Journal of
Risk and Financial Management

Article

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

Francisco Guijarro [1,*], Ismael Moya-Clemente [2] and Jawad Saleemi [3]

[1] Research Institute for Pure and Applied Mathematics, Universitat Politècnica de València, 46022 València, Spain
[2] Faculty of Business Administration and Management, Universitat Politècnica de València, 46022 València, Spain; imoya@esp.upv.es
[3] Business School, Universitat Politècnica de València, 46022 València, Spain; j.saleemi@yahoo.com
* Correspondence: fraguima@upvnet.upv.es; Tel.: +34-963877000

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

Citation: Guijarro, Francisco, Ismael Moya-Clemente, and Jawad Saleemi. 2021. Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data. *Journal of Risk and Financial Management* 14: 394. https://doi.org/10.3390/jrfm14090394

Academic Editor: George Halkos

Received: 10 July 2021
Accepted: 20 August 2021
Published: 24 August 2021

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 (Smailović 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):

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

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):

$$S_{t-1} = \frac{high_{t-1} - low_{t-1}}{close_{t-1}} \tag{2}$$

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):

$$v_t^s = \frac{v_t^{ask} - v_t^{bid}}{close_t} \tag{3}$$

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

$$\eta_t = (high_t + low_t)/2 \tag{4}$$

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:

$$v_t^{ask} = ask_t\pi + \eta_t\pi \tag{5}$$

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

$$v_t^{bid} = bid_t\pi + \eta_t\pi \tag{6}$$

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):

$$MLI_t = \sum\nolimits_{t=1}^{T} \frac{(close_t - close_{t-1})^2}{ln(vol_t)} \tag{7}$$

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):

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

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

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

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:

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

Equation (14) is further elaborated as:

$$LS_t = \begin{bmatrix} L_t \\ N_t \\ P_t \end{bmatrix}, \quad A = \begin{bmatrix} \alpha_L \\ \alpha_N \\ \alpha_P \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{bmatrix}, \quad L_t = \begin{bmatrix} L_{t-1} \\ L_{t-2} \end{bmatrix},$$
$$\gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_{31} & \gamma_{32} \end{bmatrix}, \quad N_t = \begin{bmatrix} N_{t-1} \\ N_{t-2} \end{bmatrix}, \quad \varnothing = \begin{bmatrix} \varnothing_{11} & \varnothing_{12} \\ \varnothing_{21} & \varnothing_{22} \\ \varnothing_{31} & \varnothing_{32} \end{bmatrix}, \quad (15)$$
$$P_t = \begin{bmatrix} P_{t-1} \\ P_{t-2} \end{bmatrix}, \quad \epsilon_t = \begin{bmatrix} \epsilon_{L,t} \\ \epsilon_{N,t} \\ \epsilon_{P,t} \end{bmatrix}$$

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

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

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.

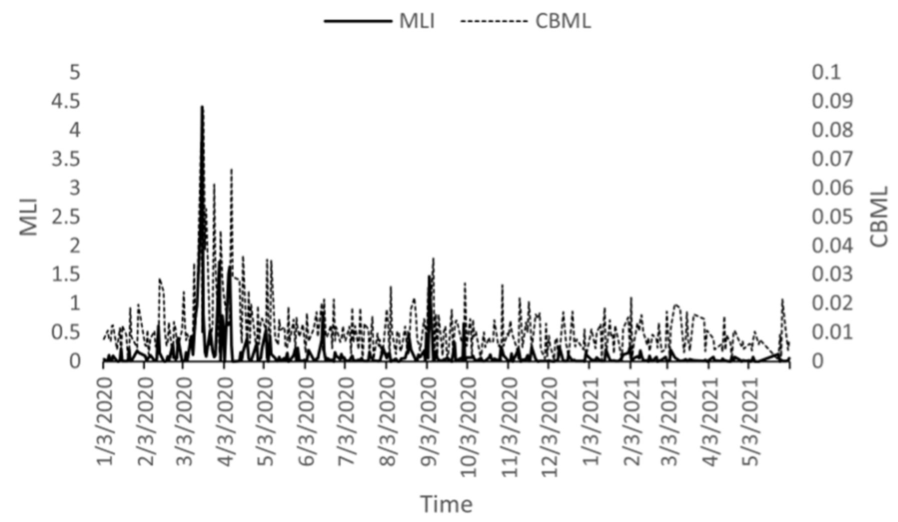

Figure 1. Time-varying market liquidity dimensions.

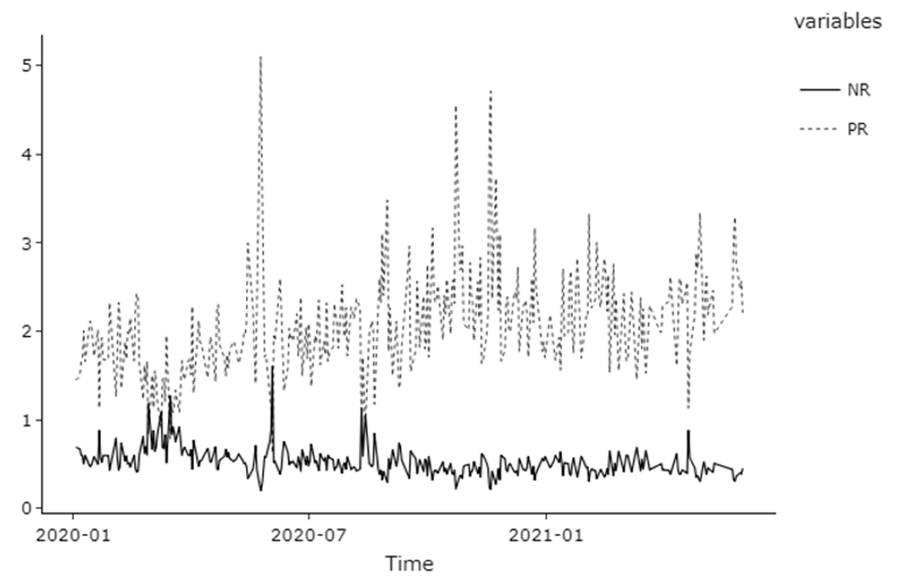

Figure 2. Time-varying investor sentiment indicators.

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
	Intercept	−0.46598	0.062395
MLI (b)	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
	$\varnothing_{11,P}$	−0.002544	0.1453
	$\beta_{12,CBML}$	−0.3748	0.000 ***
	$\gamma_{12,N}$	0.006916	0.2730
	$\varnothing_{12,P}$	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 ***
	$\varnothing_{21,P}$	−0.030498	0.25019
	$\beta_{22,CBML}$	−0.530606	0.49794
	$\gamma_{22,N}$	−0.279544	0.00375 **
	$\varnothing_{22,P}$	−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
	$\varnothing_{31,P}$	−0.463831	0.000 ***
	$\beta_{32,CBML}$	0.810466	0.7752
	$\gamma_{32,N}$	−0.054357	0.8757
	$\varnothing_{32,P}$	−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.

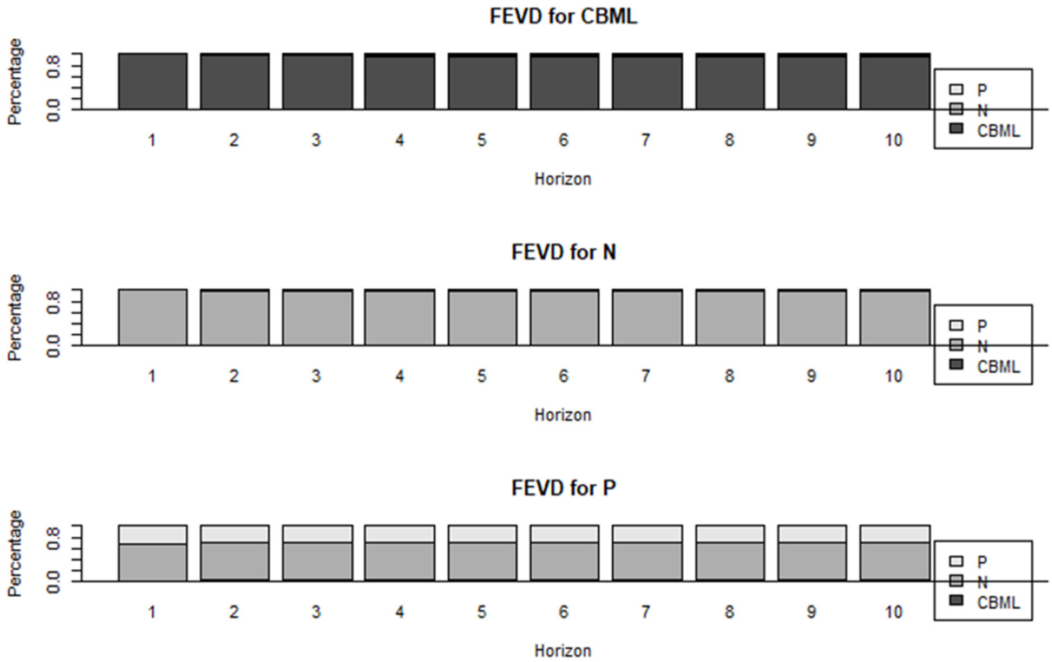

Figure 3. FEVD analysis for CBML and investor sentiments.

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 *
	$\varnothing_{11,P}$	0.0901849	0.1156
	$\beta_{12,\mathrm{MLI}}$	−0.2531608	0.000 ***
	$\gamma_{12,N}$	0.0268085	0.8976
	$\varnothing_{12,P}$	0.0373390	0.5145
	α_{MLI}	−0.0001461	0.9936
N (b)	$\beta_{21,\mathrm{MLI}}$	−0.013135	0.60219
	$\gamma_{21,N}$	−0.668877	0.000 ***
	$\varnothing_{21,P}$	−0.029397	0.26646
	$\beta_{22,\mathrm{MLI}}$	0.016084	0.51670
	$\gamma_{22,N}$	−0.276186	0.00437 **
	$\varnothing_{22,P}$	−0.020703	0.43411
	α_{N}	−0.001120	0.89491
P (c)	$\beta_{31,\mathrm{MLI}}$	0.065824	0.4712
	$\gamma_{31,N}$	0.099416	0.7738
	$\varnothing_{31,P}$	−0.474563	0.000 ***
	$\beta_{32,\mathrm{MLI}}$	−0.079112	0.3790
	$\gamma_{32,N}$	−0.047393	0.8920
	$\varnothing_{32,P}$	−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.

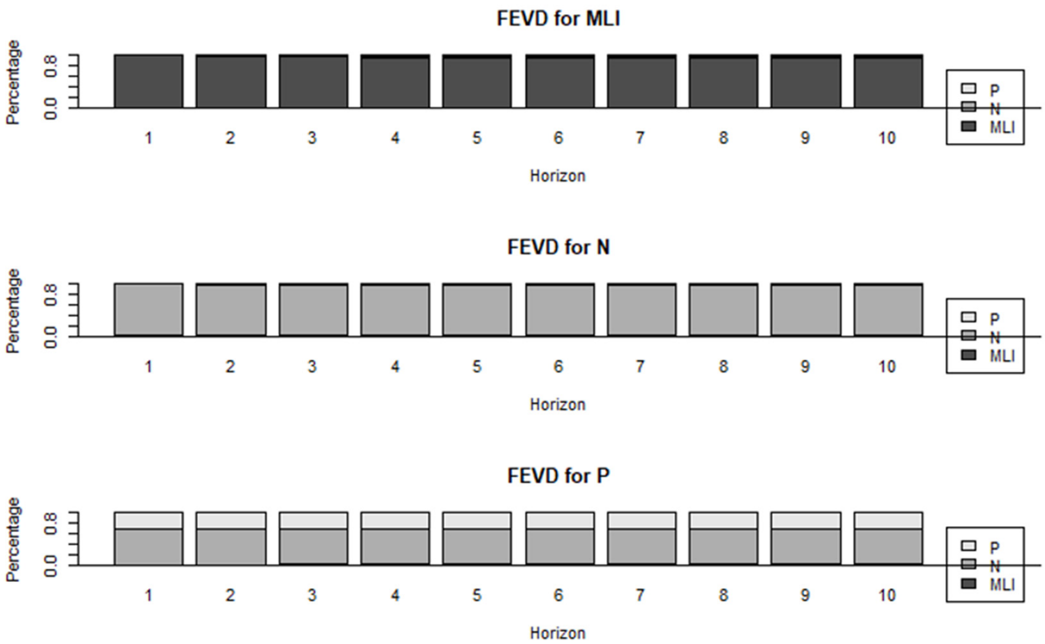

Figure 4. FEVD analysis for MLI and investor sentiments.

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.

Author Contributions: Conceptualization, F.G., I.M.-C. and J.S.; methodology, F.G., I.M.-C. and J.S.; software, F.G. and J.S.; validation, J.S.; formal analysis, I.M.-C.; investigation, F.G.; resources, J.S.; data curation, J.S.; writing—original draft preparation, J.S.; writing—review and editing, F.G. and I.M.-C.; visualization, J.S.; supervision, F.G. and I.M.-C.; project administration, F.G., I.M.-C. and J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: Two anonymous reviewers provided thoughtful and detailed comments that greatly improved the final version of this article. We would like to thank the reviewers for their valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

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]

Smailović, Jasmina, Miha Grčar, Nada Lavrač, and Martin Žnidaršič. 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]