices (OHLC prices), as well as the volume traded. Unlike high-frequency data, low-frequency data are computationally less intensive and widely accessible to the markets. In this research, the analysis is based on low-frequency data from the Australian Securities Exchange (ASX), and was run over the period 3 January 2020 to 2 June 2021.

Among the measures of liquidity, the literature devotes much attention to the bid–ask spread. The spread captures the immediacy and cost of transactions. A large spread reflects a liquidity provider's unwillingness to accept an inventory position without imposing a higher cost on the seller. Most recently, Saleemi (2020) proposed a model of the cost-based market liquidity (CBML) measure, i.e., the bid–ask spread. The CBML model estimates the possible presence of an informed trader in the financial market. Based on the general foundations of the asset pricing literature, CBML is developed from Equation (1):

$$CBML_t = \sqrt{[(S_{t-1}) - (v_t^s)]^2} \quad (1)$$

where S_{t-1} is the ratio between the price range and the closing price on day $t-1$. This value is estimated by Equation (2):

$$S_{t-1} = \frac{high_{t-1} - low_{t-1}}{Close_{t-1}} \quad (2)$$

where, $high_{t-1}$ indicates the highest price on day $t-1$; low_{t-1} refers to the lowest price of day $t-1$; and $Close_{t-1}$ is the closing price on day $t-1$. In the next trading session, the CBML method estimates the effects of asymmetric information on asset prices. v_t^s computes the ratio between the range price of an informed trader and the closing price on day t, as per Equation (3):

$$v_t^s = \frac{v_t^{ask} - v_t^{bid}}{Close_t} \tag{1}$$

Assuming risk neutrality in the next trading session, the asset is valued at:

$$\eta_t = (high_t + low_t)/2 \tag{2}$$

where η_t is the mean of high and low prices on day t. If we consider the same probability of an informed trader, the estimated ask value for which the seller would redeem his position is assumed to be conditional on a trade such as:

$$v_t^{ask} = ask_t\pi + \eta_t\pi \tag{3}$$

where the estimated bid value for which the buyer would accept the inventory position is assumed conditional on a trade such as:

$$v_t^{bid} = bid_t\pi + \eta_t\pi \tag{4}$$

The liquidity model based on the impact of price on volume mainly estimates the level of liquidity by the dispersion of the asset price and its trading quantity. The Martin Liquidity Index (MLI) estimates the link between price changes and trading volume. The MLI model assumes that price dispersion influences trading volume and, as a result, impacts market liquidity. The higher the MLI value, the greater the price dispersion relative to the quantity traded. Hence, higher price dispersion leads to lower market liquidity. The analytical expression of the MLI for period t is given in Equation (7):

$$MLI_t = \sum_{t=1}^{T} \frac{(Close_t - Close_{t-1})^2}{ln\,(Vol_t)} \tag{7}$$

where Vol_t is the quantity traded of the asset on day t. The model explains the price impact in terms of the effect that a traded unit has on the price. The illiquid asset requires less trading to move prices compared to the liquid asset. Note that our research only aims to estimate the influence of investor sentiments on stock market liquidity, therefore it is not necessary to deflate the price series. Such a deflation would make sense in a hypothetical case where one would want to analyze the profitability of an investment, as in the case where an analysis of investor sentiments could be used to derive a stock market investment strategy.

The R programming language was used to collect tweets from the ASX during the period from 3 January 2020 to 2 June 2021, using the libraries "ROAuth", "twitteR" and "rtweet". The study emphasizes pre-processing the unstructured text of the tweets. This process was carried out using the "NLP" and "tm" libraries, which allowed the original data to be cleaned and structured appropriately for further processing. Sentiment analysis tools were applied to convert intraday tweets into structured and valuable content. Tweets were structured by removing punctuation symbols, stop words and trailing spaces. In addition, the text was converted into lower case for the analysis of the microblogging financial conversation. For ethical reasons, market participants have been anonymized. For each tweet, the financial information was quantified in either a bullish (positive) or bearish (negative) score. Neutral opinions were not taken into account in the analysis. As the number of tweets posted on a single day is very large, the sentiment values for day t were aggregated for the analysis. This process was carried out through the “syuzhet” and “lubridate” libraries.

The basic sentiment indicators, i.e., the negative ratio (NR) and positive ratio (PR), were used as attributes according to Equations (8) and (9):

$$NR_t = \frac{Bear_t}{Bull_t} \tag{8}$$

$$PR_t = \frac{Bull_t}{Bear_t} \tag{9}$$

where $Bear_t$ is the aggregated bearish value on day t; and $Bull_t$ indicates the accumulated bullish value on day t.

First, we considered investor sentiment indicators as explanatory variables and liquidity dimensions as response variables, with both variables expressed in daily values. Next, the multiple linear regression model in Equation (10) was used to estimate the impact of investor sentiments on liquidity dimensions:

$$LD_t = \alpha + \beta_1 NR_t + \beta_2 PR_t + \epsilon_t \tag{10}$$

where LD_t refers to each measure of the liquidity dimension on day t; NR_t reflects the aggregated pessimistic sentiments on day t; PR_t indicates the aggregate optimistic sentiments on day t; and ϵ_t is the error term.

The following experiment is based on a multivariate forecasting algorithm, the vector autoregression (VAR) model. In this case, variables are modeled as a linear combination of their own lags and the past values of other variables. The Schwarz criterion (SC), also known as Bayesian information criterion, is applied to select the optimal lags. To estimate the impact of lags, the VAR model is structured through Equations (11)–(13):

$$L_t = \alpha_L + \beta_{11} L_{t-1} + \beta_{12} L_{t-2} + \gamma_{11} N_{t-1} + \gamma_{12} N_{t-2} + \emptyset_{11} P_{t-1} + \emptyset_{12} P_{t-2} + \epsilon_{L,t} \tag{11}$$

$$N_t = \alpha_N + \beta_{21} L_{t-1} + \beta_{22} L_{t-2} + \gamma_{21} N_{t-1} + \gamma_{22} N_{t-2} + \emptyset_{21} P_{t-1} + \emptyset_{22} P_{t-2} + \epsilon_{N,t} \tag{12}$$

$$P_t = \alpha_P + \beta_{31} L_{t-1} + \beta_{32} L_{t-2} + \gamma_{31} N_{t-1} + \gamma_{32} N_{t-2} + \emptyset_{31} P_{t-1} + \emptyset_{32} P_{t-2} + \epsilon_{P,t} \tag{13}$$

where L_t denotes each liquidity dimension on day t; L_{t-1} (L_{t-2}) refers to the lag value of each liquidity dimension on day $t-1$ ($t-2$); N_{t-1} (N_{t-2}) reflects the pessimistic sentiments on day $t-1$ ($t-2$); P_{t-1} (P_{t-2}) refers to the optimistic sentiments on day $t-1$ ($t-2$); $\epsilon_{L,t}$ is the white-noise variable; N_t refers to the negative sentiments on day t; $\epsilon_{N,t}$ is the white-noise variable; P_t refers to the positive sentiments on day t; and $\epsilon_{P,t}$ is another white-noise variable.

In the following, we represent this model in a matrix notation:

$$\begin{bmatrix} L_t \\ N_t \\ P_t \end{bmatrix} = \begin{bmatrix} \alpha_L \\ \alpha_N \\ \alpha_P \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{bmatrix} \begin{bmatrix} L_{t-1} \\ L_{t-2} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_{31} & \gamma_{32} \end{bmatrix} \begin{bmatrix} N_{t-1} \\ N_{t-2} \end{bmatrix} + \begin{bmatrix} \emptyset_{11} & \emptyset_{12} \\ \emptyset_{21} & \emptyset_{22} \\ \emptyset_{31} & \emptyset_{32} \end{bmatrix} \begin{bmatrix} P_{t-1} \\ P_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{L,t} \\ \epsilon_{N,t} \\ \epsilon_{P,t} \end{bmatrix} \quad (14)$$

Equation (14) is further elaborated as:

$$LS_t = \begin{bmatrix} L_t \\ N_t \\ P_t \end{bmatrix}, A = \begin{bmatrix} \alpha_L \\ \alpha_N \\ \alpha_P \end{bmatrix}, \beta = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{bmatrix}, L_t = \begin{bmatrix} L_{t-1} \\ L_{t-2} \end{bmatrix}, \gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_{31} & \gamma_{32} \end{bmatrix}, N_t = \begin{bmatrix} N_{t-1} \\ N_{t-2} \end{bmatrix},$$

$$\emptyset = \begin{bmatrix} \emptyset_{11} & \emptyset_{12} \\ \emptyset_{21} & \emptyset_{22} \\ \emptyset_{31} & \emptyset_{32} \end{bmatrix}, P_t = \begin{bmatrix} P_{t-1} \\ P_{t-2} \end{bmatrix}, \epsilon_t = \begin{bmatrix} \epsilon_{L,t} \\ \epsilon_{N,t} \\ \epsilon_{P,t} \end{bmatrix} \quad (15)$$

Finally, we can rewrite the VAR model as Equation (16):

$$LS_t = A + \beta L_t + \gamma N_t + \emptyset P_t + \epsilon_t \quad (16)$$

4. Results and Discussion

The descriptive statistics of the data sample are shown in Table 1. It is noted that the variables are positively skewed with fat-tailed numerical distribution. Positive skewness of the data sample indicates a right-skewed distribution, with values to the right of mean. The fat-tailed numerical distribution, or higher kurtosis, indicates extreme values in the corresponding data set. The measures applied are based on distinct theoretical assumptions, which may influence the measurement of liquidity. The measures of the liquidity dimensions are plotted in Figure 1, where it is found that they are not constant, but vary over time.

Table 1. Descriptive statistics (daily basis).

Variables	Min	Median	Mean	Max	SD	S	K
CBML	0.0000522	0.008503	0.0110089	0.0869665	0.01043	3.3169	19.4480
MLI	0.000007	0.026949	0.120687	4.408370	0.35033	7.7186	80.1896
NR	0.1961	0.4984	0.5219	1.6038	0.16152	2.0675	11.3873
PR	0.6235	2.0065	2.0712	5.1	0.57910	1.1954	7.2555

Note: Cost-based market liquidity: CBML; Martin Liquidity Index: MLI; negative ratio: NR; positive ratio: PR; standard deviation: SD; skewness: S; kurtosis: K.

The microblogging sentiment indicators are depicted in Figure 2. It is also noted that investor sentiment indicators are not constant and change over time. It is worthwhile examining whether the Twitter feeds can influence the market liquidity dimensions. In this context, the sentiment analysis tools were applied to extract valuable content from unstructured Twitter feeds and the multivariate methods were applied to disentangle the various aspects involved. In our work, we aimed to analyze the impact of microblogging content on liquidity dimensions using multiple linear regression analysis.

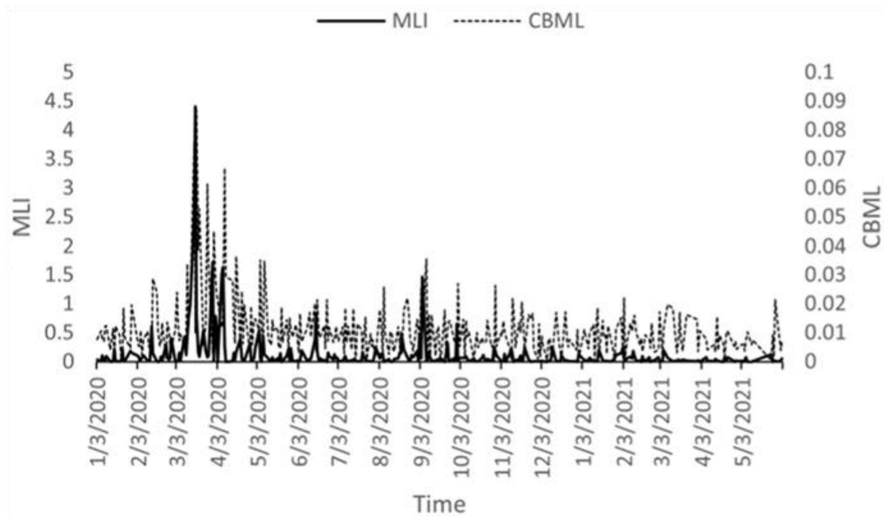


Figure 1. Time-varying market liquidity dimensions.

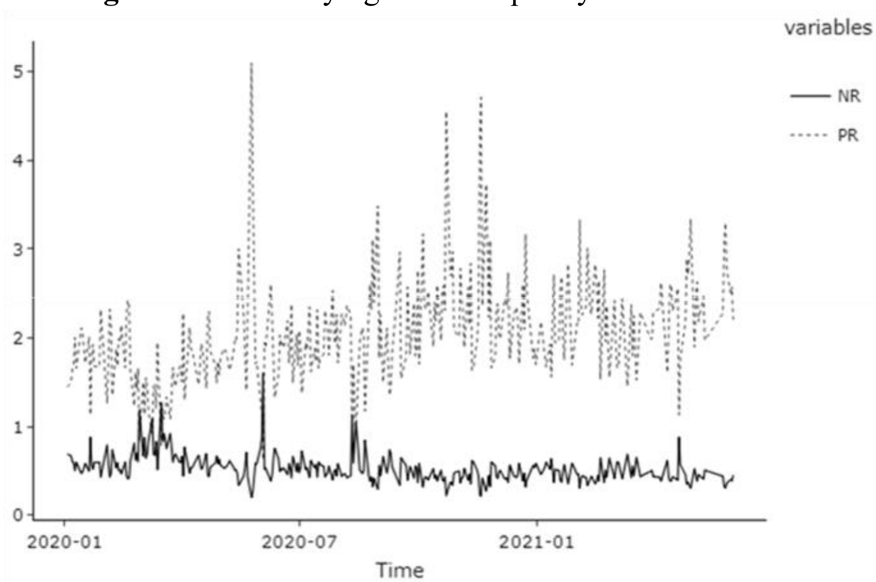


Figure 2. Time-varying investor sentiment indicators.

Table 2 presents the estimated regression values, where the investor sentiment indicators are the independent variables and each measure of the corresponding liquidity dimension acts as the dependent variable. On a daily basis, it is observed that the NR sentiment indicator is positive and significantly associated with the CBML measure. This implies that an increase in pessimistic investor sentiments leads to a higher spread. The higher spread illustrates the liquidity provider's unwillingness to accept the financial position without imposing a higher cost on the seller. A higher cost in pessimistic periods affects the speed of the transactions and therefore reduces liquidity for the ASX. Since the size of the spread is crucial in determining liquidity and its associated cost, a larger spread indicates illiquidity and a higher cost of trading in the Australian market during pessimistic periods. In contrast, the size of spread is not significantly explained by the optimistic sentiment measure, positive ratio. Therefore, optimistic sentiments based on microblogging content do not play a significant role in reducing the spread size in the Australian market.

Table 2. Regression analysis results.

Variables		Estimate	p-Value
	Intercept	− 0.006133	0.4174
CBML (a)	NR	0.021341	0.0028 **
	PR	0.002899	0.1434
MLI (b)	Intercept	− 0.46598	0.062395
	NR	0.81774	0.000533 ***
	PR	0.07718	0.237351

Note: (a) adjusted R-squared: 0.03674; F-statistic: 7.235; p-value: 0.0008427; (b) adjusted R-squared: 0.06915; F-statistic: 13.15; p-value: 0.000; significance codes: '***' < 0.001; '**' < 0.01.

The following experiment was conducted to analyze whether financial microblogging content can explain the dispersion of asset price and trading quantity. We found that pessimistic sentiments are positively and significantly associated with price impact volume-based liquidity. This indicates that a pessimistic bias in investor sentiments leads to a higher MLI value. The higher MLI value illustrates the greater price dispersion of the ASX relative to its trading volume. Therefore, investors would need a smaller amount of trades in the ASX to move its prices in the pessimistic periods. A higher MLI value, or higher price dispersion, illustrates the lack of liquidity in the ASX market. However, the optimistic mood of investors is not significantly associated with price impact volume-based liquidity.

Table 3 presents the VAR coefficients for the past time series of spreads and investor sentiments. Market liquidity and its associated cost are not significantly explained by the lagged coefficients of investor sentiments. Cost-based market liquidity is reported to be significantly correlated with its own past time series. Meanwhile, pessimistic investor sentiments are not significantly explained by the lagged coefficients of cost-based market liquidity and optimistic sentiments. Investors' optimistic sentiments are not significantly correlated with the past time series of cost-based market liquidity and pessimistic sentiments.

Table 3. Estimation of VAR coefficients and significance test values, CBML model.

Variables		Estimate	p-Value
CBML (a)	$\beta_{11,CBML}$	− 0.8168	0.000 ***
	$\gamma_{11,N}$	− 0.01081	0.0843
	$\emptyset_{11,P}$	− 0.002544	0.1453
	$\beta_{12,CBML}$	− 0.3748	0.000 ***
	$\gamma_{12,N}$	0.006916	0.2730
	$\emptyset_{12,P}$	0.0009216	0.5985
	α_{CBML}	− 0.00001418	0.9797
N (b)	$\beta_{21,CBML}$	− 0.095473	0.90380
	$\gamma_{21,N}$	− 0.677182	0.000 ***
	$\emptyset_{21,P}$	− 0.030498	0.25019
	$\beta_{22,CBML}$	− 0.530606	0.49794
	$\gamma_{22,N}$	− 0.279544	0.00375 **
	$\emptyset_{22,P}$	− 0.019775	0.45731
	α_N	− 0.001133	0.89382
P (c)	$\beta_{31,CBML}$	− 1.855197	0.5173
	$\gamma_{31,N}$	0.149640	0.6639
	$\emptyset_{31,P}$	− 0.463831	0.000 ***
	$\beta_{32,CBML}$	0.810466	0.7752
	$\gamma_{32,N}$	− 0.054357	0.8757
	$\emptyset_{32,P}$	− 0.195826	0.0429 *
	α_P	0.003898	0.8992

Note: (a) adjusted R-squared: 0.4645; F-statistic: 47.84; p-value: 0.000; ARCH test: 0.000; JB test: 0.000; (b) adjusted R-squared: 0.2608; F-statistic: 20.05; p-value: 0.000; ARCH test: 0.000; JB test: 0.000; (c) Adjusted R-squared: 0.1971; F-statistic: 14.25; p-value: 0.000; ARCH test: 0.000; JB test: 0.000; significance codes: '***' < 0.001; '**' < 0.01; '*' < 0.05.

Moreover, the results find that investors' sentiments are significantly associated with their own past time series. The Jarque–Bera (JB) test, the autoregressive conditional heteroscedastic (ARCH) test and the forecast error variance decomposition (FEVD) test are estimated. The JB test indicates that the residuals are not normally distributed. The ARCH test shows that the variables suffer from the ARCH effects. Figure 3 reveals that cost-based market liquidity and pessimistic investor sentiments are strongly influenced by their own variance shocks. Investors' optimistic sentiments are influenced by their own exogenous shocks and negative sentiments variance shocks.

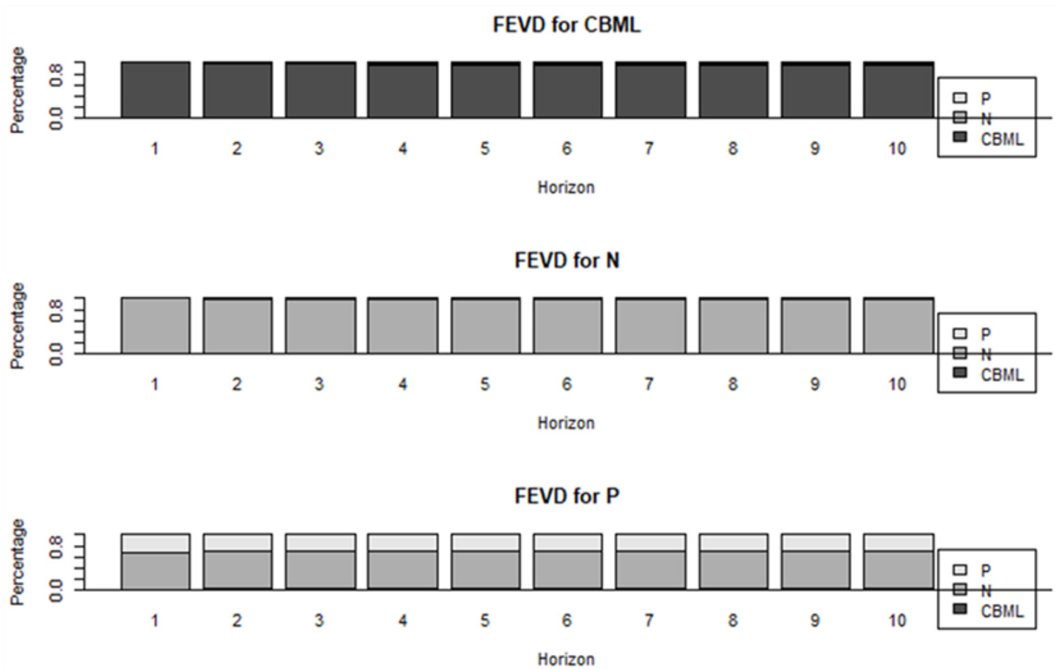


Figure 3. FEVD analysis for CBML and investor sentiments.

Based on the VAR model estimation results, the dispersion of the asset price and its trading quantity are not significantly explained by the lagged coefficients of the investor optimistic sentiments (Table 4). However, it is observed that price impact volume-based liquidity is significantly associated with its own past time series and the lag_{t-1} of pessimistic investor sentiments. Likewise, pessimistic investor sentiments are not significantly explained by the past time series of price impact volume-based liquidity and optimistic investor sentiments. Optimistic investor sentiments are not significantly associated with the lagged coefficients of price impact volume-based liquidity and pessimistic investor sentiments. Moreover, investors' sentiments are significantly explained by their own past time series. The JB test shows that the residuals are not normally distributed. The ARCH test reports that the variables suffer from the ARCH effects. Figure 4 illustrates that price impact volume-based liquidity and pessimistic investor sentiments are strongly influenced by their own exogenous shocks. Finally, investors' optimistic sentiments are influenced by their own variance shocks and exogenous negative sentiment shocks.

Table 4. Estimation of VAR coefficients and significance test values, MLI model.

Variables		Estimate	p-Value
MLI (a)	$\beta_{11,\mathrm{MLI}}$	− 0.3323240	0.000 ***
	$\gamma_{11,N}$	0.4591262	0.0268 *
	$\emptyset_{11,P}$	0.0901849	0.1156
	$\beta_{12,MLI}$	− 0.2531608	0.000 ***
	$\gamma_{12,N}$	0.0268085	0.8976
	$\emptyset_{12,P}$	0.0373390	0.5145
	α_{MLI}	− 0.0001461	0.9936
N (b)	$\beta_{21,MLI}$	− 0.013135	0.60219
	$\gamma_{21,N}$	− 0.668877	0.000 ***
	$\emptyset_{21,P}$	− 0.029397	0.26646
	$\beta_{22,MLI}$	0.016084	0.51670
	$\gamma_{22,N}$	− 0.276186	0.00437 **
	$\emptyset_{22,P}$	− 0.020703	0.43411
	α_N	− 0.001120	0.89491
P (c)	$\beta_{31,MLI}$	0.065824	0.4712
	$\gamma_{31,N}$	0.099416	0.7738
	$\emptyset_{31,P}$	− 0.474563	0.000 ***
	$\beta_{32,MLI}$	− 0.079112	0.3790
	$\gamma_{32,N}$	− 0.047393	0.8920
	$\emptyset_{32,P}$	− 0.189236	0.0491 *
	α_P	0.003904	0.8989

Note: (a) adjusted R-squared: 0.141; F-statistic: 9.862; p-value: 0.000; ARCH test: 0.000; JB test: 0.000; (b) adjusted R-squared: 0.2616; F-statistic: 20.13; p-value: 0.000; ARCH test: 0.000; JB test: 0.000; (c) adjusted R-squared: 0.1987; F-statistic: 14.39; p-value: 0.000; ARCH test: 0.000; JB test: 0.000; significance codes: ‘***’ < 0.001; ‘**’ < 0.01; ‘*’ < 0.05.

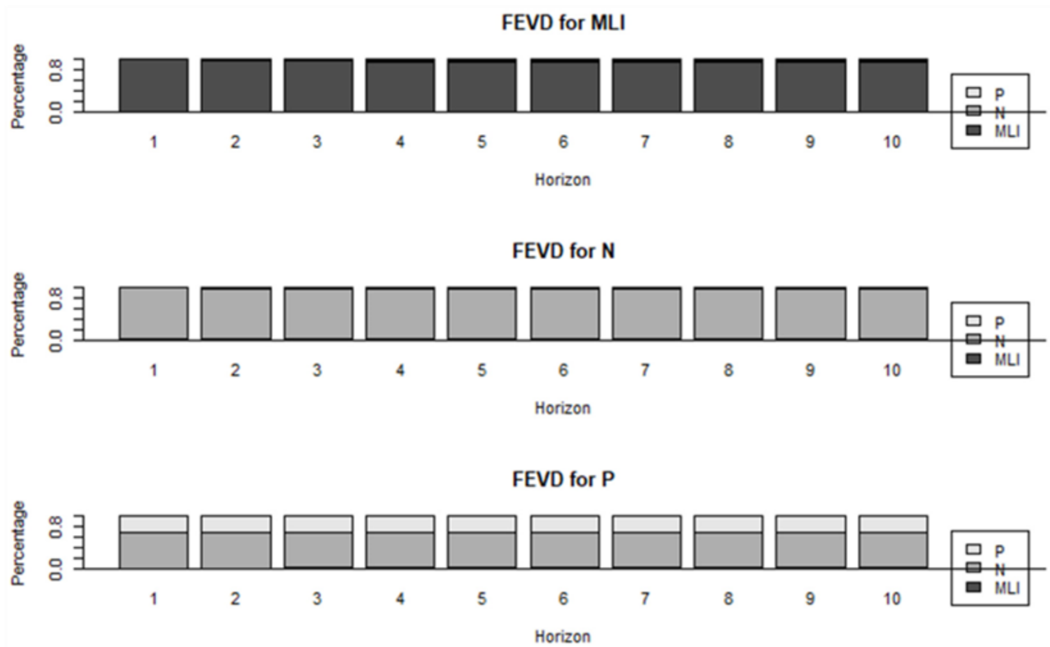


Figure 4. FEVD analysis for MLI and investor sentiments.

5. Conclusions

This research analyzed whether investor sentiments based on microblogging content influences the dimensions of market liquidity. Using time series and multivariate statistical models, the influence of investor sentiments on the liquidity of the Australian Securities Exchange was studied. To estimate investor sentiments, posts on the popular social network platform Twitter were analyzed and different liquidity measures were applied to estimate the relationship between microblogging content and liquidity dimensions. We found that investor sentiments in pessimistic periods were significantly associated with higher trading cost, illiquidity, higher price dispersion and lower trading volume. However, cost-based market liquidity and price impact volume-based liquidity were not significantly explained by optimistic investor sentiments.

From the multivariate model approach, market liquidity and its associated cost were not significantly associated with the past time series of pessimistic and optimistic investor sentiments. In contrast, price impact volume-based liquidity was found to be positive and was significantly explained by lagged pessimistic investor sentiments. Likewise, a significant relationship was found between market liquidity dimensions and their own past time series. Finally, market liquidity dimensions were discovered to be strongly influenced by their own variance shocks.

This research has important implications in terms of revealing the relationship between microblogging content and the various dimensions of liquidity that previous studies have ignored. This quantification of investor sentiments based on microblogging content may be useful for liquidity risk management and portfolio construction. Although the study fills a gap in the behavioral finance literature, the geographical dataset that was employed is a limiting element of the study. As this study covers the Australian market, the results may not be generalizable to other markets. The analysis therefore encourages other researchers to uncover the impact of microblogging content on liquidity dimensions at both the industry and the company level. This would undoubtedly provide insight into the authoritative role of microblogging content on liquidity dimensions more broadly.

References

Amihud, Yakov, and Haim Mendelson. 1991. Liquidity, maturity, and the yields on U.S. treasury securities. The Journal of Finance 46: 1411–25. [CrossRef]

Baker, Malcolm, and Jeffrey Wurgler. 2006. Investor sentiment and the cross-section of stock returns. The Journal of Finance 61: 1645–80. [CrossRef]

Bank, Semra, Evrim E. Yazar, and Ugur Sivri. 2019. Can social media marketing lead to abnormal portfolio returns? European Research on Management and Business Economics 25: 54–62. [CrossRef]

Bartov, Eli, Lucile Faurel, and Partha S. Mohanram. 2018. Can Twitter help predict firm-level earnings and stock returns? The Accounting Review 93: 25–27. [CrossRef]

Brown, Gregory W., and Michael T. Cliff. 2005. Investor sentiment and asset valuation. The Journal of Business 78: 405–40. [CrossRef]

DeGennaro, Ramon P., and Cesare Robotti. 2007. Financial Market Frictions. Economic Review 92: 1–16.

Easley, David, and Maureen O'Hara. 1992. Time and the process of security price adjustment. The Journal of Finance 47: 577–605. [CrossRef]

Easley, David, and Maureen O'Hara. 2004. Information and the cost of capital. The Journal of Finance 59: 1553–83. [CrossRef]

Fang, Lily, and Joel Peress. 2009. Media coverage and the cross-section of stock returns. The Journal of Finance 64: 2023–52. [CrossRef]

Glosten, Lawrence R., and Paul R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14: 71–100. [CrossRef]

Gorton, Gary, and Andrew Metrick. 2010. Haircuts. Federal Reserve Bank St Louis Review 92: 507–20. [CrossRef]

Gregoriou, Andros. 2013. Earnings announcements and the components of the bid-ask spread: Evidence from the London stock exchange. Journal of Economic Studies 40: 112–26. [CrossRef]

Groß-Klußmann, Axel, and Nikolaus Hautsch. 2011. When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. Journal of Empirical Finance 18: 321–40. [CrossRef]

Guijarro, Francisco, Ismael Moya-Clemente, and Jawad Saleemi. 2019. Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index. Sustainability 11: 7048. [CrossRef]

Hasbrouck, Joel, and Duane J. Seppi. 2001. Common Factors in Prices, Order Flows, and Liquidity. Journal of Financial Economics 59: 383–411. [CrossRef]

Ho, Thomas, and Hans R. Stoll. 1981. Optimal dealer pricing under transactions and return uncertainty. Journal of Financial Economics 9: 47–73. [CrossRef]

Kurov, Alexander. 2010. Investor sentiment and the stock market's reaction to monetary policy. Journal of Banking and Finance 34: 139–49. [CrossRef]

Le, Huong, and Andros Gregoriou. 2020. How do you capture liquidity? A review of the literature on Low-frequency stock liquidity. Journal of Economic Surveys 34: 1170–86. [CrossRef]

Lee, Wayne Y., Christine X. Jiang, and Daniel C. Indro. 2002. Stock market volatility, excess returns, and the role of investor sentiment. Journal of Banking and Finance 26: 2277–99. [CrossRef]

Li, Qing, Yan Chen, Jun Wang, Yuanzhu Chen, and Hsinchun Chen. 2018. Web media and stock markets: A survey and future directions from a big data perspective. IEEE Transactions on Knowledge and Data Engineering 30: 381–99. [CrossRef]

Mazboudi, Mohamad, and Samer Khalil. 2017. The attenuation effect of social media: Evidence from acquisitions by large firms. Journal of Financial Stability 28: 115–24. [CrossRef]

Oh, Chong, and Olivia Sheng. 2011. Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement. Paper presented at the International Conference on Information Systems, Shanghai, China, December 4–7.

Oliveira, Nuno, Paulo Cortez, and Nelson Areal. 2017. The impact of microblogging data for stock market prediction: Using twitter to predict returns, volatility, trading volume and survey sentiment indices. Expert Systems with Applications 73: 125–44. [CrossRef]

Prokofieva, Maria. 2015. Twitter-based dissemination of corporate disclosure and the intervening effects of firms' visibility: Evidence from Australian-listed companies. Journal of Information Systems 29: 107–36. [CrossRef]

Roll, Richard. 1984. A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. The Journal of Finance 39: 1127–39. [CrossRef]

Saleemi, Jawad. 2020. An estimation of cost-based market liquidity from daily high, low and close prices. Finance, Markets and Valuation 6: 1–11. [CrossRef]

Saleemi, Jawad. 2021. COVID-19 and liquidity risk, exploring the relationship dynamics between liquidity cost and stock market returns. National Accounting Review 3: 218–236. [CrossRef]

Sarkissian, Jack. 2016. Option pricing under quantum theory of securities price formation. SSRN Electronic Journal. [CrossRef]

Sarr, Abdourahmane, and Tonny Lybek. 2002. Measuring liquidity in financial markets. International Monetary Fund 2: 1–64. [CrossRef]

Smailovic´, Jasmina, Miha Grcˇar, Nada Lavracˇ, and Martin Žnidaršicˇ. 2013. Predictive sentiment analysis of Tweets: A stock market application. In Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data. Berlin/Heidelberg: Springer. [CrossRef]

Sprenger, Timm O., Andranik Tumasjan, Philipp G. Sandner, and Isabell M. Welpe. 2014. Tweets and trades: The information content of stock microblogs. European Financial Management 20: 926–57. [CrossRef]

Walker, Clive B. 2016. The direction of media influence: Real-estate news and the stock market. Journal of Behavioral and Experimental Finance 10: 20–31. [CrossRef]

Wei, Cen, Li Shihao, and Tong Naqiong. 2014. The Influence of Investor Attention on the Stock Return and Risk: An Empirical Study Based on the “Easy Interactive” Platform Data of Shenzhen Stock Exchange. Securities Market Herald 7: 40–47.

Yu, Yang, Wenjing Duan, and Qing Cao. 2013. The impact of social and conventional media on firm equity value: A sentiment analysis approach. Decision Support Systems 55: 919–26. [CrossRef]

Zhang, Xue, Hauke Fuehres, and Peter A. Gloor. 2011. Predicting stock market indicators through Twitter “i hope it is not as bad as I fear”. Procedia-Social and Behavioral Sciences 26: 55–62. [CrossRef]

Investor Sentiments and Liquidity Pricing: Applying the Microblogging Content to the Systematic Risk

Abstract: Investors are keenly interested in the risk of informed trading, as it can have an immediate impact on transaction costs imposed by liquidity providers. This paper examines microblogging-based informed trading as a systematic risk for liquidity in the market, focusing on the Financial Times Stock Exchange 100 Index (FTSE). Two sub-indices are constructed using the capitalization weighted average technique: one for banks and another for non-financial firms (NFF). During trading sessions, the market index liquidity was priced pessimistically. Meanwhile, the bank index liquidity was significantly exposed to systematic sentiment and liquidity risks. However, the NFF index liquidity was not affected by systematic sentiment risk, but was exposed to systematic liquidity risk. The results of the vector error correction model (VECM) show that the market index liquidity and NFF index liquidity were not associated with investor sentiment in the short or long run. However, bank index liquidity was linked with pessimistic sentiment in the long run. Short-run linkage was observed in the commonality of liquidity between the firm index and market index. Additionally, market index liquidity and firm index liquidity were found to be responsive to standard deviation shocks in investor sentiment. The findings provide valuable insights for investors and liquidity providers to better understand and manage their risks.

Keywords: Microblogging data; Investor Sentiments; Asset Pricing; Liquidity; Systematic Risk

1. Introduction

The development of information technologies has enabled access to vast amounts of information without geographical barriers. The emergence of social media has fundamentally transformed the analysis of sentiment-driven market participants (Dugast and Foucault, 2018), and a rich literature in behavioral finance has linked social networking to trading activities (Ekinci and Bulut, 2021).

In the context of user-generated information, social media can be particularly important and economically significant (Broadstock and Zhang, 2019). Extracting information from microblogging data provides a deeper understanding of the sentiment-driven behavior of financial agents (Sprenger et al., 2014). Microblogging platforms cover almost all aspects of society and can also serve decision-making purposes, including those in the financial sector.

The findings in the financial domain are multifaceted, and there is no unified approach to conclusively establish the authoritative role of microblogging-based sentiments on different attributes of the financial market (Oliveira et al., 2017; Guijarro et al., 2019). In this debate, there is still room to examine the root influence of microblogging content on the liquidity-facilitating cost across the market. This phenomenon can be even more significant when considering whether investor sentiments within a broader market are priced in the systematic liquidity risk.

Liquidity is understood as the ease of redeeming an asset while incurring lower costs. Liquidity facilitators tend to reduce their risk against informed traders (Gorton and Metrick, 2010), leading to costs borne by the counterparty, such as a higher bid-ask spread (Saleemi, 2020). A large spread size indicates illiquidity or higher conditioning costs to facilitate liquidity for financial assets. The informed counterparty impacts trading, and its risk should be priced by liquidity providers (Saleemi, 2022).

The aim of this study is to investigate whether microblogging-based investor sentiments in a systematic context are exposed to liquidity risk across the market. Commonality in liquidity between the market and its individual assets is often attributed to a common market (Brunnermeier and Pedersen, 2009). Therefore, it is essential to understand the impact of systematic sentiment risk on liquidity at both market and firm levels. Our research represents the first attempt to examine this aspect, and it has significant implications for the management of systematic liquidity risk more broadly.

The paper is structured as follows: Section 2 provides a brief review of the literature, while Section 3 discusses the benchmark models and the data collection process. Section 4 presents the research findings, and Section 5 summarizes the main outcomes of the study.

2. Literature Review

Sentiment analysis is a subfield of natural language processing that can assist in analyzing investor opinions, particularly in the context of binary quantification (bullish vs bearish) or multi-level attributes. Measurement of investor emotions on social media has emerged as a popular research topic in recent years (Oliveira et al., 2013; Poria et al., 2017). The fundamental value of the investment is crucial in executing financial transactions (Cervelló-Royo and Guijarro, 2020).

One social media platform that has gained popularity for modeling financial securities is microblogging, particularly Twitter (Sprenger et al., 2014). Quantifying microblogging data can provide insights into market and investor information (Zhang et al., 2022). However, the unstructured nature of microblogging data in its initial stages necessitates the application of sentiment analysis to arrange it for further analysis. Identifying patterns from a large amount of information can be a critical factor for investors (Guijarro et al., 2019).

Analysis of the extracted content from the microblogging network can provide insights into various aspects of the market behavior, including returns (Groß-Klußmann and Hautsch, 2011), price directional movements (Oh and Sheng, 2011; Smailović et al, 2013), market performance (Yu et al., 2013), stock trends (Li et al., 2018), firm’s earnings (Bartov et al., 2018; Bank et al., 2019), and market liquidity dimensions (Guijarro et al., 2021). Microblogging content may also be accessed more conveniently on a real-time basis than traditional sentiment measures (Oliveira et al., 2017).

Alleviating rumors related to investment concerns on social media is crucial for the market, investors, and corporations alike (Wei et al., 2014). Business engagement on microblogging networks can reduce information asymmetry (Prokofieva, 2015) and mitigate bearish market reactions (Mazboudi and Khalil, 2017). Rumors regarding earning expectations in the market can also influence transaction execution (Chen et al., 2011; Zhang et al., 2022).

Market liquidity is an essential indicator of asset value in the financial market (Amihud, 2002; Easley and O’Hara, 2004; Corwin and Schultz, 2012). Specialists secure trading against the risk of an informed counterparty (Glosten and Milgrom, 1985; Saleemi, 2020), which is often considered a priced factor (Amihud et al., 2015; Saleemi, 2022). Information transparency about the fundamental value of an asset is critical in determining market liquidity (Gorton and Metrick, 2010).

Market liquidity can impact the cost of capital (Acharya and Pedersen, 2005), corporate investment decisions (Amihud and Mendelson, 2008), funding liquidity (Brunnermeier and Pedersen, 2009), asset prices (Bao et al., 2011), and yields on investment (Amihud et al., 2015). Investors are particularly concerned with uncertainty related to liquidity (Brunnermeier and Pedersen, 2005), and liquidity is considered a priced risk factor in uncertain environments (Saleemi, 2021).

The concept of liquidity is a multidimensional debate and there is currently no unified method for its estimation in the financial market (Goyenko et al., 2009; Abdi and Ranaldo, 2017). Over time, several models focusing either on bid-ask spread or price impact volume have been proposed. The bid-ask spread represents the transaction immediacy at possible trading cost (Roll, 1984; Corwin and Schultz, 2012), while another stream in the field emphasizes the relationship between price variations and trading quantity (Amihud, 2002).

Despite specific assumptions in the construction of different spread models, market frictions are common determinants of liquidity (Degennaro and Robotti, 2007). These frictions can be classified into explicit costs, such as taxes or brokerage fees that are generally observable in advance of trading, and implicit costs, which represent a large fraction of the total trading cost and are less observable before the transaction takes place.

The spread is a popular cost-based liquidity proxy that estimates almost all costs associated with trading (Huang and Stoll, 1997; Sarr and Lybek, 2002). An asset is quoted in two major elements: the ask (high) price and the bid (low) price. Market makers would accept an inventory at the lowest bid price and redeem the position at the best highest ask price, earning yields on the investment. The spread size indicates the cost that the liquidity supplier tends to impose on the counterparty. A higher spread reflects illiquidity in the market (Corwin and Schultz, 2012).

3. Data Sampling and Benchmark Models

This study aims to contribute to the debate on systematic risk by exploring the potential of microblogging data in determining liquidity-facilitating costs for individual assets and their respective markets. To achieve this, we adopted a weighted market capitalization technique and created an index of banks and non-financial firms listed in the FTSE market. Specifically, Table 1 provides a detailed overview of the stocks considered, with all banks listed in the market included, while other firms were selected through a simple random sampling technique to enable a broader study of systematic risk.

Table 1. List of assets in the data sampling.

	Stocks	Symbol	Speciality
Banks	Standard Chartered	STAN.L	Banking & Financial Services
	NatWest Group	NWG.L	Banking & Financial Services
	Lloyds Banking Group	LLOY.L	Banking & Financial Services
	HSBC	HSBA.L	Banking & Financial Services
	Barclays	BARC.L	Banking & Financial Services
Non-Financial Firms (NFF)	Antofagasta	ANTO.L	Mining
	Ashtead Group	AHT.L	Support Services
	Associated British Foods	ABF.L	Food Producers
	AstraZeneca	AZN.L	Pharmaceuticals & Biotechnology
	Auto Trader Group	AUTO.L	Media
	AVEVA Group	AVV.L	Software & Computer Services
	BAE Systems	BA.L	Aerospace & Defense
	Barratt Developments	BDEV.L	Household Goods & Home Construction
	Convatec Group	CTEC.L	Health Care

To establish a link between individual assets and their respective markets, we constructed a firm index following (1),

$$IC_t = IC_{t-1}(1 + IR_t) \tag{5}$$

where IC_t (IC_{t-1}) refers to the index closing price of day t ($t-1$), and IR_t states the index yield of day t. IR_t is estimated according to (2):

$$IR_t = \sum_{i=1}^{n} WMC_{i,t}\left[LN\left(\frac{C_{i,t}}{C_{i,t-1}}\right)\right] \tag{6}$$

where $WMC_{i,t}$ denotes the weighted market capitalization of the individual asset on day t; and $C_{i,t}$ ($C_{i,t-1}$) represents the closing price of the asset on day t ($t-1$). The weighted market capitalization of the stock is computed as (3).

$$WMC_{i,t} = \frac{(S_{i,t} \times C_{i,t})}{\sum_{i=1}^{n} MC_{i,t}} \tag{7}$$

$S_{i,t}$ indicates the outstanding shares of individual securities on day t; and $\sum_{i=1}^{n} MC_{i,t}$ depicts the accumulated market capitalization of assets on day t. The market capitalization, $MC_{i,t}$, is estimated by multiplying the outstanding shares of an asset by its closing price on day t. IH_t reflects the index highest price of day t; $H_{i,t}$ ($H_{i,t-1}$) denotes the highest price of asset i on day t ($t-1$); IL_t the index lowest price of day t; $L_{i,t}$ ($L_{i,t-1}$) the lowest price of asset i on day t ($t-1$).

$$IH_t = \left(\sum_{i=1}^{n} WMC_{i,t}\left[1 + LN\left(\frac{H_{i,t}}{H_{i,t-1}}\right)\right]\right) IC_t \tag{8}$$

$$IL_t = \left(\sum_{i=1}^{n} WMC_{i,t}\left[1 + LN\left(\frac{L_{i,t}}{L_{i,t-1}}\right)\right]\right) IC_t \tag{9}$$

The liquidity is estimated using the cost-based market liquidity (CBML) approach. Recognizing the presence of asymmetric information during trading, the CBML model can effectively estimate liquidity and its associated facilitating cost (Saleemi, 2020). The CBML method is formulated according to Equation (6).

$$CBML_t = \sqrt{\left[\left(\frac{Range_{t-1}}{EP_{t-1}}\right) - E_t^S\right]^2} \tag{10}$$

Here, EP_{t-1} refers to the execution price of the transaction on day *t-1*, and $Range_{t-1}$ represents the difference between the highest and lowest quoted prices of the previous trading session. Equation (7) models asymmetric information, assuming equal probability for the informed trader.

$$E_t^S = \frac{E[ask_t] - E[bid_t]}{EP_t} \tag{11}$$

EP_t represents the execution price of the transaction on day t, while $\mathrm{E}[\mathrm{ask_t}]$ denotes the expected highest price at which a liquidity provider may be willing to redeem the financial position, and $\mathrm{E}[\mathrm{bid_t}]$ indicates the expected lowest value that a liquidity provider would pay to accept the financial inventory. The calculation of $\mathrm{E}[\mathrm{ask_t}]$ is contingent upon a trade, as specified in Equation (8):

$$E[ask_t] = H_t\theta + \left(\frac{QS_t}{2}\right)\theta \tag{12}$$

Here, θ represents the probability of asymmetric information, H_t denotes the highest quoted price on day t, and QS_t is the sum of the quoted prices during the same trading session. The calculation of $E[\mathrm{bid}_t]$ is conditioned on a transaction and can be expressed as (9):

$$E[bid_t] = L_t\theta + \left(\frac{QS_t}{2}\right)\theta \tag{13}$$

Similarly, L_t represents the lowest quoted price on day t. The liquidity measure is derived from low-frequency data, covering the period from June 05, 2020, to October 27, 2022. The attributes of the low-frequency data pertain to the closing, highest, and lowest prices (CHL).

To analyze unstructured microblogging data and gain insights into liquidity-providing costs, the R programming language was employed. The microblogging data was initially organized according to market symbols, such as FTSE 100, and collected for the period from June 05, 2020, to October 27, 2022. To prepare the unstructured text for further processing and construct sentiment indicators, the text underwent cleaning using the "NLP" and "tm" libraries. This cleaning process involved removing punctuation, stop words, trailing spaces, and converting the text to lowercase.

Each tweet was classified as either bullish or bearish, with neutral market participants excluded from the analysis. Given the large volume of data for day t, the process of aggregating sentiments is illustrated in Equations (10) and (11):

$$\sum_{t=1}^{T} Bullish_t = Bullish_1 + Bullish_2 + Bullish_3 + \cdots + Bullish_T \tag{14}$$

$$\sum_{t=1}^{T} Bearish_t = Bearish_1 + Bearish_2 + Bearish_3 + \cdots + Bearish_T \quad (15)$$

where T represents the total number of bullish or bearish sentiments on day t. $\sum_{t=1}^{T} Bullish_t$ denotes the cumulative bullish score for the day, while $\sum_{t=1}^{T} Bearish_t$ represents the aggregated bearish score for the day. This aggregation process was performed using the "syuzhet" and "lubridate" libraries.

Furthermore, Equation (12) examines the linear regression relationship between variables. The liquidity cost of the FTSE market is selected as the response variable, while the sentiment indicators serve as explanatory variables.

$$MLC_t = \alpha + \beta_1 Bearish_t + \beta_2 Bullish_t + \epsilon_t \quad (16)$$

MLC_t represents the cost associated with facilitating liquidity for the entire market on day t. $Bearish_t$ reflects the aggregated negative sentiments for the day, while $Bearish_t$ represents the accumulated positive sentiments for the day. ϵ_t represents the error term. Equation (6) is utilized to estimate market liquidity and its associated facilitating cost.

Additionally, the dataset for the same trading session is examined to determine whether individual assets are exposed to systematic sentiment and liquidity risk. In this context, the dataset is modeled according to Equation (13):

$$ILC_t = \alpha + \beta_1 Bearish_t + \beta_2 Bullish_t + \beta_3 MLC_t + \epsilon_t \quad (17)$$

where ILC_t illustrates the liquidity-facilitating cost of the bank Index or non-financial firm Index on day t (6).

The Vector Error Correction Model (VECM) explores the dynamic relationships between time series variables in both the short run and the long run. Specifically, it analyzes the effects of changes in market liquidity-facilitating costs on day t, taking into account not only its own lagged changes but also the historical variations in investor sentiments (14).

$$\Delta MLC_t = \beta_0 + \sum_{i=1}^{n} \delta_i \, \Delta MLC_{t-i} + \sum_{i=1}^{n} \emptyset_i \Delta \, Bearish_{t-i} + \sum_{i=1}^{n} \gamma_i \Delta \, Bullish_{t-i} + \varphi ECT_{t-1} + \epsilon_t \quad (18)$$

where ΔMLC_t (ΔMLC_{t-i}) represents the change in the liquidity-providing cost of the entire market on day t ($t-i$); $\Delta Bearish_{t-i}$ and $\Delta Bullish_{t-i}$ indicate the previous changes in bearish and bullish sentiments, respectively, on day $t-i$; ECT_{t-1} represents the error correction term of day $t-1$. The optimal lags are derived using the Hannan-Quinn (HQ) criterion technique, and their values are provided in Equations (15)-(17) :

$$\Delta MLC_{t-i} = \delta_1 \Delta MLC_{t-1} + \delta_2 \Delta MLC_{t-2} + \delta_3 \Delta MLC_{t-3} + \delta_4 \Delta MLC_{t-4} \quad (19)$$

$$\Delta Bearish_{t-i} = \emptyset_1 \Delta Bearish_{t-1} + \emptyset_2 \Delta Bearish_{t-2} + \emptyset_3 \Delta Bearish_{t-3} + \emptyset_4 \Delta Bearish_{t-4} \quad (20)$$

$$\Delta Bullish_{t-i} = \gamma_1 \Delta Bullish_{t-1} + \gamma_2 \Delta Bullish_{t-2} + \gamma_3 \Delta Bullish_{t-3} + \gamma_4 \Delta Bullish_{t-4} \quad (21)$$

Equation (18) investigates the relationship between the change in the liquidity-facilitating cost of individual assets on day t and its corresponding previous changes, as well as previous changes in sentiment indicators and cost-based liquidity for the entire market:

$$\Delta ILC_t = \beta_0 + \sum_{i=1}^{n} \psi_i \, \Delta ILC_{t-i} + \sum_{i=1}^{n} \emptyset_i \Delta \, Bearish_{t-i} + \sum_{i=1}^{n} \gamma_i \Delta \, Bullish_{t-i} + \sum_{i=1}^{n} \delta_i \, \Delta MLC_{t-i} + \varphi ECT_{t-1} + \epsilon_t \quad (22)$$

where ΔILC_t (ΔILC_{t-i}) represents the change in the cost-based liquidity of the bank index or non-financial firm index on day t ($t - i$). Using the Hannan-Quinn (HQ) criterion approach, the optimal lags are computed using Equation (19):

$$\Delta ILC_{t-i} = \psi_1 \Delta ILC_{t-1} + \psi_2 \Delta ILC_{t-2} + \psi_3 \Delta ILC_{t-3} + \psi_4 \Delta ILC_{t-4} \quad (23)$$

4. Analysis and Discussion

Table 2 displays the descriptive attributes of the data sampling. The analysis indicates that the variables exhibit positive skewness along with higher kurtosis values. The positive skewness indicates a right-skewed distribution, where the majority of numeric values are situated to the right of the mean. The higher kurtosis signifies a fat-tailed distribution within the numerical dataset. To visually represent the dataset, Figure 1 plots the data measurements, revealing the non-constant behavior of the variables. The fluctuations observed among the variables are first assessed as a linear combination within the same trading session.

Table 2. Descriptive attributes (daily basis).

Variables	Min	Median	Mean	Max	SD	Skewness	Kurtosis
MLC	0.0039	0.5597	0.7298	4.6253	0.6084	1.7031	7.3657
Bearish	0.030	0.840	1.291	14.660	1.4216	4.2639	29.2174
Bullish	0.090	0.990	1.589	30.560	1.9128	7.1787	92.6034
ILC_B	0.0064	1.1411	1.4595	11.8198	1.2734	3.4752	23.2924
ILC_{NFF}	0.0088	1.0828	1.2739	13.0906	1.0125	4.5543	41.3208

Note: Liquidity cost for the entire market (MLC); Bank Index Liquidity cost (ILC_B); Liquidity cost for index of non-financial firms (ILC_{NFF}); Standard deviation (SD); Significance level codes: *** < 0.001; ** < 0.01; * < 0.05.

The model presented in Equation (12) employs investor sentiments as predictors of the liquidity-facilitating cost for the FTSE market. The findings from Table 3 suggest a significant and positive association between pessimistic sentiments and cost-based liquidity. This relationship indicates that an increase in negative sentiments leads to a wider spread size of the market index. A larger spread size is often associated with higher transaction costs and increased market illiquidity. During periods of pessimism, liquidity providers are more hesitant to accept financial inventory without imposing higher costs on the counterparty. As a result, negative sentiments appear to be priced into the overall market liquidity. Conversely, the underlying drivers of optimistic investor sentiments towards market index liquidity have not been identified.

To further explore the presence of systematic sentiment and liquidity risks at the individual asset level, the stocks are divided into financial and non-financial sectors, and corresponding indices are constructed using Equations (1)-(5). This approach enables a more comprehensive analysis of systematic risk. The model specified in Equation (13) investigates either the bank index or the NFF index in terms of the common market for liquidity and investor sentiments.

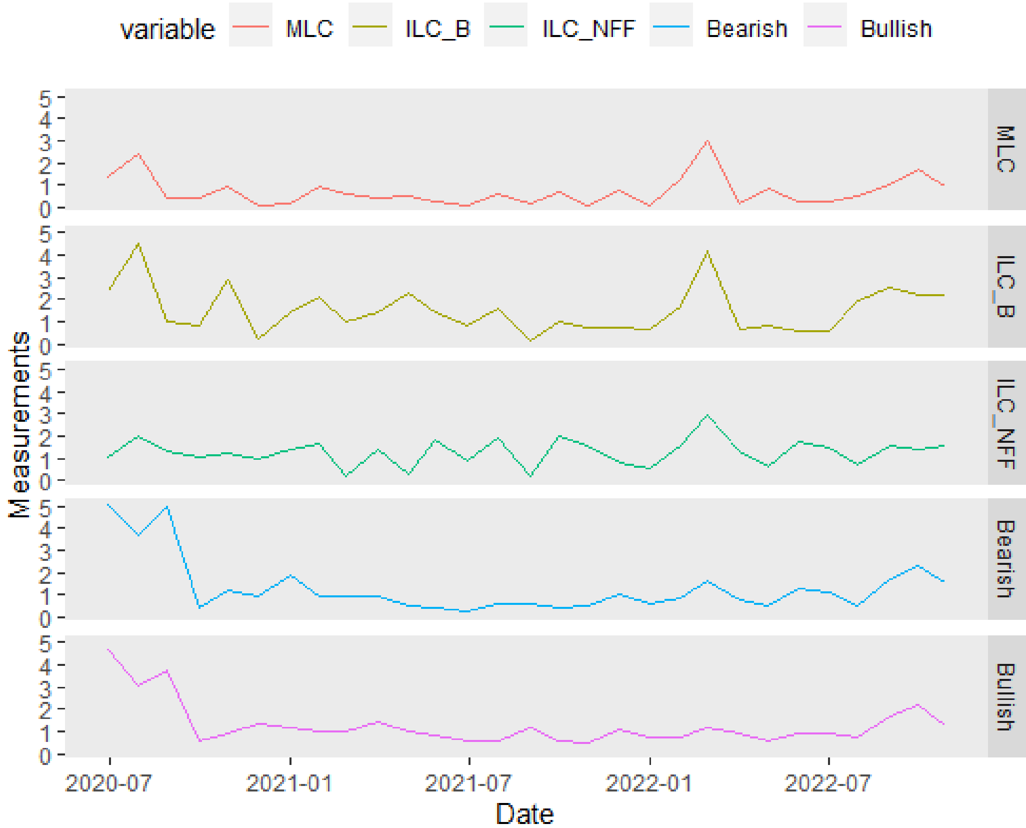


Figure 1. Time-varying measurements of different variables (monthly basis).

Table 3 reports that the liquidity-facilitating cost for the bank index is positively and significantly influenced by market pessimistic sentiments. This indicates that during bearish market periods, the spread size of the bank index increases. A wider spread signifies the liquidity provider's reluctance to accept the bank index without imposing higher costs on the counterparty. The analysis reveals that the bank index exhibits a significant response to pessimistic sentiments within the broader market, highlighting the pricing of bearish sentiments in the liquidity of banking assets.

In contrast, optimistic sentiments in the broader market show a negative and significant correlation with bank index spreads. This relationship suggests that the spread size of the bank index decreases in response to positive investor sentiments. As the FTSE market enters optimistic periods, liquidity providers seem more willing to execute bank index transactions at a lower cost, resulting in increased liquidity for bank stocks. Additionally, the liquidity of the bank index is positively influenced by the liquidity of its corresponding market index. These findings indicate that bank stocks are exposed to both systematic sentiment risk and systematic liquidity risk.

However, the cost of providing liquidity for the non-financial firm index is not significantly explained by bearish and bullish sentiments. Nevertheless, the liquidity of individual non-financial assets is positively associated with the liquidity of the corresponding market index. Therefore, non-financial firms appear to be less impacted by systematic sentiment risk but remain exposed to systematic liquidity risk.

Table 3. Regression quantification (daily basis).

Variables		Estimate	p-value
MLC (I)	Intercept	0.6335	0.000 ***
	Bearish	0.0613	0.017 *
	Bullish	0.0108	0.571
ILC_B (II)	Intercept	0.7233	0.000 ***
	Bearish	0.2548	0.000 ***
	Bullish	-0.1348	0.000 ***
	MLC	0.8514	0.000 ***
ILC_{NFF} (III)	Intercept	0.9236	0.000 ***
	Bearish	-0.0276	0.510
	Bullish	0.0360	0.246
	MLC	0.4505	0.000 ***

Note: I) Adjusted R-squared: 0.025; F-statistic: 8.961; p-value: 0.000; (II) Adjusted R-squared: 0.213; F-statistic: 55.75; p-value: 0.000; (III) Adjusted R-squared: 0.072; F-statistic: 16.74; p-value: 0.000.

To examine the dynamics of the relationship, a VECM approach is utilized, starting with the assessment of unit roots and cointegration in the system. The Augmented Dickey-Fuller (ADF) test results, shown in Table 4, indicate stationarity in the time series. Cointegration, denoted as term r in Table 5, is analyzed using the Johansen technique. The Trace statistics exceeding the critical values suggest the presence of cointegration among the time series variables.

The VECM model presented in Equation (14) investigates the relationship between changes in the cost of accepting positions in the FTSE index on day t and its own previous changes, as well as past changes in bearish and bullish sentiments. The results for the optimal lags, based on Equations (15)-(17), are reported in Table 6. The findings indicate that changes in the cost ΔMLC_t are not significantly explained by previous changes in investor sentiments. This suggests that changes in market index liquidity on day t are not influenced by past changes in investor sentiments in either the short or long run. However, changes in the cost of facilitating liquidity for the market index on day t are associated with its own past series, with the exception of lag $t-4$.

To examine systematic risk, Equation (18) is employed in the VECM approach, where changes in the liquidity of the bank index for the following trading session are analyzed in relation to corresponding lags and past series changes of other variables. The results, presented in Table 6, show that changes in $\Delta ILC_{B,t}$ (bank index liquidity) are not significantly correlated with changes in previous sentiment series, except for lag (t-4) of bearish sentiments. This suggests a long-run association between bank index liquidity and bearish sentiments. Conversely, changes in the liquidity-facilitating cost for the bank index on day t are linked to changes in the past series of market index liquidity, except for lags $t-3$ and $t-4$. This indicates a short-run relationship between bank index liquidity and its corresponding market liquidity. Furthermore, changes in the cost of providing liquidity for the bank index on day t are significantly explained by changes in its own previous series.

Table 4. Unit roots test.

Variables	ADF Statistics	p-value	1% CV	5% CV	10% CV
MLC	-6.8945	0.000	-2.58	-1.95	-1.62
Bearish	-5.8757	0.000	-2.58	-1.95	-1.62
Bullish	-6.9173	0.000	-2.58	-1.95	-1.62
ILC_B	-7.3347	0.000	-2.58	-1.95	-1.62
ILC_{NFF}	-7.266	0.000	-2.58	-1.95	-1.62

Note: Critical value (CV).

Table 5. Cointegration analysis results.

Cointegrated Relationship	Trace Statistics	10% CV	5% CV	1% CV
MLC & Sentiments				
$r > 2$	22.64	7.52	9.24	12.97
$r > 1$	99.52	17.85	19.96	24.60
$r > 0$	198.78	32.00	34.91	41.07
ILC_B, Sentiments & MLC				
$r > 3$	21.82	7.52	9.24	12.97
$r > 2$	99.55	17.85	19.96	24.60
$r > 1$	195.77	32.00	34.91	41.07
$r > 0$	348.77	49.65	53.12	60.16
ILC_{NFF}, Sentiments & MLC				
$r > 3$	23.19	7.52	9.24	12.97
$r > 2$	100.46	17.85	19.96	24.60
$r > 1$	199.36	32.00	34.91	41.07
$r > 0$	352.54	49.65	53.12	60.16

Note: $r > 0$: cointegration exists at least one in the system; $r > 1$: cointegrated relationship between two series; $r > 2$: three cointegrated vectors; $r > 3$: cointegration is greater than 3.

Table 6. VECM quantification.

ΔMLC_t	Estimates	$\Delta ILC_{B,t}$	Estimates	$\Delta ILC_{NFF,t}$	Estimates
ECT	-0.506	ECT	-0.642	ECT	-0.587
	(0.069)***		(0.090)***		(0.079)***
Intercept	0.173	Intercept	0.145	Intercept	0.165
	(0.033)***		(0.054)**		(0.047)***
ΔMLC_{t-1}	-0.478	$\Delta ILC_{B,t-1}$	-0.388	$\Delta ILC_{NFF,t-1}$	-0.389
	(0.068)***		(0.082)***		(0.073)***
$\Delta Bearish_{t-1}$	0.042	$\Delta Bearish_{t-1}$	-0.039	$\Delta Bearish_{t-1}$	0.016
	(0.030)		(0.064)		(0.046)
$\Delta Bullish_{t-1}$	0.006	$\Delta Bullish_{t-1}$	0.082	$\Delta Bullish_{t-1}$	0.001
	(0.019)		(0.042)		(0.033)
ΔMLC_{t-2}	-0.368	ΔMLC_{t-1}	-0.528	ΔMLC_{t-1}	-0.398
	(0.065)***		(0.130)***		(0.096)***
$\Delta Bearish_{t-2}$	0.011	$\Delta ILC_{B,t-2}$	-0.305	$\Delta ILC_{NFF,t-2}$	-0.313
	(0.032)		(0.074)***		(0.067)***
$\Delta Bullish_{t-2}$	0.011	$\Delta Bearish_{t-2}$	-0.052	$\Delta Bearish_{t-2}$	-0.007
	(0.022)		(0.071)		(0.053)
ΔMLC_{t-3}	-0.154	$\Delta Bullish_{t-2}$	0.088	$\Delta Bullish_{t-2}$	-0.0003
	(0.056)**		(0.046)		(0.037)
$\Delta Bearish_{t-3}$	0.035	ΔMLC_{t-2}	-0.386	ΔMLC_{t-2}	-0.227
	(0.032)		(0.135)**		(0.103)*
	0.008	$\Delta ILC_{B,t-3}$	-0.196	$\Delta ILC_{NFF,t-3}$	-0.235
$\Delta Bullish_{t-3}$	(0.021)		(0.062)**		(0.056)***
	-0.034	$\Delta Bearish_{t-3}$	-0.054	$\Delta Bearish_{t-3}$	-0.048
ΔMLC_{t-4}	(0.039)		(0.069)		(0.053)
	0.031	$\Delta Bullish_{t-3}$	0.078	$\Delta Bullish_{t-3}$	0.055
$\Delta Bearish_{t-4}$	(0.027)		(0.046)		(0.037)
	0.0198	ΔMLC_{t-3}	-0.151	ΔMLC_{t-3}	-0.059
$\Delta Bullish_{t-4}$	(0.019)		(0.124)		(0.095)
		$\Delta ILC_{B,t-4}$	-0.193	$\Delta ILC_{NFF,t-4}$	-0.144
			(0.044)***		(0.041)***
		$\Delta Bearish_{t-4}$	0.134	$\Delta Bearish_{t-4}$	-0.060
			(0.060)*		(0.047)
		$\Delta Bullish_{t-4}$	0.032	$\Delta Bullish_{t-4}$	0.048
			(0.041)		(0.033)
		ΔMLC_{t-4}	0.088	ΔMLC_{t-4}	0.004
			(0.090)		(0.069)

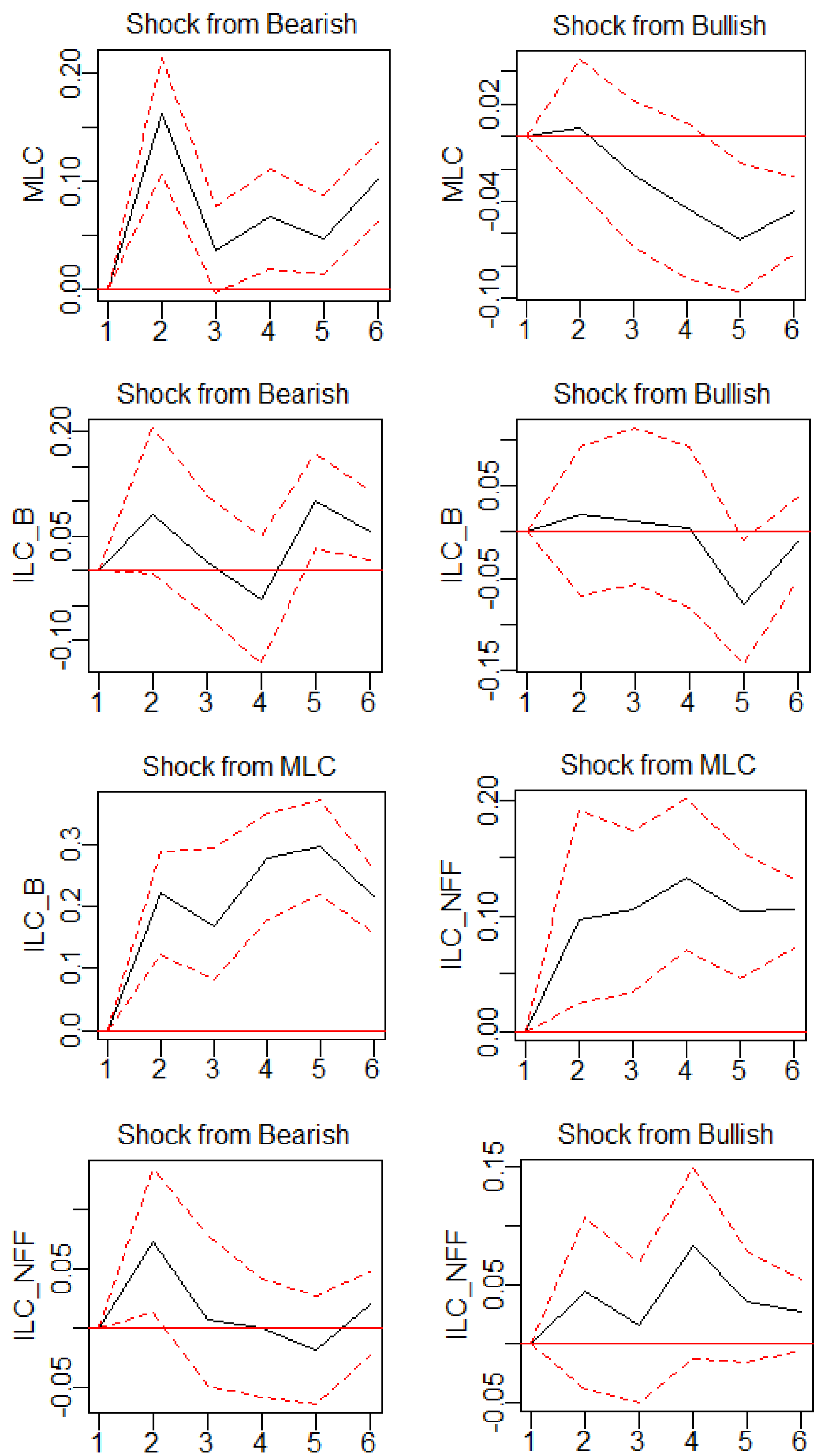


Figure 2. Impulse Response analysis. Bootstrap 95% confidence interval.

Table 6 also reveals that changes in the trading cost for the non-financial firm index on day t are not significantly linked to changes in past sentiment series. Thus, previous changes in sentiment series, whether in the short or long run, are not applicable for estimating $\Delta ILC_{NFF,t}$ for the next trading session. However, changes in the cost of facilitating liquidity for the NFF index on period t are associated with past series changes in market index liquidity, except for lags $t-3$ and $t-4$. This suggests a short-run linkage of liquidity commonality between the non-financial firm index and its corresponding market. Additionally, changes in NFF index liquidity for the next trading period are significantly influenced by changes in its own past series.

Finally, this study conducts an impulse response analysis using the Bootstrap 95% confidence interval, as illustrated in Figure 2. The results demonstrate that the cost of facilitating liquidity for the market index responds to shocks in investor sentiments. Thus, standard deviation shocks in investor sentiments can impact market index liquidity during the observed responsive periods. Similarly, the cost of trading the bank index is influenced by shocks in both investor sentiments and market index liquidity. In this regard, standard deviation shocks in investor sentiments and market index liquidity play a significant role in changing bank index liquidity during each responsive period. Likewise, the liquidity-facilitating cost for the NFF index shows considerable responsiveness to standard deviation shocks in investor sentiments and market index liquidity.

5. Conclusions

This research is focused on analyzing systematic risk by utilizing microblogging data as a source of investor sentiments to examine liquidity pricing across the market. The analysis considered firms categorized by industry, constructing a bank index and a non-financial assets index for analysis purposes. The outcomes aimed to provide insights into systematic liquidity risk within the broader market in relation to investor sentiments.

The findings revealed that during the same trading periods, the cost of accepting positions in the FTSE index was positively influenced by pessimistic investor sentiments. This suggests higher trading costs or illiquidity in response to negative sentiments. Consequently, market index liquidity appears to be influenced by bearish market periods. However, no association was observed between market index liquidity and bullish market periods within the same trading sessions.

During the same trading periods, the liquidity of the bank index was significantly related to investor sentiments and market index liquidity. The cost of facilitating liquidity was positively associated with bearish sentiments, indicating higher transaction costs in response to pessimistic sentiments. Thus, bank index liquidity was priced based on negative investor sentiment. Additionally, the bank index spread showed a negative correlation with optimistic sentiments, implying that a bullish market leads to a decrease in the cost of trading the bank index. Furthermore, bank index liquidity exhibited a positive relationship with market index liquidity. On the other hand, no significant relationship was observed between investor sentiments and the trading cost of the non-financial firm index. Nevertheless, the liquidity of the non-financial firm index was positively linked to its corresponding market index liquidity.

The VECM analysis indicated that changes in market index liquidity for the following trading session were not significantly explained by changes in past sentiment series. This suggests that these variables are not associated in the short or long run. Similarly, changes in the liquidity of the non-financial firm index for the next trading period were not significantly explained by changes in previous sentiment series. Meanwhile, changes in bank index liquidity for the next period were not significantly associated with changes in past sentiment series, except for lag $t-4$ of bearish sentiments. This suggests a long-run relationship between bank index liquidity and pessimistic sentiments. However, changes in bank index liquidity or non-financial firm index liquidity for the following trading session were linked to changes in the past series of market index liquidity, excluding lags $t-3$ and $t-4$. This indicates a short-run linkage of liquidity commonality between individual assets and the market index. Moreover, standard deviation shocks in investor sentiments had a considerable impact on market index liquidity and firm index liquidity. The liquidity of the firm index was also responsive to standard deviation shocks in market index liquidity.

These findings have important implications for quantifying liquidity within a broader market in relation to the systematic risk associated with microblogging-based sentiments. The results may be applicable in managing systematic liquidity risk across the market. However, it is important to acknowledge that there may be limitations related to geographical aspects in this research. In the context of systematic sentiment risk, the findings suggest that researchers should consider including assets from different markets to provide a more comprehensive understanding of systematic risk.

References

Abdi, F., & Ranaldo, A. (2017). A simple estimation of bid-ask spreads from daily close, high, and low prices. *The Review of Financial Studies, 30*(12), 4437–4480. doi:10.1093/rfs/hhx084

Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics, 77*(2), 375-410. doi:10.1016/j.jfineco.2004.06.007

Amihud, Y. (2002). Illiquidity and stock returns cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56. doi:10.1016/S1386-4181(01)00024-6

Amihud, Y., & Mendelson, H. (2008). Liquidity, the value of the firm, and corporate finance. *Journal of Applied Corporate Finance, 20*(2), 32–45. doi:10.1111/j.1745-6622.2008.00179.x

Amihud, Y., Hameed, A., Kang, W., & Zhang, H. (2015). The Illiquidity Premium: International Evidence. *Journal of Financial Economics, 117*(2), 350–368. doi:10.1016/j.jfineco.2015.04.005

Bank, S., Yazar, E. E., & Sivri, U. (2019). Can social media marketing lead to abnormal portfolio returns? *European Research on Management and Business Economics, 25*, 54-62. doi:10.1016/j.iedeen.2019.04.006

Bao, J., Pan, J., & Wang, J. (2011). The Illiquidity of Corporate Bonds. *The Journal of Finance, 66*(3), 911-946. doi:10.1111/j.1540-6261.2011.01655.x

Bartov, E., Faurel, L., & Mohanram, P. (2018). Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review, 93*(3), 25-27. doi:10.2308/accr-51865

Broadstock, D., & Zhang, D. (2019). Social-media and intraday stock returns: The pricing power of sentiment. *Finance Research Letters, 30*(C), 116-123. doi:10.1016/j.frl.2019.03.030

Brunnermeier, M. K., & Pedersen, L. H. (2005). Predatory trading. *The Journal of Finance, 60*(4), 1825–1863. doi:10.1111/j.1540-6261.2005.00781.x

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *The Review of Financial Studies, 22*(6), 2201–2238. doi:10.1093/rfs/hhn098

Cervelló-Royo, R., & Guijarro, F. (2020). Forecasting stock market trend: a comparison of machine learning algorithms. *Finance, Markets and Valuation, 6*(1), 37–49. doi:10.46503/NLUF8557

Chen, H., De, P., Hu, Y., & Hwang, B. H. (2011). Sentiment revealed in social media and its effect on the stock market. *IEEE Statistical Signal Processing Workshop (SSP)*, 25–28. doi:10.1109/SSP.2011.5967675

Corwin, S. A., & Schultz, P. (2012). A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *The Journal of Finance, 67*(2), 719-760. doi:10.1111/j.1540-6261.2012.01729.x

Degennaro, R. P., & Robotti, C. (2007). Financial Market Frictions. *Economic Review, 92*(3), 1-16.

Dugast, J., & Foucault, T. (2018). Data Abundance and Asset Price Informativeness. *Journal of Financial Economics, 130*(2), 367-391. doi:10.1016/j.jfineco.2018.07.004

Easley, D., & O'Hara, M. (2004). Information and the cost of capital. *The Journal of Finance, 59*(4), 1553-1583. doi:10.1111/j.1540-6261.2004.00672.x

Ekinci, C., & Bulut, A. E. (2021). Google search and stock returns: A study on BIST 100 stocks. *Global Finance Journal, 47*, 100518. doi:10.1016/j.gfj.2020.100518

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*, 71–100. doi:10.1016/0304-405X(85)90044-3

Gorton, G., & Metrick, A. (2010). Haircuts. *Federal Reserve Bank St Louis Review, 92*(6), 507–520. doi:10.20955/r.92.507-20

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*(2), 153–181. doi:10.1016/j.jfineco.2008.06.002

Groß-Klußmann, A., & Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Empirical Finance, 18*(2), 321-340. doi:10.1016/j.jempfin.2010.11.009

Guijarro, F., Moya-Clemente, I., & Saleemi, J. (2019). Liquidity Risk and Investors' Mood: Linking the Financial Market Liquidity to Sentiment Analysis through Twitter in the S&P500 Index. *Sustainability, 11*, 7048. doi:10.3390/su11247048

Guijarro, F., Moya-Clemente, I., & Saleemi, J. (2021). Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data. *Journal of Risk and Financial Management, 14*(9), 394. doi:10.3390/jrfm14090394

Huang, R. D., & Stoll, H. R. (1997). The Components of the Bid-Ask Spread: A General Approach. *The Review of Financial Studies, 10*(4), 995–1034. doi:10.1093/rfs/10.4.995

Li, Q., Chen, Y., Wang, J., Chen, Y., & Chen, H. (2018). Web media and stock markets: A survey and future directions from a big data perspective. *IEEE Transactions on Knowledge and Data Engineering, 30*(2), 381–399. doi:10.1109/TKDE.2017.2763144

Mazboudi, M., & Khalil, S. (2017). The attenuation effect of social media: Evidence from acquisitions by large firms. *Journal of Financial Stability, 28*(C), 115-124. doi:10.1016/j.jfs.2016.11.010

Oh, C., & Sheng, O. (2011). Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement. *Proceedings of the International Conference on Information Systems*, 1-18. https://aisel.aisnet.org/icis2011/proceedings/knowledge/17

Oliveira, N., Cortez, P., & Areal, N. (2013). On the predictability of stock market behavior using stocktwits sentiment and posting volume. *In Progress in artificial intelligence. In Lecture notes in computer science, 8154*, 355–365. doi:10.1007/978-3-642-40669-0_31

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: using twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications, 73*, 125-144. doi:10.1016/j.eswa.2016.12.036

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion, 37*, 98–125. doi:10.1016/j.inffus.2017.02.003

Prokofieva, M. (2015). Twitter-based dissemination of corporate disclosure and the intervening effects of firms' visibility: Evidence from Australian-listed companies. *Journal of Information Systems, 29*(2), 107-136. doi:10.2308/isys-50994

Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance, 39*(4), 1127-1139. doi:10.1111/j.1540-6261.1984.tb03897.x

Saleemi, J. (2020). An estimation of cost-based market liquidity from daily high, low and close prices. *Finance, Markets and Valuation, 6*(2), 1-11. doi:10.46503/VUTL1758

Saleemi, J. (2021). COVID-19 uncertainty and Bitcoin market, linking the liquidity cost to the cryptocurrency yields. *Finance, Markets and Valuation, 7*(1), 1-11. doi:10.46503/BJWT6248

Saleemi, J. (2022). Asymmetric information modelling in the realized spread: A new simple estimation of the informed realized spread. *Finance, Markets and Valuation, 8*(1), 1–12. doi:10.46503/JQYH3943

Sarr, A., & Lybek, T. (2002). Measuring liquidity in financial markets. *International Monetary Fund, 2*, 1–64. doi:10.5089/9781451875577.001

Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2013). Predictive sentiment analysis of Tweets: a stock market application. *In Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*, 77-88. doi:10.1007/978-3-642-39146-0_8

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: the information content of stock microblogs. *European Financial Management, 20*(5), 926-957. doi:10.1111/j.1468-036X.2013.12007.x

Wei, C., Shihao, L., & Naqiong, T. (2014). The Influence of Investor Attention on the Stock Return and Risk: An Empirical Study Based on the "Easy Interactive" Platform Data of Shenzhen Stock Exchange. *Securities Market Herald, 7*, 40-47.

Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems, 55*(4), 919-926. doi:10.1016/j.dss.2012.12.028

Zhang, H., Chen, Y., Rong, W., & Wang, J. (2022). Effect of social media rumors on stock market volatility: A case of data mining in China. *Frontiers in Physics, 10*, 987799. doi:10.3389/fphy.2022.987799

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter ‘’i hope it is not as bad as I fear’’. *Procedia-Social and Behavioral Sciences, 26*, 55-62. doi:10.1016/j.sbspro.2011.10.562