Credit and Capital Markets, Volume 54, Issue 4, pp. 589 – 639
Scientific Papers

The Impact of Media Attention on the Illiquidity of Stocks: Evidence from the Global FinTech Sector

Eduard Gaar*, Valentin Moritz* and Dirk Schiereck*

Abstract

As a result of technological innovations in data processing, the exploitation of Internet usage data in relation to search engines or social networks is becoming increasingly intriguing for understanding and anticipating stock market movements. We analyze the impact of three alternative investor attention variables, i. e. Google search volume, Wikipedia page views, and stock market-relevant news on the rapidly growing FinTech sector. The result of the simultaneous correlation analysis reveals a highly significant correlation between the trading activities of the FinTech sector and the three investor attention variables. The time-delayed regression analysis complements the results by identifying substantial changes of the effects within one week considering the order of magnitude and sign. Furthermore, multivariate regression analysis highlights that the explanatory power for future stock trading activities and illiquidity primarily depends on Google search volume and stock market-relevant news volume, while the simultaneous correlations are best explained by the number of visits to the corresponding Wikipedia page.

Keywords: Capital Market Liquidity, Illiquidity, Internet Search, Media Attention, FinTech

JEL Classification: D82, D83, G12, G14, L82, L86

I. Introduction

In 1955, Nobel laureate Herbert A. Simon stated that the first step in the decision-making process of a market participant consists of the initial gathering of information (*Simon*, 1955, p. 106). Since then, the introduction of new technologies drastically changed the informational landscape providing access to a nearly unlimited amount of information through the Internet. The growing potential of the Internet as an unlimited source of analyzable data inherently offers the ability to accurately explain various phenomena and to predict behavioral patterns in numerous spheres of interest on a large scale. The related term Big Data Analytics relies on various sources of data including search engines, social

* Technical University of Darmstadt, Department of Business Administration, Economics and Law, 64289 Darmstadt, Germany.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

networks, discussion forums and news, which can be used to calculate simulation models with intangible value for both business and scientific research. One particular subject of attention concerns the application of Big Data Analytics on the financial market, so for example the Google search engine which provides the possibility to capture the interest of individuals, as a function of the number of search queries. First results in 2011 for the US market identified a significant correlation between stock price development and trading volume and the search volume of the Internet search engine Google (*Da/Engelberg/Gao*, 2011). Since then, there are some more studies which show an influence of Googles search volume on market indicators. *Pöppe* et al. (2014) focus on the global agricultural sector and *Bank* et al. (2011) on the German stock market and underline the influence of search volume on the stock market liquidity and return. Next to Google are other important sources for information about companies as the Wikipedia website and general media news and blogs. Wikipedia is the 6th most visited website overall in 2015 (*Hinnosaar* et al., 2015) and therefore a very popular starting point for initial information gatherings. LexisNexis, and especially Newstex, as a service for company profiles, legal information and other news like market insights offers a huge database for real-time information. Newstex as a collection of magazines, newspapers and blogs seems to be an important source for content for investors. As all these sources catch the attention of investors, they can affect the price formation on security markets.

At this point, it is important to differentiate between different companies and thus internet search habits. *Pöppe* et al. (2019) point out that while for business-to-business companies the majority of search queries probably come from interested investors, for consumer-focused companies the majority of searches are product-related and thus probably from potential buyers. In addition, these search queries occur in an international setting and concern companies that operate worldwide. As these searches should not have a direct impact on liquidity, it is interesting to see how an international, fast-growing and heterogeneous industry like the FinTech industry, which in our data sample consists mainly of B2B companies, is influenced by these searches. The relatively young financial technology sector (FinTech), which "… is used to describe a variety of innovative business models and emerging technologies that have the potential to transform the financial services industry …" (Report of the Board of IOSCO, 2017, p. 4), presents a particularly interesting market. Furthermore, it is in the nature of the FinTech sector that it has a very high monetary value as well as a high degree of adaptability and flexibility that is needed to exploit synergies in both the financial sector and the technology industry (*Smith/Tran/Perera Tino*, 2016, p. 3; *Hendrikse/Bassens/van Meeteren*, 2018, p. 162). This paper provides empirical evidence on the extent to which the attention of stock market participants positively or negatively effects the stock's trading activities for an international data set of FinTechs. For this purpose, the paper examines the development of

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

three proxies for investor attention (Google Search, Wikipedia Search and news volume) that have been confirmed in literature separately and for different industries in the past and their implications for trading activity and liquidity of the associated companies over a period of 5 years. In contrast to *Da* et al. (2011) and *Bank* et al. (2011) who each focus only on one country and on industries that have the highest average attention in their respective countries we extend the focus. In addition to comparing three different investor attention proxies, we analyze stocks with a lower basic attention and try to apply already known illiquidity patterns to a new industry. This lower basic attention should rule out the possibility that non-investor-relevant search queries, such as for gifts in the run-up to Christmas, lead to random significances. Further, the international setting of FinTechs limits country-specific search queries that affect all stocks of a market index.

Further, we contrast with literature such as *Duan* et al. (2020), *Li* and *Yu* (2012), *Smales* (2021) or *Vozlyublennaia* (2014), whose focus is on analyzing the impact of investor attention on returns and predictability. Since the literature in this area is already extensive, we look at the impact on illiquidity and volume in an attempt to add value. In this regard, the emphasis lies in the comparison between the three different proxies and their capability to explain concurrent and time-delayed developments in trading activity. The results presented below identify the Google search volume as a significant measure for the subsequent stock's illiquidity and therefore corroborates previous findings from *Bijl* et al. (2016). Beyond that, the results also indicate a significant change in the relationship between the proxies for investor attention and the key figures for liquidity over the course of one week from a positive to a negative correlation.

The rest of this paper is structured as follows. First, chapter 2 gives an overview of previous research in this area. After that chapter 3 introduces the proxies for investor attention and the key indicators for trading activity and liquidity. Chapter 4 presents the results of a correlation analysis of liquidity and search activity before chapter 5 continues with a multivariate regression analysis. Chapter 6 finishes with a conclusion.

II. Previous Research

Early research studies indicate proxies for investor attention such as the number of traded shares and extreme returns (*Barber/Odean*, 2008) or quantifies investor attention as media coverage in the form of news headlines (*Tetlock*, 2007). These two papers show that the media coverage referring to a company, like the number of mentions in the Wall Street Journal, increases trading volume, high stock returns and results in a significantly increased demand from investors. The results are associated with the assumption that considerable changes in

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

market activities like higher trading volume or an unusually high number of references in newspapers do grab the attention of investors. *Da* et al. (2011) criticize this hypothesis by referring to reduced attention from investors due to the enormous amount of available information. They propose the Google search volume as a direct measure for investor attention and find evidence that the Google search volume correlates with the US stock prices differently from other proxies for investor attention (*Da* et al., 2011). Further research studies the relation of the Google search volume and stock market movements of specific countries and sectors. *Bank* et al. (2011) conclude that the Google search volume provides a convenient proxy for company recognition and investor attention by analyzing German companies traded on the Xetra trading system. More specifically, the Google search volume presents a positively correlated proxy for companies' trading volume and a negatively correlated proxy for companies' illiquidity, which functions as a metric for the price impact. Furthermore, based on those findings concerning investor attention, they provide the argument that the naïve Google search volume does not represent the overall investor attention, but primarily the attention of uninformed individual investors and thereby agree with the findings of *Da* et al. (2011).

The dynamic of individual investors is further examined by *Dimpfl* and *Jank* (2016). The analysis of the relation between stock market volatility and investor attention based on the Dow Jones Industrial Average from 2006 to 2011 introduces evidence for noise trading as an explanation for the immediate stock market response. Based on the hypothesis of noise traders, investors conduct trades without actual knowledge of new informational value regarding the company or without fully taking advantage of the information available. Similar to the noise trader model, *Redding* (1996) argues that herding behavior can be used to explain a portion of the deviation of stock prices. In this case, the investor consciously or unconsciously follows the masses instead of relying on his own information. The research of *Pöppe* et al. (2014) continues the work of *Bank* et al. (2011) by examining a specific industry on a global scale. The investigation of the global agricultural sector, as a B2B sector with a generally low degree of public attention, indicates a strong explanatory and predictive power between the investor attention proxy Google search volume and the negatively correlated liquidity of a company's stock. Regarding the search terms for the Google search data, both papers argue for the removal of company-specific legal terms like "LTD", "INC" or "LLC" if possible, to include more investor attention (*Bank* et al., 2011; *Pöppe* et al., 2014).

In contrast to research mainly using Google search queries, *Bordino* et al. (2012) offer insights into the relation between stock market movements and online search query volume based on data of the Yahoo search engine and the NASDAQ 100. In addition to using a different online search tool, the companies' unique ticker symbols are used. The results of the different settings comply

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

with previous findings of a correlation between online search volume and trading volume *Bordino* et al. (2012). A more recent paper by *Bijl* et al. (2016) studies the influence of Google search volume on stock returns of companies listed on the S&P 500 in the period between 2008 to 2013. The results generally agree with earlier findings concerning the relationship between the Google search volume and stock market return predictions. Further, they show that an increase in investor attention precedes a decrease in stock returns, whereas previous research identified an initial increase of stock returns with a subsequent decrease in the following two weeks (*Bijl* et al., 2016).

Other sources of information for investors are social media platforms like Twitter or Facebook. Amplified by mobile social media applications, they provide a stage to connect and circulate information with a considerable reach and speed. Consequently, there seems to be a certain predictive power to social media for real-world performances. *Asur* and *Huberman* (2010) show a strong correlation between Twitter post frequency and sentiment and box office revenues of movies that outperform other revenue prediction models. *Bollen* et al. (2011) examine the public mood resulting from the extracted sentiments from Twitter posts and indicate that the use of Twitter feed as a public mood indicator increases the predictive power of changes in the closing price of the Dow Jones Industrial Average Index. In order to capture the public mood concerning specific companies, *Smales* (2012) examines news articles from the Dow Jones newswire and the Wall Street Journal. Moreover, by extracting sentiments, the degree of relevance and novelty, the paper extends the research of *Tetlock* (2007) in an attempt to account for noise alongside the actual news. In agreement with *Groß-Klußmann* and *Hautsch* (2011), both papers provide evidence for a significant and positive influence of news sentiment and news volume on the volatility and the trading volume of specific companies. In addition to these findings, *Smales* (2012, 2014) also shows that stock market reactions are more strongly influenced by negative news.

A different approach to investor attention is chosen in the work of *Rubin* and *Rubin* (2010) by focusing on information processing. The paper introduces the editing-frequency of Wikipedia pages as an indicator of the public interest in a company. Wikipedia as a web-based open-source encyclopedia presents a tool for individuals to express their interest in specific topics through reading the associated articles as well as creating or editing Wikipedia articles. “The unique attribute of Wikipedia that allows individuals to actively participate in the information gathering process allows us to quantify cross-sectional variation in internet information processing” (*Rubin*/*Rubin*, 2010). Their research examining the link between the Wikipedia editing frequency and the number of informed investors and analysts confirms their theory of increased information processing’s correlation with the degree to which market participants are informed about companies. The level of information shows a reduced deviation between ana-

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

lysts' forecasts and actual stock market reactions, which are quantified through increases in bid-ask spreads in this research (*Rubin/Rubin*, 2010). Furthermore, *Xu* and *Zhang* (2013) show that Wikipedia as a tool for information aggregation functions as a measure to alleviate some of the information asymmetries for investors. They argue that the different types of information aggregation, as opposed to traditional ways for investors to get informed, inhibits biases to a certain degree.

Further research from *Antweiler* and *Frank* (2004) investigate the information content of stock message boards. The paper analyzes over 1.5 million messages posted on the Yahoo Finance and Raging Bull message board concerning 45 companies listed on the Dow Jones Industrial Average and the Dow Jones Internet Index. Under the premises of the financial theory that disagreement induces trade, the frequency of messaging and the degree of bullishness show a significant, but compared to transaction costs economically small, negative influence on stock market returns and volatility. The results for daily trading volumes support the financial theory depending on the degree of disagreement. Up to a certain level of disagreement on the message board, trading volume is positively impacted the next day. This finding is especially relevant for smaller-sized trades. However, the paper provides evidence that as soon as the detected degree of disagreement passes a certain threshold, the influence of next day trading volumes reverses to a negative correlation (*Antweiler/Frank*, 2004). Extending the research on stock market message boards, *Sabherwal* et al. (2011) analyze herding behavior in relation to stock board messages in the absence of actual material news. The product of their research supports earlier findings that disagreement is a factor for trading activity. More precisely, the abnormal returns correlate positively one day ahead of the increased disagreement, followed by a negative correlation in the subsequent two days. The effect is particularly noticeable for stocks with small market capitalization, weak financials and low institutional holdings (*Sabherwal* et al., 2011). Both mentioned papers attribute a high influential power over stock market movements to message boards and propose further research on that subject for example in topics like artificially created conversations and disagreements to profit from short-term herding behavior.

The extensive research on proxies for investor attention highlights Google search volume, news volume, Wikipedia article editing frequency, and message board posts as four valid approaches to detect the influence of investor attention on stock market activity. With the previous research in mind, this paper extends the work of *Pöppe* et al. (2014) and *Bank* et al. (2011) in three essential dimensions. First, the selected data set consists exclusively of globally composed companies operating in the FinTech sector and thereby includes B2B as well as B2C firms with characteristics of both the financial and technology sector (*Pöppe* et al., 2014). Second, the analyzed time period from 2015–2019 offers insights

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

into whether the predictive power of investor attention proxies during ordinary market conditions is similar to the financial crisis situation. Moreover, *Bijl* et al. (2016) argue that the incorporation of investor attention into the market occurs more rapidly than one decade ago and therefore indicates that there is a need for new research with recent data. And third, the introduction of Wikipedia-article visits and FinTech market-related news as additional investor attention proxies beyond the Google search volume allows a sector-specific analysis of the effects of the proxies on trading activities and of the relationship among the proxies.

III. Method

1. Data Sources

a) Google

The introduction of the Internet to the public at the beginning of the 1990s provoked the emergence of myriads of websites to such an extent that simple indexing was no longer a viable option for cataloging all websites. Subsequently, commercial search engines offered a solution by sorting the websites according to their relative frequency of the searched keyword. Based on a Ph.D. research project from Stanford University regarding search engine optimization in 1998, Google was founded. The main feature of Google's new search engine in contrast to the conventional keyword-based algorithms was the introduction of the relevance of each site in the form of the number and importance of other sites that are linked to it (*Moore/Tambini*, 2018).

The established dominance in the search engine market and a multitude of new ventures in the following years, led to Google's initial public offering in 2004, which generated a total market capitalization of 23 billion US-dollars. Since then, Google has capitalized on the increasing importance of the Internet. In 2012, approximately a third of the world's population, around 2.4 billion people, adopted the internet as a means of communication and information procurement. The search engine market share in the same year amounted to approximately 91 % which translates into 2.1 billion users. In 2017, Google's market dominance remained with around 88 % of the desktop market shares and 95 % of the mobile search market (*Statista*, 2019; *Moore/Tambini*, 2018). The undisputable leading position of Google in the search engine market requires Google to function as the dataset source and omits the necessity to include other search engines in the study.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

b) Wikipedia

The open-access encyclopedia Wikipedia is characterized by a high diversity, a large content and the collaborative approach to produce content. Founded in 2001, it contains more than 5 million articles and about 400 million cross-references through links. The website receives a considerable amount of web-traffic and can be viewed as a standard reference source as it is the 6th most visited website overall in 2015 (*Hinnosaar* et al., 2015). In order to maintain reliable and published sources for the Wikipedia articles, the guidelines dictate that all majority and significant minority views have to be covered by sources and original research. Furthermore, Wikipedia provides discussion platforms for each article with the objective to encourage users to exchange their opinions concerning the articles validity. The concept of no gatekeeping function, no proof of identification or qualification facilitates the usage of the open editing system. Moreover, Wikipedia stores every version of an article entry, which allows Wikipedia to discard any kind of falsified articles and revert to the saved previous version (*Rubin/Rubin*, 2010).

c) LexisNexis

As one of the subsidiaries of the RELX Group, formerly Reed Elsevier, LexisNexis specializes in the fields of economics and law. Their news and their legal database contain around 109 billion documents. Its online research platform offers a professional search database for international sources and information in full text. More specifically, Nexis news and business content include 40,000 sources in 30 languages covering more than 150 countries and it is composed of several thousands of journals, magazines, and newspapers from all around the world with their respective archives included. In addition, their offer covers further areas, such as legal information, biographies, market insights and company profiles (*Poley/Kuffer*, 2020; *LexisNexis Legal/Professional*, 2020). In 2004, LexisNexis established Newstex as a service with the objective to provide "[…] real-time news and commentary from thousands of branded newswires, newspapers, magazines, financial and business sources, official government feeds and weblogs" (*Newman*, 2008). Moreover, the news aggregation service Newstex has been rewarded for its innovativeness concerning its real-time information feed with multiple digital media and content awards, which gives them the status of a key player in the digital content industry. Therefore, the platform offers the potential to provide extended access to news on specific topics.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

2. FinTech Sector

In literature the notion "FinTech" is referred to as a portmanteau word as it is a combination of the words "Financial Service" and "Technology". Accordingly, companies is this sector aim at utilizing innovative information technology to provide specific financial services to customers and or businesses (*Alt/Puschmann*, 2016; *Gulden*, 2019; *Tiberius/Rasche*, 2017). The emergence of FinTechs and more importantly their substantial growth in investments are timed around the global financial crisis of 2008. Associated with the crisis, a rising sentiment of distrust in the banking industry was compounded by a growing customer base of digital natives (*Wang* et al., 2013), who were receptive to mobile services and functioned as a catalyst for financial innovation (*Lee/Shin*, 2018; *Menat*, 2016; *Tiberius/Rasche*, 2017). In addition to these initial facilitators for growth, the application of disruptive technological innovations on personalized services and specifically banking areas allows FinTechs to disintermediate traditional finance companies (*Lee/Shin*, 2018). The report of the International Trade Association predicts that "over $4.7 trillion of revenue at traditional financial services companies is at risk of disruption by the new FinTech entrants" (*Smith* et al., 2016). In 2014 the growth in investment volume of FinTech companies outperformed global venture capital investments with 201 % to 63 % according to *Skan* et al. (2015) and the investment volume further increases as FinTech investments more than doubled from 2017 to 2018.[1] Estimates show that the compound annual growth rate (CAGR) of Fintech companies is expected to increase by 20 – 30 % until 2025 (QYResearch, 2019).

The nature of FinTech companies as a combination of the Financial Services and Technology sector raises pertinent questions regarding its regulatory environment. Especially the Financial sector as a response to the financial crisis of 2008 became subject to increased regulatory scrutiny (*Heath* et al., 2015). FinTech companies, also operating on the financial service market, are subject to the increased regulatory standards (*Smolinski* et al., 2017). However, the novelty of FinTechs and the different approaches to financial services result in different regulatory approaches dependent on the country and on the FinTechs sub-sector (*Loesch*, 2018). The pace of innovation in the industry is much higher compared with the frequency of regulatory changes and consequently, governments face the issue of continuously redrafting outdated regulations (*Lee/Shin*, 2018).

The companies included in the survey and thus constituting the data set have to show the following characteristics:

- The company has to be publicly traded.

[1] From $ 50.8 Billion US Dollar in 2017 to $ 111.8 Billion US Dollar in 2018 (*Blackman,* 2019).

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

- The company utilizes technology to create financial products and provides financial services.
- The distribution of the product or service is carried out exclusively in electronic form.
- The companies do not or only to a limited extent maintain physical locations.
- The companies expected revenue mixes are mainly based on fees.
- The company is not classified as a Penny Stock to avoid distortion due to strong fluctuations.

The introduced criteria apply to the 48 companies of the KBW Nasdaq Financial Technology Index (KFTX) (see Table A.1 in the Appendix for the full index). Of these 48 companies, 33 are pure B2B FinTechs, 5 operate only as B2C companies, and 10 companies offer both business relationships. This shows that the FinTechs studied are mainly B2B and that distinguishing and analyzing B2C and B2B FinTechs would not add any value. The companies without available data at the beginning of the observation period in July 2014 are included after the completion of their initial public offering. The resulting data set is therefore a panel data set and thus consists of cross-sectional and time series data with 189956 existing data points that satisfy the defined criteria. One feature of panel data is the ability to deal with unmeasured explanatory variables that have an influence on the behavior of the examined firms and potentially cause biases in estimations, called heterogeneity. Furthermore, the combination of variations across the different individuals with variation over time increases the efficiency of the estimations, particularly due to the mitigation of issues with multicollinearity. Another crucial feature of panel data is the superior analysis of dynamic adjustments, because capitalizing on the dynamic reactions of every individual separately compensates for an extremely long time series (*Kennedy*, 2009).

3. The Selection and Inquiry of Financial Key Figures and Internet Search Activity

The relevant economic indicators of the individual companies are retrieved from the Thomson Reuters Datastream service. We use the weekly data to be able to compare the weekly economic indicators with the weekly Google search volumes without further conversion. In reference to previous research from *Bank* et al. (2011) and *Pöppe* et al. (2014) on the relationship between online investor attention and stock market reactions, the following key figures serve as proxies:

- Traded Volume (TV)
- Turnover Rate (TR)

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

- Return of a Stock (R)
- Illiquidity (ILLIQ)
- Turnover Price Impact (TPI)
- Google search volume (GSV)
- Wikipedia search volume (Wiki)
- News volume (News)

The Traded Volume, as the first measure for trading activity, in thousand US Dollars (USD) TW_{iyw} of the company's stock i in the year y and the week w is calculated by the natural logarithm of the weekly traded shares VO_{iyw} multiplied by the weekly closing price P_{iyw}:

(1) $$TV_{iyw} = \ln\left(VO_{iyw} * P_{iyw}\right)$$

Moreover, in previous research the Turnover Rate TW_{iyw} of stock i in year y and week w has been identified as an intuitive proxy for trading activity (*Lo/Wang*, 2000). It represents the fraction of traded shares VO_{iyw} relative to the number of outstanding shares $NOSH_{iyw}$ of stock i in year y and week w. More precisely, an elevated Turnover Rate indicates that the average holding period of the stock reduces because the Turnover Rate can be viewed as the reciprocal value of the average holding period (*Bank* et al., 2011):

(2) $$TR = \frac{VO_{iyw}}{NOSH_{iyw}}$$

The measure to quantify how fast and to what extent a share can be bought and sold on the market is defined as the liquidity of a share. In related research of Amihud (2002), the measure illiquidity is introduced and represents the ratio of the return of the share to the volume of shares traded during the same period. The illiquidity or Amihud-ratio $ILLIQ_{iyw}$ functions as a proxy for the price impact, because it mirrors the price changes caused by a particular quantity of trading volume. Consequently, a high trading volume which does not result in a price change suggests a market that is capable of compensating the high trading volume and therefore the illiquidity is low. The absolute value of the return of the share R_{iyw} divided by the previously calculated Traded Volume TW_{iyw} of share i in year y and in week w produces the Illiquidity $ILLIQ_{iyw}$:

(3) $$ILLIQ_{iyw} = \frac{\left|R_{iyw}\right|}{TV_{iyw}}$$

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

(4) $$R_{iyw} = \frac{P_{iyw} - P_{iy(w-1)}}{P_{iy(w-1)}} * 100$$

The Return of the share $R_{i,y,w}$ of the company *i* in year *y* and week *w* the is the difference between the price of the share in week *w* $P_{i,y,w}$ and week $(w-1)$ $P_{i,y,w-1}$ divided by the price of the share in week $(w-1)$ $P_{i,y,w-1}$ (*Bank* et al., 2011; *Pöppe* et al., 2014).

An alternative proxy for the price impact, which is based on *Amihud* (2002), is the Turnover Price Impact TPI_{iyw} for share *i* in year *y* and week *w*. Instead of calculating the quotient on the basis of the absolute value of the return of the share and the traded volume, the Turnover Price Impact divides $R_{i,y,w}$ by the Turnover Rate $TR_{i,y,w}$:

(5) $$TPI_{iyw} = \frac{\left|R_{i,y,w}\right|}{TR_{i,y,w}}$$

The change to the $ILLIQ_{iyw}$ reduces the impact of market capitalization on the TPI_{iyw} and thereby might be more adequate for an extended observation period of approximately five years (*Pöppe* et al., 2014).

(6) $$GSV_{i,t} = \frac{GSV_{i,t,a} - \overline{GSV_i}}{ó_{GSV_i}}$$

Internet search activity represents the number of queries for a certain search term. Among the numerous search engines available on the internet, Google positions at the top with the highest market share by far. In comparison to its competitors, Google also allows users to enquire about search volumes with their service Google Trends. The platform returns a representative sample of Google search activity concerning a certain category or keyword.[2] In addition to the sampling, Google Trend normalizes the search data with the objective of improving comparability between different search terms and returns the demand on a scale of 0 to 100. More specifically, "each data point is divided by the total searches of the geography and time range it represents to compare relative popularity" (Google, 2020). The observation period determines the granularity of the data. For example, an observation period of more than five years returns monthly data, whereas a period of fewer than five years returns weekly data and a period of fewer than 90 days provides daily data. The adjustment to general

[2] Google processes billions of search queries per day and in order to provide fast processing of the query representative samples of the data are deployed.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

trends as a result of the increased popularity of the Internet also inhibits the comparison of variation between companies and therefore the examination is restricted to the variation in search volume within each company (*Bank* et al., 2011).

The search volume for each company is queried via the company name and its corresponding time series between July 2014 and October 2019 and returns the global search volume on a weekly basis. The search term corresponds to the company's name[3] omitting, the terms in their names describing the legal form like LTD, HLD or Inc. The usage of legal terms is restricted to cases with a potential distortion of the data due to ambiguous meanings (e.g. "Square" as geometrical shape and "Square Inc" the FinTech company). The examined companies operate on a global scale and therefore geographical and linguistic filtering could limit the significance of the data. A list of all company names or tickers used for the Google search volume is included in the Appendix. For each company, there is a series of at least eight consecutive weeks and the data set also includes very low search volumes like a relative demand of zero, provided the corresponding Stock market data is available for the period.

In order to be able to compare different observations we standardize the search volume. There are many possibilities to standardize Google search volume so we follow the research from *Pöppe* et al. (2014) and use the classical standardization. The mean value of the population $\overline{GSV_i}$ of the company i is subtracted from the raw value $GSV_{i,t}$ of company i in week t and next divided by the standard deviation of the population σ_{GSV_i}.

(7)
$$WIKI_{iyw} = \sum_{d=1}^{D_{ijw}} WIKI_{iywd}$$

The daily visits on the Wikipedia page $WIKI_{iyd}$ of the company i in year y and day d are being aggregated on a weekly basis in order to match the time intervals of the Google Search Volume GSV_{iyw}. The online statistic tool Pageviews provides daily or monthly information about the number of visits to a specific Wikipedia-Article. In addition, the tool allows the user to specify the query in order to reduce noise in the results. The importance of noise and distortion reduction applies to Wikipedia as agents like web crawler for example from Google or automated programs can be used to mine data or optimize the search engine. Consequentially the data potentially contains peaks that do not represent a heightened interest of individual users and results in a distorted set of data (*Forns*, 2020) and therefore, the filter only includes individual users. In addition, due to the increasing use of mobile devices, the data includes search queries

[3] All company names are provided by Thomson Reuters Datastream.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

from desktops and mobile phones, which accounts for mobile web and mobile app queries.

$$NEWS_{yw} = \sum_{d=1}^{D_{ijw}} NEWS_{yd} \tag{8}$$

The daily amount of the search term related news $NEWS_{yd}$ of year y and day d are being aggregated on a weekly basis in order to match the time intervals of the Google Search Volume GSV_{iyw}. The student portal Nexisuni uses Lexis-Nexis's database regarding legal and economic issues. Contrary to the data from Wikipedia and Google Trends, Nexisuni collects all news related to the keyword FinTech and its variations[4] to avoid missing articles due to a different spelling of the search term. In addition to the search term optimization, the data exclusively contains English articles published by the news aggregation service Newstex. The decision to use Newstex as the only source for data is based on its aggregation of news and blogs for a more diversified data collection. After the application of these filters, NexisUni provides a data set containing news articles and blog posts related to the keyword FinTech between 2014 and 2019 on a daily basis. The other feature of NexisUni offers the possibility to perform sentiment analysis for English articles with the aim of identifying negative news articles. To classify the news data, Nexis iterates through a list of negatively connotated words in their truncated form and tries to find matches in the near vicinity of the keyword[5]. A match in the same paragraph of the search term, as well as the number of occurrences of negatively connotated words, decides whether it is a negative article or more specifically an article that discusses the subject FinTech in a negative manner. In order to further improve the detection rate, the original list of NexisUni can be extended and therefore additional negatively connotated words, specifically related to the financial or technical industry, were manually added.

IV. Correlation Analysis of Liquidity and Search Activity

The correlation analysis investigates the concurrent relationship between the proxies for investor attention and stock trading activities. For this purpose, starting with a visual examination of this relationship based on an exemplary FinTech company, the correlation will be examined based on absolute values and weekly changes. Subsequently, discrepancies in the correlation results are investigated by partitioning the data set to ensure a reasonable degree of robust-

[4] More specifically, keywords like Fintechs, Financial Technology, Financial Tech, etc.
[5] In this case Fintech or Financial Technology.

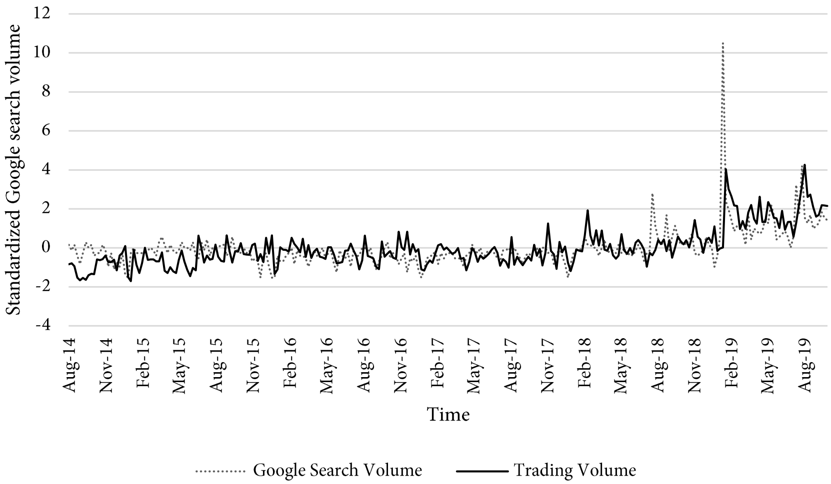
DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ness. In order to facilitate the comparability between the relationships, the data is standardized with a mean of 0 and a standard deviation of 1.

The following figures 1–3 showcase the relationship between the trading activity and the different proxies for investor attention for the FinTech company Fiserv Inc.

Figure 1 depicts the relationship between the two time-series of the Google search volume and the company trading volume over the entire observation period. The figure shows similar curves for both time-series, which suggests the notion that the relationship between the two time-series is not random. On closer inspection, the figure indicates a lagged response of the company's trading volume to the company's Google search volume, especially visible from September 2018 until the end of the observation period.

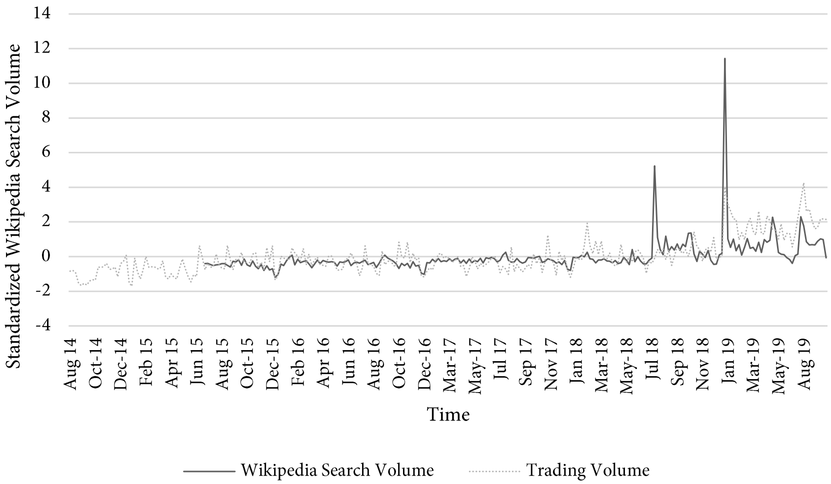
Figure 2 shows the trading volume of the same company and the volume of visits to the companies Wikipedia page. Similar to the comparison between Google search volume and trading volume, the curves in Figure 2 indicate a substantial connection between the Wikipedia search and trading volume. The comparison between Figure 1 and 2 illustrates that the signs for a lagged response of the trading volume to the Google search volume cannot be observed to the same degree in Figure 2.



Note(s): Company Fiserv Inc.

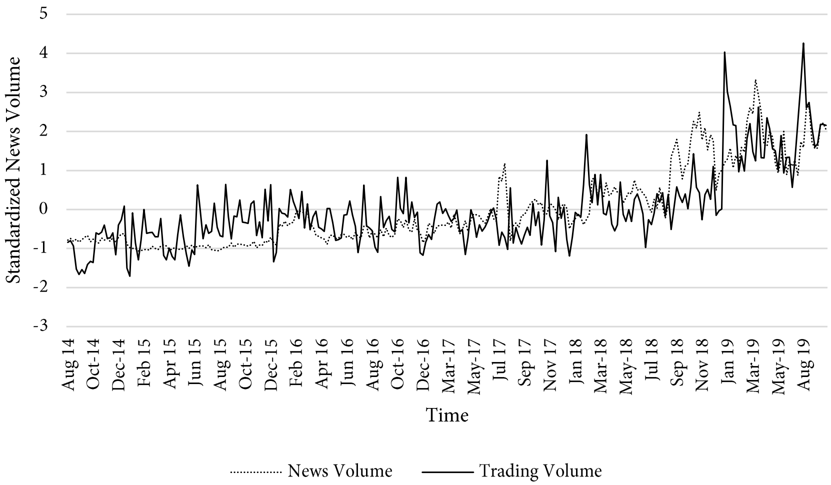
Figure 1: The Development of Standardized Google Search and Trading Volume

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/



Note(s): Company Fiserv Inc.

Figure 2: The Development of Standardized Wikipedia Search and Trading Volume



Note(s): Company Fiserv Inc.

Figure 3: The Development of Standardized News and Trading Volume

Lastly, the relationship between FinTech related news and the trading volume of the company is displayed in Figure 3. The closer examination of the two time-series indicates a link between the two variables as the two curves show a similar development.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Therefore, all three figures suggest that in the case of an elevated traded volume also the theoretical proxies for investor attention are elevated. Moreover, the similar course of the figures for the Google and Wikipedia search volume indicates a possible correlation between those variables.

In addition to the visual examination of the company's relationship between the trading volume and potential proxies for investor attention, the Pearson correlation coefficient matrix in Table 1 provides statistical measures for the relationships of the entire data set. The three proxies for investor attention Google search volume (*GSV*), Wikipedia search volume (*Wiki*) and FinTech related news volume (*News*) show a highly significant and positive correlation with the trading volume. More specifically, the FinTech related news volume demonstrates the highest correlation coefficient of 0.442 followed by the Wikipedia search volume with 0.288 and the Google search volume with 0.131. Apart from the high correlation coefficients for the trading volume, the results exhibit a significant positive, but considerably lower, correlation coefficient for the turnover rate. Furthermore, the variable illiquidity shows a significant negative correlation coefficient for the Google and Wikipedia search volume, whereas the results suggest an insignificant correlation between the news volume and illiquidity.

The relationship between the Google search volume, Wikipedia search volume and the FinTech-related news volume exhibits a highly significant positive correlation. On closer inspection, the coefficients indicate that the Google and Wikipedia search volume (correlation coefficient: 0.253), as well as the Wikipedia search and the news volume (0.281), react positively to each other to a much higher degree than the news volume and the Google search volume positively correlate (0.040). Consistently with the positive and significant correlation coefficients for the trading volume (GSV: 0.131, Wiki: 0.288, News: 0.442), the negative and partly significant correlation coefficients for the illiquidity

Table 1

Correlation Matrix of Absolute Values

Variables	*GSV*	*Wiki*	*News*	*TV*	*TR*	*ILLIQ*
GSV	1	0.253***	0.040***	0.131***	0.052***	–0.019**
Wiki		1	0.281***	0.288***	0.111***	–0.024**
News			1	0.442***	0.045***	–0.012
TV				1	0.650***	0.148***
TR					1	0.114***
ILLIQ						1

*p < 0.1; **p < 0.05; ***p < 0.01

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

(GSV: –0.019, Wiki: –0.024, News: 0.012) further highlights the validity of the results, as an increased trading volume is connected to reduced illiquidity of the company's shares.

In this context it is necessary to consider the panel data regarding stationarity. The augmented Dickey Fuller test based on *Choi* (2015) is used. This test is also suitable for unbalanced panel data sets due to its more general assumptions. The null hypothesis for this test, which assumes that all panels are unit-root nonstationary can be rejected for all considered variables in the correlation analysis and therefore indicates that at least one individual of the panel data set exhibits stationarity.[6] Furthermore, the absolute values as well as the differences are examined in order to provide more meaningful results.

Table 2 provides the correlation coefficients of the weekly changes in investor attention proxies and trading activity key figures. The correlation between the changes in trading activity and investor attention proxies is calculated from the difference between the week t and the respective previous week t–1.

The news volume changes (ΔNews) coincides with increasing trading activity ΔTV (0.010) and ΔTR (0,067), the correlation is significant at the 1 % level and the correlation coefficient for the change in illiquidity ΔILLIQ (–0.062) is negative at the 1 % level. These results align with the previous correlation analysis for absolute values. Regarding the change in Google search volume ΔGSV, Table2 provides evidence for a negative correlation between the ΔGSV and the trading activity, as the correlation coefficient for ΔTV (–0.076) and ΔTR (–0.079) is negative and highly significant. The negative coefficient for ΔILLIQ (–0.006) indicates no significant correlation, whereas a positive correlation coefficient can

Table 2

Correlation Matrix of Weekly Differences

Variables	*ΔGSV*	*ΔWiki*	*ΔNews*	*ΔTV*	*ΔTR*	*ΔILLIQ*	*ΔR*
ΔGSV	1	0.091***	–0.029***	–0.076***	-0.079***	–0.006	0.016**
ΔWiki		1	0.062***	0.139***	0.112***	0.020**	0.002
ΔNews			1	0.010***	0.067***	–0.062***	0.007
ΔTV				1	0.841***	–0.047***	0.034***
ΔTR					1	–0.045***	–0.070***
ΔILLIQ						1	0.133***
ΔR							1

*p<0.1; **p<0.05; ***p<0.01

[6] The complete test results are displayed conditions in Table A.3 in the Appendix.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

be observed between the change in return ΔR and ΔGSV at the 5 % level (0.016). This would suggest that an increase in Google search volume is associated with a decrease in trading activity and simultaneously entails an increase in return. In contrast to the Google search volume, the change in Wikipedia search volume ΔWiki correlates positively and significantly with changes in the trading activities ΔTV (0.139), ΔTR (0.067), and ΔILLIQ (0.020). Particularly noteworthy is the inconsistency between the increased trading activity and the simultaneously rising illiquidity. Furthermore, the change in news volume largely confirms the findings of Table 1 as it shows a significant positive correlation with ΔTV (0.010) and ΔTR (0,067) while indicating a negative relation with ΔILLIQ (–0.062) due to a highly significant negative coefficient.

The positive and significant correlation of the illiquidity with rising returns (0.133) is a well-known phenomenon, as investors demand a yield premium for lack of liquidity. Furthermore, the negative correlation at the 1 % level between ΔILLIQ and ΔTV (–0.047) is consistent with the definition of illiquidity provided by the quotient of return and trading volume. Therefore, increasing trading volume should be accompanied by decreasing illiquidity. According to the market microstructure literature, there are three main factors impacting the illiquidity (*Bank* et al., 2011):

1. Explicit trading costs like fees or taxes.
2. The existence of asymmetric distribution of information.
3. The inventory risk of market-makers.

The negative and partly significant correlation between the absolute proxies for investor attention GSV, Wiki and News, and the illiquidity implies that one of the three sources for illiquidity is related to these proxies. The relationship with explicit trading costs or inventory risk of market-makers may not be plausibly justified in this context. One reason for this is the international positioning of the companies and the various stock exchange platforms associated with the companies' stocks, of which only some can be characterized as a quote-driven market. Therefore, the results of the correlation analysis suggest that the relationship between illiquidity and the proxies for investor attention may be attributed to the existence of asymmetric distribution of information. More specifically, the negative correlation potentially indicates a reduction in asymmetric information costs and thus would identify the uninformed investor as a primary source for a rise in trading activity and liquidity or a decrease in illiquidity. The interaction ΔGSV, ΔWiki and ΔNews partially reflects the results of Table 1 as changes in ΔWiki again positively and significantly correlate with changes in ΔGSV (0.091) and ΔNews (0.062). Furthermore, the results of Table 2 imply a decrease in ΔNews at the same time as there is an increase in ΔGSV (–0.029).

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

The correlation analysis carried out in relation to the absolute values and the weekly differences support the initial hypothesis of a significant relationship between the proxies for investor attention and the relevant ratios for quantifying the liquidity of a stock. However, the comparison between absolute values and weekly differences highlights discrepancies concerning the manifestation of the correlation and therefore their validity is verified in the next section.

1. Partitioned Correlation Analysis of Search Activity and Liquidity

The preceding univariate correlation analysis is evaluated for robustness by partitioning the data set. A partial correlation shows the correlation of two variables with a number of control variables removed that could have a confounding effect on the two analyzed variables. This seems necessary due to the differences between the correlation of the absolute values and the weekly differences. Furthermore, it is possible to show which group is driving the overall correlation. Therefore, on a weekly basis according to the absolute value and the amount and direction of change of the search volume over time the data set is divided into groups. Identical to the previous section the same hypotheses are being tested, whereby in this case the focus is on the amount and direction of the search volume. As a consequence, the influence of distorting factors like the size of the company on the analysis is limited, since the weekly observation points of a specific company can be in different partitions depending on their value.

For each of the three potential proxies for investor attention, there are three subdivisions of the data set. The first partition in Table 3 divides the data set based on the absolute values of the search volume and key figures. The second partition in Table 4 considers the positive and negative changes in the search volume and divides the observation points according to the changes of the order of magnitude. The last partition differentiates based on whether there is a positive, negative or no change in search volume value.[7]

The data set of each partition is standardized in order to maintain an equal weighting in the evaluation and a two-sided t-test determines the statistical significance of the difference in the mean values. Comparing the Google search volume with the two other potential proxies for investor attention in Table 4a shows that positive changes in the Wikipedia search volume (High ΔWiki: –0.0684, Middle ΔWiki: –0.0599, Low ΔWiki: –0.1949) and news volume (High ΔNews: –0.0008, Middle Δ News: –0.1177, Low Δ News: –0.1446) respectively correlate negatively with the Google search volume. Moreover, the correlation coefficients exhibit a comparatively strong negative correlation in case of positive changes in Wikipedia search volumes.

[7] An overview of the criteria used for the partitions in tabular form and because of reasons of space the last partition can be found in Table A.4 and A.5 in the Appendix.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 3

Results of Partitioning by the Amount of Change in GSV, Wiki and News

Variables	High GSV	Middle GSV	Low GSV	High – Low
Wiki	0.1416	0.226	–0.0312	0.1728***
News	–0.6527	0.0868	–0.1266	–0.5261
TV	0.0204	0.0469	0.111	–0.0906***
TR	0.1032	–0.0887	0.007	0.0962***
R	–0.0897	–0.0886	–0.1388	0.0491
ILLIQ	–0.1934	–0.0977	–0.2639	0.0705***
Variables	High Wiki	Middle Wiki	Low Wiki	High – Low
GSV	0.4496	0.0638	–0.3769	0.8265***
News	–0.0395	–0.1792	–0.0108	–0.0287***
TV	0.374	0.4738	–0.5074	0.8814***
TR	0.117	0.1109	0.2856	–0.1686***
R	–0.1895	–0.0964	–0.0446	–0.1449
ILLIQ	–0.427	–0.3248	0.262	–0.689***
Variables	High News	Middle News	Low News	High – Low
GSV	0.1778	–0.0861	–0.1969	0.3747
Wiki	0.1202	–0.1157	–0.1406	0.2608***
TV	–0.1909	–0.0038	–0.1116	–0.0793***
TR	–0.0281	–0.1757	–0.0592	0.0311***
R	–0.0797	–0.0353	0.0355	–0.1152
ILLIQ	–0.1622	–0.1182	–0.0645	–0.0977***

Note(s): The variables TV, TR, R and ILLIQ are standardized to the mean of zero and the variance of one. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

Table 4a

Results of Partitioning by the Amount and Positive Direction of Change in GSV, Wiki and News

Variables	High ΔGSV	Middle ΔGSV	Low ΔGSV	High – Low
Wiki	0.0495	–0.0402	–0.1882	0.2377
News	–0.0422	–0.1192	–0.261	0.2188
TV	–0.0773	–0.1824	–0.2667	0.1894**
TR	–0.1	–0.1401	–0.1351	0.0351***
R	–0.0941	–0.1382	–0.2042	0.1101
ILLIQ	–0.1088	–0.1423	–0.1034	–0.0054**

(continue next page)

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

(Table 4a continued)

	High ΔWiki	Middle ΔWiki	Low ΔWiki	High – Low
GSV	−0.0684	−0.0599	−0.1949	0.1265
News	−0.008	−0.0928	0.0621	−0.0701**
TV	0.0739	−0.0273	0.004	0.0699***
TR	0.072	−0.0097	0.0521	0.0199**
R	−0.2226	−0.0972	−0.1825	−0.0401
ILLIQ	−0.0025	−0.1014	−0.1907	0.1882
	High ΔNews	Middle ΔNews	Low ΔNews	High – Low
GSV	−0.0008	−0.1177	−0.1446	0.1438***
Wiki	0.0204	−0.0757	−0.1705	0.1909
TV	−0.202	−0.2702	−0.0466	−0.1554***
TR	−0.1308	−0.2453	−0.0134	−0.1174
R	−0.1671	−0.1517	−0.1613	−0.0058***
ILLIQ	−0.2242	−0.0286	−0.3994	0.1752

Note(s): The variables TV, TR, R and ILLIQ are standardized to the mean of zero and the variance of one. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

Ultimately, it should be noted that the correlation analysis of the partitioned data sets confirms the effects observed in the previous section and thus attributes a degree of robustness to the findings of the correlation analysis. The partition analysis has verified that the changes in value of the proxies exert a negative influence on stock trading regardless of the size of the company, while when looking at the absolute values a constant positive effect for the trading activities and negative effect for the illiquidity can be observed.

Table 4b

Results of Partitioning by the Amount and Negative Direction of Change in GSV, Wiki and News

Variables	High ΔGSV	Middle ΔGSV	Low ΔGSV	High – Low
Wiki	0.0688	−0.0523	0.0525	0.0163
News	−0.0821	−0.078	−0.0982	0.0161 *
TV	−0.3441	−0.2357	−0.1781	−0.166**
TR	−0.3038	−0.1716	−0.0614	−0.2424**
R	−0.0552	−0.1502	−0.2316	0.1764
ILLIQ	−0.2154	−0.1264	−0.0969	−0.1185

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

	High ΔWiki	Middle ΔWiki	Low ΔWiki	High – Low
GSV	0.0045	0.0521	0.0186	–0.0141
News	–0.0667	–0.0725	–0.0284	–0.0383
TV	–0.0344	–0.0212	–0.1658	0.1314
TR	0.0276	–0.1178	–0.1611	0.1887
R	–0.1604	–0.175	–0.1426	–0.0178
ILLIQ	–0.0952	–0.0259	–0.1022	0.007
	High ΔNews	Middle ΔNews	Low ΔNews	High – Low
GSV	–0.1205	–0.1115	0.0165	–0.137***
Wiki	0.0206	–0.2014	–0.0066	0.0272
TV	0.0316	–0.019	–0.0548	0.0864
TR	0.0982	–0.1664	–0.0238	0.122
R	–0.4569	0.0018	–0.0351	–0.4218**
ILLIQ	–0.0918	–0.1216	–0.0724	–0.0194***

Note(s): The variables TV, TR, R and ILLIQ are standardized to the mean of zero and the variance of one. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

V. Multivariate Regression Analysis

1. Model

The multivariate panel regression model is based on an unbalanced panel data set due to missing key figures for certain observation periods. The aggregation of the weekly data to monthly data in order to create a more balanced panel data set can potentially reduce the impact of short-term, strongly influencing events (*Pöppe* et al., 2014). However, the results of the correlation analysis in the previous section display significant correlations for the panel data on a weekly basis. Furthermore, previous research on the illiquidity and Google search volume shows that the results for a weekly and a monthly aggregation are qualitatively similar (*Bank* et al., 2011). The variables of the panel data set are standardized with a mean of 0 and a standard deviation of 1 to enable the comparison across the different dependent variables. For the independent variables, a 1-week lag structure is applied with the objective to take potential endogenous interdependencies between the dependent variable and the control variables into consideration.

$$
\begin{aligned}
ILLIQ_{i,t} &= a + b_1 ILLIQ_{i,t-1} + b_2 Proxy_{i,t-1} + b_3 lnMV_{i,t-1} + b_4 R_{i,t-1} \\
&\quad + {}_5 \qquad {}_{i,t-1} + {}_6 \left(\qquad {}_{i,t-1} * \qquad {}_{i,t-1} \right) \\
&\quad + b_7 \left(Proxy_{i,t-1} \,/\, ILLIQ_{i,t-1} * B2C \right) + c_i + \mu_t
\end{aligned} \tag{9}
$$

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Equation (9) illustrates the estimated multivariate panel regression model with the illiquidity $ILLIQ_{i,t}$ of the company i in week t as the dependent variable. The regression model considers the autocorrelation by including the illiquidity $ILLIQ_{i,t-1}$ of the previous week. In reference to *Bank* et al. (2011) and *Pöppe* et al. (2014), the regression model's two central independent variables are $Proxy_{i,t-1}$, which introduces the potential proxy for investor attention in the form of the GSV, Wiki and News as a regressor and $Proxy_{i,t-1} * lnMV_{i,t-1}$ as an interaction term. The importance of the interaction term is related to the possible control for a company's Google search, Wikipedia search or news volume in regard to the company's market capitalization. According to *Bank* et al. (2011), the regression coefficient b_2 for the Google search volume as $Proxy_{i,t-1}$ and the coefficients b_6 for the interactive term with the Google search volume as $Proxy_{i,t-1}$ exhibit opposing signs and therefore opposing influences on the dependent variable $ILLIQ_{i,t}$. The underlying assumption suggests that companies with a comparatively high market capitization demonstrate a weaker relationship between the Google search volume and the illiquidity than companies with a lower market capitalization.

The remaining part of the regression model comprises the following variables, the logarithmized market capitalization $lnMV_{i,t-1}$ at the end of the previous week $t-1$, the returns $R_{i,t-1}$ and the variable $Trading\ Activity_{i,t-1}$ of the previous week. The latter variable functions as a placeholder for the variables Trading Volume $TV_{i,t-1}$ and Turnover Rate $TR_{i,t-1}$ in order to observe their influences separately. To account for different search habits, we introduce another interaction term. According to *Pöppe* et al. (2019), the majority of search queries for business-to-business companies are probably from interested investors while potential buyers are looking for products and thus for consumer-focused companies. Because the dataset being a panel data set and we already account for company specific effects with a fixed-effects model a simple B2B/B2C dummy variable isn't sufficient. So, we introduce an interaction term with B2C being 1 if the company is either a purely business-to-customer company or has a mixed business model. Next to controlling search habits we change the proxy variable in the interaction term with the autocorrelation variable $ILLIQ_{i,t-1}$ to see the effects for illiquidity for different companies. The results are in the Appendix in Table A6 for illiquidity and Table A7 for TPI. Furthermore, the regression model accounts for a time-independent error term c_1 that captures the company-specific variance and a company-independent error term μ_1 for the time-dependent variance since both these variances are not represented in the model.

In addition to the illiquidity, the presented regression model is tested with an alternative ratio for illiquidity or liquidity as a dependent variable in reference to the research of *Bank* et al. (2011) and *Pöppe* et al. (2014). This results in the following model with TPI as the dependent variable:

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

$$
\begin{aligned}
TPI_{i,t} &= a + b_1 TPI_{i,t-1} + b_2 Proxy_{i,t-1} + b_3 lnMV_{i,t-1} + b_4 R_{i,t-1} \\
&\quad + b_5 Trading\ activity_{i,t-1} + b_6 \left(Proxy_{i,t-1} * lnMV_{i,t-1}\right) \\
&\quad + b_7 \left(Proxy_{i,t-1} / ILLIQ_{i,t-1} * B2C\right) + c_i + \ _t
\end{aligned} \tag{10}
$$

In summary, the presented regression models, which show heteroscedasticity and serial correlation, are corrected by clustered standard errors and will be calculated under the assumption of fixed effects.

2. *Results*

a) Google Search Volume

Table 5 summarizes the results for seven combinations of the first regression model with ILLIQ as a dependent variable. More specifically, the first four models aim at explaining the $ILLIQ_{i,t}$ of a specific stock of a company as a function of its time-delayed value $ILLIQ_{i,t-1}$ in combination with the Google search volume $GSV_{i,t-1}$. Moreover, the control variables $lnMV_{i,t-1}$ and $R_{i,t-1}$ are added gradually in the subsequent specifications. In the model specifications (4) through (7) the variables $TV_{i,t-1}$ ($TR_{i,t-1}$) control for trading activity, while the added interaction term $GSV_{i,t-1} * lnMV_{i,t-1}$ functions as a measure to capture the non-linear relationship between the illiquidity and the Google search volume proportional to the company‘s respective market capitalization while the interaction term $GSV_{i,t-1} * B2C$ captures the effects of Google search volume on business-to-customer companies. The results of the table represent a panel regression with the GSV according to *Da* et al. (2011), as it shows robustness to short-term jumps and low frequency seasonalities. The F-statistic indicates that at least one coefficient of the individual specifications is significantly different from zero. Moreover, the R-square suggests that between 1.36 % and 6.18 % of the total sample variation in the dependent variable can be explained by the independent variable.

The auto-correlation of the illiquidity $ILLIQ_{i,t-1}$ is significantly positive on the 1 % level in each model specification and exhibits a positive effect on the illiquidity $ILLIQ_{i,t}$ with coefficients between 0.0666 and 0.1101. The delayed Google search volume $GSV_{i,t-1}$ exhibits a positive influence on the illiquidity ranging from 0.0356 to 0.0515 with a p-value that indicates a significance at least on the 10 % level. Thus, suggesting that the hypothesis of decreasing illiquidity due to an increase in public attention in the form of Google search volume should be rejected in this regression models. This result is in direct contradiction to the preceding research by *Pöppe* et al. (2014) and *Bank* et al. (2011), which identified highly significant negative coefficients between the illiquidity and the

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Google search volume. The interaction term shows at the same time negative significant coefficients which point in the same direction like *Pöppe* et al. (2014) and *Bank* et al. (2011) if the company is B2C. This could be because we expect informed investors to act immediately and uninformed investors more slowly. We imply that the kind of public interest that is captured by lagged Internet search activity represents more likely the degree to which a respective asset is traded by uninformed investors. Another line of reasoning for this result could be also that all investors act faster than one week, so that a lagged variable of one week measures the illiquidity after the trades are done and the trading volume starts to decline and the illiquidity increases.

Furthermore, significant coefficients can be observed from the lagged market capitalization $lnMV_{i,t-1}$ for all specifications, the highly significant positive effect of 0.0825 suggests that higher market capitalization is accompanied by decreased liquidity of the share in the following week. The interaction term of search volume and market capitalization displays an opposite sign to the coefficient of $GSV_{i,t-1}$ in models (5) and (7) and is not significant. Further, the results display a significant influence of the time-delayed returns $R_{i,t-1}$ on the 1 % level with a negative coefficient for all specifications ranging from –0.0345 to –0.0439. The change in sign compared to the correlation analysis is in line with previous research, which argues that preceding high returns negatively affect the subsequent illiquidity as a result of elevated awareness from the investor regarding a specific share due to extraordinary returns (*Pöppe* et al., 2014).

Moreover, Table 5 illustrates a negative influence of the trading activities on the illiquidity. The lagged trading volume $TV_{i,t-1}$ in the models (4) and (5) exhibits a comparatively strong negative influence on the illiquidity with a coefficient of –0.284 on a 1 % significance level. Similarly, the lagged turnover rate $TR_{i,t-1}$ coefficients in the models (6) and (7) indicate a negative influence on the illiquidity with a coefficient of –0.175 on the 1 % significance level. In line with current economic theory, the negative coefficients confirm that the low liquidity of a share is likely to imply lower trading activity (*Amihud* et al., 2002).

The panel regression analysis with illiquidity as a dependent variable partially confirms the results and associated hypotheses of the correlation analysis. The hypotheses concerning the relationship between illiquidity and the variables associated with trading activities, as well as the negative impact of returns on illiquidity, can thus be confirmed. In return, the results of Table 5 suggest that findings of previous research regarding increased liquidity due to comparatively higher market capitalization cannot be reproduced based on the complete panel data set. Only the interaction term with B2C companies point in this direction. Furthermore, the examination of the Google search volume and its influence on the illiquidity implies a differentiation between the time-parallel and time-delayed consideration. As a result, the illiquidity decreases in the case of an in-

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 5

RegressionRresults of the Panel Data Set with ILLIQ as the Dependent Variable

Variables	*Specifications*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ILLIQ_{i,t-1}$	0.1062***	0.1047***	0.1101***	0.0667***	0.0666***	0.0847***	0.0845***
	(0.0180)	(0.0179)	(0.0184)	(0.0157)	(0.0157)	(0.0180)	(0.0180)
$GSV_{i,t-1}$	0.0356*	0.0403*	0.0401**	0.0508***	0.0515***	0.0470**	0.0478**
	(0.0214)	(0.0206)	(0.0204)	(0.0184)	(0.0183)	(0.0198)	(0.0197)
$lnMV_{i,t-1}$		−0.0569*	−0.0535*	0.0816***	0.0825***	−0.0690**	−0.0680**
		(0.0290)	(0.0286)	(0.0260)	(0.0260)	(0.0269)	(0.0268)
$R_{i,t-1}$			−0.0394***	−0.0348***	−0.0345***	−0.0439***	−0.0435***
			(0.0123)	(0.0105)	(0.0105)	(0.0107)	(0.0107)
$TV_{i,t-1}$				−0.2843***	−0.2845***		
				(0.0133)	(0.0134)		
$TR_{i,t-1}$						−0.1746***	−0.1751***
						(0.0101)	(0.0100)
$lnMV_{i,t-1} * GSV_{i,t-1}$					−0.0121		−0.0151
					(0.0117)		(0.0130)
$GSV_{i,t-1} * B2C$	−0.0933***	−0.0915***	−0.0928***	−0.0753**	−0.0755**	−0.0782***	−0.0784***
	(0.0297)	(0.0292)	(0.0290)	(0.0301)	(0.0301)	(0.0290)	(0.0290)
Observations	12,519	12,519	12,519	12,519	12,519	12,519	12,519
R^2	0.0138	0.0159	0.0172	0.0666	0.0668	0.0483	0.0485
Adjusted R^2	−0.0119	−0.0098	−0.0086	0.0420	0.0421	0.0232	0.0234
F Statistic	57.0514***	49.4181***	42.6006***	145.044***	124.638***	103.101***	88.8522***

Note(s): The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 6

Regression Results of the Panel Data Set with TPI as the Dependent Variable

Variables	*Specifications*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$TPI_{i,t-1}$	0.1103***	0.1104***	0.1165***	0.0762***	0.0762***	0.0789***	0.0788***
	(0.0191)	(0.0191)	(0.0196)	(0.0166)	(0.0166)	(0.0159)	(0.0159)
$GSV_{i,t-1}$	0.0347*	0.0354**	0.0352**	0.0456***	0.0458***	0.0434***	0.0439***
	(0.0181)	(0.0173)	(0.0171)	(0.0159)	(0.0159)	(0.0158)	(0.0157)
$lnMV_{i,t-1}$		−0.0087	−0.0051	0.1286***	0.1289***	−0.0219	−0.0214
		(0.0306)	(0.0302)	(0.0266)	(0.0264)	(0.0268)	(0.0267)
$R_{i,t-1}$			−0.0463***	−0.0423***	−0.0422***	−0.0505***	−0.0503***
			(0.0126)	(0.0105)	(0.0105)	(0.0103)	(0.0103)
$TV_{i,t-1}$				−0.2763***	−0.2764***		
				(0.0141)	(0.0141)		
$TR_{i,t-1}$						−0.2090***	−0.2093***
						(0.0117)	(0.0117)
$lnMV_{i,t-1} * GSV_{i,t-1}$					−0.0044		−0.0087
					(0.0116)		(0.0125)
$GSV_{i,t-1} * B2C$	−0.0648**	−0.0645**	−0.0661**	−0.0482*	−0.0482*	−0.0483*	−0.0484*
	(0.0277)	(0.0279)	(0.0278)	(0.0286)	(0.0286)	(0.0263)	(0.0263)
Observations	12,519	12,519	12,519	12,519	12,519	12,519	12,519
R^2	0.0136	0.0136	0.0153	0.0617	0.0618	0.0590	0.0591
Adjusted R^2	−0.0122	−0.0122	−0.0106	0.0371	0.0370	0.0343	0.0343
F Statistic	55.8680***	42.0497***	37.8223***	133.775***	114.696***	127.507***	109.447***

Note(s): The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

crease in Google search volume as shown in the correlation analysis, but it increases due to elevated Google search volume in the previous week.

In the next step, the results are tested for robustness by applying alternative metrics for the illiquidity, which capture the price impact. The results for the multivariate regression with TPI as the dependent variable are depicted in Table 6. The autocorrelation of TPI exhibits significant positive coefficients ranging from 0.0762 to 0.1165, suggesting that similarly to the illiquidity an increase in TPI the previous week elevates the current TPI.

Moreover, the $GSV_{i,t-1}$ positively influences the $TPI_{i,t}$ with a coefficient in an interval from 0.0347 to 0.0458 thereby indicating a similar influence like on $ILLIQ_{i,t}$, but with a more significant coefficient than the results in Table 5. In terms of the coefficients, the variable for market capitalization yields the highest significant coefficient of 0.1286 on the 1% level in model (4). Consequently, a company‘s increasing market capitalization would be expected to have an elevated TPI. Moreover, $R_{i,t-1}$ exhibits a significant influence on the dependent variable in all specifications and produces a negative coefficient on the 1% level, whereas the sign and magnitude of the coefficients confirm the results in Table 5. The time-delayed trading activities $TV_{i,t-1}$ and $TR_{i,t-1}$ report consistent results in terms of the level of significance, sign and magnitude regardless of the dependent variable in Table 5 and 6. Furthermore the interaction term $GSV_{i,t-1} * lnMV_{i,t-1}$, consistent with the results in Table 5, exhibits non-significant results in the models (5) and (7). The interaction term $GSV_{i,t-1} * B2C$ shows again a negative significant coefficient indicating that the illiquidity of B2C companies is lower when the Google search volume increases than for B2B companies. An examination of the adjusted R-squared implies that, in comparison, the specifications in Table 6 exhibit a greater percentage of the sample variation in the dependent variable explained by the independent variables.

In summary, the results on panel regression in this chapter show that liquidity, measured by the illiquidity measure of *Amihud* et al. (2002) and the TPI of listed companies in the FinTech sector, generally demonstrate a positive relationship with Google's search volume. The comparison of the correlation analysis and the regression analysis regarding the Google search volume reveals a change in sign. More specifically, $GSV_{i,t-1}$ exhibits a positive coefficient and $GSV_{i,t}$ a negative coefficient in regard to their effect on the stock's illiquidity. From this, it could be concluded that, depending on the time of the increases in Google search volume, the illiquidity responds either with an increase or decrease. Thus, providing evidence that the liquidity of the stock surges only temporarily and within a week, the Google search volume induces the opposite effect. This pattern confirms the findings of earlier research, which likewise reveal time-variable correlations in respect of Google's search volume and stock market reactions. For example, the results of *Bijl* et al. (2016), demonstrated a similar effect

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

when considering the return on investment. The autocorrelation in combination with the time-delayed Google search volume explains a substantial part of the following development of $ILLIQ_{i,t}$ and $TPI_{i,t}$.

b) Wikipedia Search Volume

Analogous to the regression analysis of the Google search volume, the same regression models apply in this section, using the Wikipedia search volume as a proxy for investor attention. For reasons of efficiency and place, each individual variable will not be elaborated on explicitly in the following, but rather focusing in greater detail on the proxies for investor attention and significant differences in the control variables with respect to the previous chapter. The respective Table A.8 is in the Appendix.

Table A.8 displays the results for the Wikipedia search volume as the proxy for investor attention. The time-delayed Wikipedia search volume $Wiki_{i,t-1}$ exhibits a positive influence on the illiquidity $ILLIQ_{i,t}$ for all seven specifications but without any significant coefficient. As a consequence of this observation, there is the suggestion that increased visits to the company's Wikipedia page in the previous week, similar to an uptick in Google search volume, reduces the liquidity of the corresponding stock. Moreover, the results in Table A.8 illustrate that the time-delayed market capitalization $lnMV_{i,t-1}$ effect on $ILLIQ_{i,t}$ differs depending on the model specification. In specification (4) and (5) a significant positive coefficient can be observed on the 10 % level, which agrees with the results in the previous section, whereas the remaining specifications exhibit a negative coefficient. In terms of significance, sign, and order of magnitude, the remaining control variables behave comparatively analogous to panel regression with the Google search volume as a proxy with the exception of both interaction terms which are not significant. The regressions with TPI as the dependent variable show a significant F-statistics and this time effect of $Wiki_{i,t-1}$ on the illiquidity measure is significant for the specifications (11) to (14) on the 5 % and 10 % level. Furthermore, the previously detected negative coefficients of the $lnMV_{i,t-1}$ are now throughout positive. The autocorrelation's highly significant coefficient consistently indicates a positive effect on the TPI, while the variable $R_{i,t-1}$ displays significant negative coefficients in all models. Moreover, the control variables for the trading activity $TV_{i,t-1}$ and $TR_{i,t-1}$ as well as the interaction terms $lnMV_{i,t-1} * Wiki_{i,t-1}$ and $Wiki_{i,t-1} * B2C$ largely align in terms of significance, sign, and order of magnitude with the $ILLIQ_{i,t}$ as the dependent variable.

In a holistic assessment of the panel regression analysis with the Wikipedia search volume as a proxy, the results present evidence for a missing dependency between the liquidity of a share and the proxy in the FinTech sector, based on a five-year observation period. The significant coefficients for most of the specifi-

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

cations for $Wiki_{i,t-1}$ in Table A.8 occur only with TPI as the dependent variable. Moreover, the magnitude of the coefficients indicates a marginal effect on the liquidity, especially in comparison to the considerably stronger influences of the autocorrelation or control variables such as $lnMV_{i,t-1}$ or TV. The evaluation of the R-Squared values for the two panel regressions, show that the introduction of new control variables only marginally increases the overall explanatory power of the regression model. The comparison with the correlation analysis further supports the assumption of a highly reactive market, which compensates a large extent of the effects of a change in investment attention within a week.

c) News Volume

The following panel regression introduces the FinTech-related news volume according to official digital papers and blogs as a further proxy for investor attention. All results depicted in Table A.9 in the Appendix are significant on the 1 % level for the illiquidity measures. Additionally, the panel regressions based on the total number of news items and on the exclusively negative news items are considered separately.

The specifications (1) to (7) illustrates the results for the panel regression with the $ILLIQ_{i,t}$ as the dependent variable. The examination of the FinTech related news volume shows that all specifications exhibit a significant positive influence on the 1 % level. The interaction term $lnMV_{i,t-1} * News_{i,t-1}$ in models (5) and (7) shows an insignificant influence on the illiquidity in the next week whereas the interaction term with the business model is positive but mostly not significant indicating no difference between different business models. Moreover, the control variables $R_{i,t-1}$, $TV_{i,t-1}$ and $TR_{i,t-1}$ consistently and highly significantly exhibit a negative coefficient and in terms of magnitude similar values. The $lnMV_{i,t-1}$ coefficient analogous to the previous regressions shows a significant positive effect on the $ILLIQ_{i,t}$ for the specifications (4) and (5). The tests for the robustness of the panel regression confirm the findings. The coefficient $News_{i,t-1}$ displays one significant effect for the $TPI_{i,t}$ (specification (11)) whereas significant negative coefficients for the control variables $R_{i,t-1}$, $TV_{i,t-1}$ and $TR_{i,t-1}$ for the $TPI_{i,t}$ as well as a significant positive influence of the $lnMV_{i,t-1}$ can be observed. The results from Table A.9 suggest that the general news flow regarding FinTechs has no significant impact on the illiquidity or general trading volume of the company.

The results of a regression analysis based exclusively on negative news and blog entries are summarized in Table A.10. An examination of the regression coefficients for the volume of negative news $NegNews_{i,t}$ reveals no significant coefficients for $ILLIQ_{i,t}$. The interaction term $NegNews_{i,t} * lnMV_{i,t-1}$ produces positive coefficients. The remaining control variables depicted in Table A.10 do

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

not differ from the results in Table A.9 with regard to the sign, the significance and order of magnitude. An initial examination of the results reveals the preliminary evidence of a relationship between the increase of negative news volume followed by a decrease in adverse selection costs and therefore in stock illiquidity of a specific company, regardless of the content in the news or blog article. For the purpose of verifying the robustness of the results the $TPI_{i,t}$ is employed as a dependent variable. Considering the results of the panel regression for $TPI_{i,t}$, the coefficient for the variable $Neg\,News_{i,t-1}$ is consistently negative and only significant in specification (11). The interaction term $Neg\,News_{i,t} * lnMV_{i,t-1}$ exhibits in the specification (12) and (14) negative coefficients indicating that negative news for big companies lead to a decrease of TPI. Thus, suggesting a similar effect compared to $Wiki_{i,t}$ and $GSV_{i,t}$ as the negative coefficient implies a stronger influence of the negative news volume on companies with a comparatively elevated market capitalization. Consequently, the panel regression results in Table A.10 provide evidence for decreasing liquidity of a share in connection with increased negative sentiments in news or blogs and the same time no difference between companies with different business models. The examination of the R-Squared values for the two-panel regressions in this section demonstrates that the introduction of new control variables slightly increases the overall explanatory power of the regression model. The specifications (4) and (5), and (11) and (12) are particularly noteworthy, as the two provide the highest adjusted R-Squared values for both general and exclusively negative news. Thus providing evidence, that the illiquidity is more accurately explained with the turnover volume as a proxy instead of the turnover rate for the trading activities of a company's stock.

d) Comparison of the Results

As one major question is the difference in the investor attention proxies, this section summarizes and reviews the individual results on the basis of exemplary selected regression specifications. Table 7 depicts specification (7) for all investor attention proxies as the F-test null hypothesis shows a 1 % significance level for the different proxies Google search volume, Wikipedia page volume and news volume in general as well as explicitly negative news volume. The close comparison of the individual proxies shows that only the variable $GSV_{i,t-1}$ produces a significant positive coefficient. In this context, the variable News and negative News are considered in a differentiated way, since the regression is based on fewer observations and therefore a direct comparison with the other proxies requires a reserved assessment. The variable $Neg\,News_{i,t-1}$ also regularly fails to produce a p-value sub 0.1, whereas considering the significant specifications, the difference in coefficients is substantially smaller. The positive coefficient of the significant variable supports the notion that the liquidity of a com-

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 7

Comparison of Alternative Proxies for the Panel Regression with IllIQ as the Dependent Variable

Variables	$GSV_{i,t-1}$	$Wiki_{i,t-1}$	$News_{i,t-1}$	$Neg\,News_{i,t-1}$
Coefficient	0.0478**	0.0157	0.0803	0.0007
	(0.0197)	(0.0130)	(0.1260)	(0.1591)
Variables	$GSV_{i,t-1} * lnMV_{i,t-1}$,	$Wiki_{i,t-1} * lnMV_{i,t-1}$,	$News_{i,t-1} * lnMV_{i,t-1}$	$Neg\,News_{i,t-1} * lnMV_{i,t-1}$
Coefficient	–0.0151	–0.0083	0.0202	0.0027
	(0.0130)	(0.0123)	(0.0180)	(0.0237)
Variables	$GSV_{i,t-1} * B2C$	$Wiki_{i,t-1} * B2C$	$News_{i,t-1} * B2C$	$Neg\,News_{i,t-1} * B2C$
Coefficient	–0.0784***	–0.0189	0.0622	0.0427
	(0.0290)	(0.0214)	(0.0435)	(0.0388)
Observations	12,519	8,083	7,827	7,827
R^2	0.0591	0.0328	0.0367	0.0357
Adjusted R^2	0.0343	0.0006	0.0031	0.0020
F-statistic	109.447***	37.929***	41.150***	39.9742***

Note(s): This is only an excerpt. The complete regressions are found in the text and Appendix. The variables are standardized to the mean of zero and the variance of one. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

pany's stock reacts negatively to an increase of the Google search volume, more sensitive than to the other proxies.

In addition, Table 7 compares the interaction terms for each proxy, which only registers a significant value for the coefficient $GSV_{i,t-1} * B2C$. Therefore, the non-linear relationship between the proxies and the illiquidity proportional on the company's business model seems to occur exclusively in the case of $GSV_{i,t-1}$ as a proxy for investor attention. In summary, it can be observed that mostly Google search volume demonstrates significant results. In the case of $GSV_{i,t-1}$, the illiquidity of a company is influenced independently of its market capitalization, whereas the influence of $B2C$ shows a negative sign. This suggests that the illiquidity of business-to-customer companies is more sensitive to google search habits, while the influence of Google search volume on the illiquidity presumably affects big and small companies equally. The robustness analysis with TPI as the dependent variable generally supports the results in Table 7.

VI. Conclusion

The results identify in most specifications the Google search volume as a significant measure for the stock's illiquidity of the following week. Beyond that, the results also indicate a significant change in the relationship between the proxies for investor attention and the key figures for liquidity over the course of one week from a positive to a negative correlation. The predictive power of the significant measure Google search volume for the illiquidity is considerably smaller than the control variables and autocorrelation. The comparison of these results with similar research approaches shows that the selection of the observation period and the choice of companies under examination strongly influences the relationship between the proxies for investor attention and the illiquidity of stocks traded. The interaction terms with a dummy variable for B2C show a difference between B2B and B2C companies and confirm *Pöppe* et al. (2019) that different companies have customers with different search habits.

The findings provide a basis for the following assumptions: First, the general information-gathering process before and during the trades for investors regarding the FinTech sector relies to a certain degree on online search queries like Google. Hence, the FinTech companies' online profile exerts a considerable influence on the illiquidity of their shares, whereby the active management of the dissemination of information from other sources besides its own homepage must be factored in. A comparison with previous research demonstrates that Google's search volume in particular exhibits a significant measurable influence on the company's liquidity from various industry sectors with B2B and B2C business models, whereby the degree of influence varies which also applies to the FinTech sector. Second, the Google search volume function as an indicator

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

of short-term trading activities. Their incorporation into the companies' predictive models enables more accurate projections with regard to the development of stock's liquidity. However, the substantial differences between the delayed and real-time influences of the proxies support the assumption that the market is reacting faster to the information inherent in Google's search volume, news volume or Wikipedia search volume. The possible deductions would be on the one hand to support the finding of *Bijl* et al. (2016) of a decreasing reaction time of the market to asymmetric information in recent years, and on the other hand to emphasize the special nature of the FinTech sector as a market which, due to its economic, technical and regulatory environment, compensates more rapidly to changes in investor attention. Overall, the measures Google search volume, Wikipedia search volume, and FinTech-related news volume capture the attention of the individual investor in different phases of the information gathering process. As a result, companies have the opportunity to use online search engines as an objective and direct indicator for the underlying opinions of entire populations or clearly defined target groups. For example, this insight can enable companies to better anticipate the reaction to ad hoc announcements or to identify optimal time frames for new issues of equities and other securities. Likewise, investors can integrate these indicators into their own decision-making process and thus plan their buying and selling policy according to a more precise information situation.

References

Alt, R./*Puschmann*, T. (2016): Digitalisierung der Finanzindustrie. Berlin/Heidelberg: Springer Berlin Heidelberg.

Amihud, Y./*Mendelson*, H./*Pedersen*, L. H. (2002): Illiquidity and Stock Returns Cross-Section and Time-Series Effects, 110 – 136.

Antweiler, W./*Frank*, M. Z. (2004): Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. The Journal of Finance, 59(3), 1259 – 1294.

Asur, S./*Huberman*, B. A. (2010): Predicting the Future with Social Media. In 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology – Workshops. WI-IAT 2010 Workshops: proceedings, 31 August – 3 September 2010, Toronto, Ontario, Canada (pp. 492 – 499). [Los Alamitos, Calif.]: IEEE Computer Society.

Bank, M./*Larch*, M./*Peter*, G. (2011): Google search volume and its influence on liquidity and returns of German stocks. Financial Markets and Portfolio Management, 25(3), 31.

Barber, B. M./*Odean*, T. (2008): All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. The Review of Financial Studies, 21(2), 785 – 818.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Bijl, L./*Kringhaug*, G./*Molnár*, P./*Sandvik*, E. (2016): Google Searches and Stock Returns. International Review of Financial Analysis, 45, 150 – 156.

Blackman, P. (2019): Global fintech investment rockets to a record $111.8B in 2018, driven by mega deals: KPMG Pulse of Fintech – KPMG Global. Retrieved February 17, 2020, from https://home.kpmg/xx/en/home/media/press-releases/2019/02/global-fintech-investment-hits-record-in-2018.html.

Bollen, J./*Mao*, H./*Zeng*, X. (2011): Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1 – 8.

Bordino, I./*Battiston*, S./*Caldarelli*, G./*Cristelli*, M./*Ukkonen*, A./*Weber*, I. (2012): Web Search Queries Can Predict Stock Market Volumes. PLOS ONE, 7(7), e40014, from https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0040014&type=printable.

Choi, C. Y./*Ham*, H. B. (2015): Theoretical Study on the Business Model of Fintech Enterprises. The e-Business Studies, (16), 85 – 100.

Da, Z. H. I./*Engelberg*, J./*Gao*, P. (2011): In Search of Attention. The Journal of Finance, 66(5), 1461 – 1499.

Dimpfl, T./*Jank*, S. (2016): Can internet search queries help to predict stock market volatility?, European Financial Management, 22(2), 171 – 192.

Duan, J./*Liu*, J./*Chen*, Q. (2020): Research on the relationship between FinTech attention and its sector returns. International Journal of Economics, Finance and Management Sciences, 8(1), 57.

Forns, M. R. (2020): Pageviews Analysis – Häufig gestellte Fragen. Retrieved January 10, 2020, from https://tools.wmflabs.org/pageviews/faq/.

Google (2020): FAQ about Google Trends data – Trends Help. Retrieved January 10, 2020, from https://support.google.com/trends/answer/4365533?hl=en.

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.

Gulden, J. (2019): Automatisierte Geldanlage. Wiesbaden: Springer Fachmedien Wiesbaden.

Heath, P. T./*Schwartz*, D./*Sun*, T. (2015): The Future of Finance: The Socialization of Finance. USA: Goldman, Sachs &. Co.

Hendrikse, R./*Bassens*, D./*van Meeteren*, M. (2018): The Appleization of finance: Charting incumbent finance's embrace of FinTech. Finance and Society, 4(2), 159 – 180, from http://financeandsociety.ed.ac.uk/article/download/2870/3924.

Hinnosaar, M./*Hinnosaar*, T./*Kummer*, M. E./*Slivko*, O. (2015): Does Wikipedia Matter? The Effect of Wikipedia on Tourist Choices. SSRN Electronic Journal, 1 – 17.

Kennedy, P. (2009): A guide to econometrics (6th ed., [Nachdr.]). Malden, Mass.: Blackwell.

Lee, I./*Shin*, Y. J. (2018): Fintech: Ecosystem, business models, investment decisions, and challenges. Business Horizons, 61(1), 35 – 46.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

LexisNexis Legal & Professional (2020): Legal. Retrieved January 09, 2020, from https://www.relx.com/our-business/market-segments/legal.

Li, J./*Yu*, J. (2012): Investor attention, psychological anchors, and stock return predictability. Journal of Financial Economics, 104(2), 401 – 419.

Lo, A./*Wang*, J: (2000). Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory. Cambridge, MA.

Loesch, S. (2018): A Guide to Financial Regulation for Fintech Entrepreneurs. Chichester, UK: John Wiley & Sons, Ltd.

Lütkepohl, H./*Xu*, F. (2012): The role of the log transformation in forecasting economic variables. Empirical Economics, 42(3), 619 – 638.

Menat, R. (2016): Why We're so Excited About FinTech. In S. Chishti & J. Barberis (eds.), The FinTech Book (pp. 10 – 12). Chichester, UK: John Wiley & Sons, Ltd.

Moore, M./*Tambini*, D. (2018): Digital Dominance: The Power of Google, Amazon, Facebook, and Apple. Oxford: Oxford University Press Incorporated.

Newman, O. (2008): Online Business Sourcebook: De Gruyter.

Poley, C./*Kuffer*, J. (2020): DBIS. Retrieved January 09, 2020, from http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib_id=tuda&colors=&ocolors=&lett=fs&titel_id=1670#.

Pöppe, T./*Kiesel*, F./*Kolaric*, S./*Schiereck*, D. (2019): Information or Noise: How Twitter Facilitates Stock Market Information Aggregation. Darmstadt Technical University, Department of Business Administration, Economics and Law, Institute for Business Studies (BWL).

Pöppe, T./*Schiereck*, D./*Zielinski*, F. (2014): Das Google-Suchvolumen als Liquiditätsindikator des Aktienhandels: Evidenz aus dem weltweiten Agrarsektor. Credit and Capital Markets – Kredit und Kapital, 47(4), 611 – 640.

QYResearch (2019): FinTech Market Report: Global Market size, Market Forecast, Market Share, Industry outlook, Demand Analysis, Market Report 2019 – 2025: Valuates Reports. Retrieved February 13, 2020, from https://reports.valuates.com/market-reports/QYRE-Othe-2W194/fintech-market.

Redding, L. S. (1996): Noise Traders and Herding Behavior. IMF Working Papers: Working Paper No. 96/104. Washington, D.C: International Monetary Fund.

Report of the Board of IOSCO (2017): FR02/2017 IOSCO Research Report on Financial Technologies (Fintech), from https://www.iosco.org/library/pubdocs/pdf/IOSCOPD554.pdf.

Rubin, A./*Rubin*, E. (2010): Informed Investors and the Internet. Journal of Business Finance & Accounting, 37(7 – 8), 841 – 865.

Sabherwal, S./*Sarkar*, S. K./*Zhang*, Y. (2011): Do Internet Stock Message Boards Influence Trading? Evidence from Heavily Discussed Stocks with No Fundamental News. Journal of Business Finance & Accounting, 38(9 – 10), 1209 – 1237.

Simon, H. A. (1955): A Behavioral Model of Rational Choice. The Quarterly Journal of Economics, 69(1), 99 – 118.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Skan, J./*Dickerson*, J./*Masood*, S. (2015): The Future of Fintech and Banking: Digitally disrupted or reimagined?, Accenture, from https://www.accenture.com/_acnmedia/accenture/conversion-assets/dotcom/documents/global/pdf/dualpub_11/accenture-future-fintech-banking.pdf.

Smales, L. A. (2012): Non-Scheduled News Arrival and High-Frequency Stock Market Dynamics: Evidence from the Australian Securities Exchange. SSRN Electronic Journal.

Smales, L. A. (2014): News sentiment and the investor fear gauge. Finance Research Letters, 11(2), 122 – 130.

Smales, L. A. (2021): Investor attention and global market returns during the COVID-19 crisis. International Review of Financial Analysis, 73, 101616.

Smith, S./*Tran*, V./*Perera*, T. (2016): 2016 Top Markets Report Financial Technology: A Market Assessment Tool for U.S. Exporters, from https://legacy.trade.gov/topmarkets/financial-tech.asp.

Smolinski, R./*Gerdes*, M./*Siejka*, M./*Bodek*, M. C. (2017): Innovationen und Innovationsmanagement in der Finanzbranche. Wiesbaden: Springer Fachmedien Wiesbaden.

Statista (2019): Search engine market share worldwide 2019 | Statista. Retrieved December 20, 2019, from https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/.

Tetlock, P. C. (2007): Giving Content to Investor Sentiment: The Role of Media in the Stock Market. Journal of Finance, 62(3), 1139 – 1168.

Tiberius, V./*Rasche*, C. (2017): FinTechs. Wiesbaden: Springer Fachmedien Wiesbaden.

Vozlyublennaia, N. (2014): Investor attention, index performance, and return predictability. Journal of Banking & Finance, 41, 17-35.

Wang, Q./*Myers*, M. D./*Sundaram*, D. (2013): Digital Natives und Digital Immigrants. Wirtschaftsinformatik, 55(6), 409 – 420.

Xu, S. X./*Zhang*, X. (2013): Impact of Wikipedia on Market Information Environment: Evidence on Management Disclosure and Investor Reaction. MIS Quarterly, 37(4), 1043 – 1068.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Appendix

Table A.1

List of the Company Names

Company	*Google Search Term*	*Wikipedia Site*	*Classification*
ACI Worldwide Inc.	ACI WORLDWIDE	Yes	Payment Transaction
Alliance Data Systems Inc.	ALLIANCE DATA SYS	Yes	Financial Data Analytics
American Express	AMERICAN EXPRESS	Yes	Payment transactions
Axos Financial Inc.	AXOS FINANCIAL	Yes	Payment transactions
Black Knight Inc.	BLACK KNIGHT	Yes	Financial Software
Bottomline Technologies Inc.	BOTTOMLINE TECH	No	Financial Software
Broadridge Financial Solutions Inc.	BROADRIDGE Financial Solutions	Yes	Financial Software
CARDTRONICS	CARDTRONICS	No	Financial Software
Cboe Global Markets Inc.	CBOE GLOBAL MARKETS	Yes	Platforms
CME Group Inc.	CME GROUP	Yes	Platforms
CORELOGIC	CORELOGIC	Yes	Financial Data Analytics
CoStar Realty Information Inc.	COSTAR GROUP	Yes	Financial Data Analytics
Envestnet	ENVESTNET	Yes	Financial Data Analytics
Equifax Inc.	EQUIFAX	Yes	Financial Data Analytics
Euronet Worldwide Inc.	EURONET WORLDWIDE	Yes	Payment transactions
Evertec Inc.	EVERTEC	No	Financial Software
FactSet Research Systems Inc.	FACTSET RESEARCH SYS	Yes	Financial Software
Fair Isaac Corporation	FAIR ISAAC	Yes	Platforms
Fidelity National Information Services Inc.	FIS	Yes	Financial Software
Fiserv Inc.	Fiserv	Yes	Payment Transaction
FLEETCOR TECHNOLOGIES INC.	FLEETCOR TECH	Yes	Payment Transaction

(continue next page)

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

(Table A1 continued)

Company	*Google Search Term*	*Wikipedia Site*	*Classification*
Global Payments Direct Inc.	GLOBAL PAYMENTS	Yes	Financial Software
Green Dot Corporation	GREEN DOT	Yes	Platforms
IHS Markit	IHS MARKIT	Yes	Financial Data Analytics
Intercontinental Exchange Inc.	INTERCONTINENTAL XCH	Yes	Financial Data Analytics
Jack Henry & Associates Inc.	JACK HENRY & ASSOC	Yes	Financial Software
LendingClub Corporation	LENDINGCLUB CORP	Yes	Platforms
LendingTree Inc.	LENDINGTREE	Yes	Platforms
MarketAxess Holdings Inc.	MARKETAXESS HOLDINGS	Yes	Platforms
Mastercard	MASTERCARD	Yes	Payment Transaction
Moody's Investors Service Inc.	MOODY'S CORP	Yes	Financial Data Analytics
MSCI Inc.	MSCI	Yes	Financial Data Analytics
Nasdaq Inc.	NASDAQ	Yes	Platforms
OnDeck	ON DECK CAPITAL	Yes	Financial Software
PayPal Inc.	PAYPAL HOLDINGS	Yes	Platforms
RealPage Inc.	REALPAGE	Yes	Financial Data Analytics
S&P Global	S&P GLOBAL	Yes	Financial Data Analytics
SEI Investments Company	SEI INVESTMENTS	Yes	Platforms
SQUARE CL.A	SQUARE	Yes	Financial Software
SS&C Technologies Inc.	SS&C TECHNOLOGIES	Yes	Financial Software
Thomson Reuters Corp	THOMSON REUTERS	Yes	Financial Data Analytics
TransUnion LLC.	TRANSUNION	Yes	Financial Data Analytics
Verisk Analytics Inc.	VERISK ANALYTICS	Yes	Financial Data Analytics
VIRTU Financial Inc.	VIRTU FINANCIAL	Yes	Financial Software
Visa Inc.	VISA	Yes	Payment Transaction

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Company	*Google Search Term*	*Wikipedia Site*	*Classification*
Western Union Holdings Inc.	WESTERN UNION	Yes	Payment Transaction
WEX Inc.	WEX	Yes	Payment Transaction
WisdomTree Investments Inc.	WISDOMTREE	Yes	Platforms

Table A.2

Statistical Properties of the Regression Variables

Variable	*Mean*	*Median*	*Std. Dev.*	*Min*	*Max*	*Max – Min*	*Observations*
GSV	45.62994	47	22.02963	0	100	100	12690
Wiki	3070.17	1091	7005.101	1	275809	275808	8535
News	100.9259	69	92.95284	2	411	409	12690
Neg News	29.20741	26	25.65587	0	190	190	12690
TV	19.2816	19.28329	1.330421	15.19033	23.31961	8.129285	12304
TR	0.0405765	0.0282608	0.0597467	0.0018176	2.680001	2.678184	12304
lnMV	8.991054	8.98887	1.357927	5.451081	12.67847	7.22739	12408
R	0.3594096	0.4044118	4.100068	–50.56338	45.96599	96.52937	12297
ILLIQ	1.096023	0.5241009	1.552584	0	21.51641	21.51641	12297
TPI	98.70867	67.11609	129.4804	0	2550.007	2550.007	12297

Table A.3

Results of the Augmented Dickey Fuller Test for Stationarity

Variable	*Inverse* γ^2	*p-Value*
GSV	916,5754	0
Wiki	541,6967	0
News	271,4669	0
TV	623,4979	0
TR	945,7916	0
ILLIQ	1343,4234	0
R	1973,13	0

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.4

Criteria for the Partitioned Data

	Partition 1		
	GSV	*Wiki*	*News*
Low	0 < GSV < 40	0 < Wiki800	0 < News < 30
Middle	40 < GSV < 71	800 < Wiki < 2000	30 < News < 114
High	71 < GSV	2000 < Wiki	114 < News
	Partition 2a		
	GSV	*Wiki*	*News*
Low	$0 < \Delta GSV < 8$	$0 < \Delta Wiki < 75$	$0 < \Delta News < 8$
Middle	$8 < \Delta GSV < 20$	$75 < \Delta Wiki < 200$	$8 < \Delta News < 21$
High	$20 < \Delta GSV$	$200 < \Delta Wiki$	$21 < \Delta News$
	Partition 2b		
	GSV	*Wiki*	*News*
Low	$0 > \Delta GSV > -8$	$0 > \Delta Wiki > -75$	$0 > \Delta News > -8$
Middle	$-8 > \Delta GSV > -20$	$-75 > \Delta Wiki > -200$	$-8 > \Delta News > -21$
High	$-20 > \Delta GSV$	$-200 > \Delta Wiki$	$-21 > \Delta News$
	Partition 3		
	GSV	*Wiki*	*News*
Positive	$0 < \Delta GSV$	$0 < \Delta Wiki$	$0 < \Delta News$
Negative	$0 > \Delta GSV$	$0 > \Delta Wiki$	$0 > \Delta News$

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.5

Results of Partition 3

Variables	*Positive changes in GSV*	*Negative changes in GSV*	*Positive – Negative*
Wiki	–0.0045	–0.1158	0.1113
News	–0.1284	–0.1175	–0.0109 ***
TV	–0.1473	–0.3472	0.1999***
TR	–0.1132	–0.3327	0.2195***
R	–0.1689	–0.1469	–0.022 **
ILLIQ	–0.1333	–0.1619	0.0286

Variables	*Positive changes Wiki*	*Negative changes in Wiki*	*Positive – Negative*
GSV	–0.1228	–0.0169	–0.1059***
News	–0.0176	–0.0552	0.0376***
TV	0.0432	–0.0319	0.0751***
TR	0.048	0.0228	0.0252***
R	–0.1828	–0.1409	–0.0419
ILLIQ	–0.0583	–0.0944	0.0361***

Variables	*Positive changes News*	*Negative changes in News*	*Positive – Negative*
GSV	–0.0126	–0.1424	0.1298***
Wiki	–0.0172	–0.0111	–0.0061
TV	–0.1476	0.086	–0.2336***
TR	–0.0918	0.027	–0.1188***
R	–0.2162	–0.1626	–0.0536***
ILLIQ	–0.1327	–0.0831	–0.0496***

Note(s): This table shows the results of partitioning by the direction of change in GSV, Wiki and News. The variables TV, TR, R and ILLIQ are *standardized* to the mean of zero and the variance of one. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.6

Results of ILLIQ with a Lagged ILLIQ in the Interaction Term

Variables	*Specifications*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ILLIQ_{i,t-1}$	0.1002***	0.0977***	0.1026***	0.0580***	0.0580***	0.0758***	0.0757***
	(0.0211)	(0.0210)	(0.0216)	(0.0191)	(0.0191)	(0.0213)	(0.0212)
$GSV_{i,t-1}$	0.0057	0.0111	0.0105	0.0271*	0.0277*	0.0223	0.0230
	(0.0175)	(0.0169)	(0.0169)	(0.0157)	(0.0156)	(0.0164)	(0.0162)
$lnMV_{i,t-1}$		–0.0586**	–0.0553*	0.0810***	0.0819***	–0.0708**	–0.0698**
		(0.0298)	(0.0294)	(0.0272)	(0.0271)	(0.0279)	(0.0278)
$R_{i,t-1}$			–0.0387***	–0.0343***	–0.0340***	–0.0434***	–0.0430***
			(0.0123)	(0.0105)	(0.0105)	(0.0106)	(0.0107)
$TV_{i,t-1}$				–0.2863***	–0.2865***		
				(0.0136)	(0.0137)		
$TR_{i,t-1}$						–0.1763***	–0.1769***
						(0.0100)	(0.0099)
$lnMV_{i,t-1} * GSV_{i,t-1}$					–0.0117		–0.0148
					(0.0116)		(0.0129)
$ILLIQ_{i,t-1} * B2C$	0.0245	0.0276	0.0286	0.0307	0.0304	0.0316	0.0313
	(0.0373)	(0.0370)	(0.0367)	(0.0304)	(0.0304)	(0.0335)	(0.0335)
Observations	12,519	12,519	12,519	12,519	12,519	12,519	12,519
R^2	0.0118	0.0140	0.0152	0.0654	0.0656	0.0470	0.0472
Adjusted R^2	–0.0140	–0.0118	–0.0106	0.0408	0.0409	0.0219	0.0221
F Statistic	48.5218***	43.3902***	37.6589***	142.283***	122.248***	100.227***	86.3630***

Note(s): This table shows the regression results of the panel data set with ILLIQ as the dependent variable and an interaction term with the lagged illiquidity. The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.7

Results of TPI with a Lagged TPI in the Interaction Term

Variables	*Specifications*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$TPI_{i,t-1}$	0.0891***	0.0890***	0.0949***	0.0566***	0.0567***	0.0586***	0.0586***
	(0.0175)	(0.0176)	(0.0184)	(0.0168)	(0.0168)	(0.0159)	(0.0159)
$GSV_{i,t-1}$	0.0145	0.0155	0.0148	0.0310**	0.0311**	0.0288**	0.0292**
	(0.0139)	(0.0134)	(0.0134)	(0.0130)	(0.0130)	(0.0126)	(0.0125)
$lnMV_{i,t-1}$		−0.0110	−0.0075	0.1268***	0.1271***	−0.0240	−0.0235
		(0.0309)	(0.0306)	(0.0272)	(0.0270)	(0.0272)	(0.0272)
$R_{i,t-1}$			−0.0458***	−0.0420***	−0.0419***	−0.0502***	−0.0500***
			(0.0127)	(0.0106)	(0.0106)	(0.0104)	(0.0104)
$TV_{i,t-1}$				−0.2769***	−0.2769***		
				(0.0139)	(0.0139)		
$TR_{i,t-1}$						−0.2096***	−0.2099***
						(0.0111)	(0.0111)
$lnMV_{i,t-1} * GSV_{i,t-1}$					−0.0039		−0.0083
					(0.0115)		(0.0123)
$TPI_{i,t-1} * B2C$	0.0700*	0.0705*	0.0710*	0.0639*	0.0638*	0.0660*	0.0658*
	(0.0410)	(0.0412)	(0.0409)	(0.0344)	(0.0345)	(0.0339)	(0.0339)
Observations	12,519	12,519	12,519	12,519	12,519	12,519	12,519
R^2	0.0138	0.0138	0.0155	0.0622	0.0622	0.0596	0.0596
Adjusted R^2	−0.0119	−0.0119	−0.0103	0.0375	0.0375	0.0348	0.0348
F Statistic	56.7782***	42.8253***	38.3581***	134.877***	115.632***	128.755***	110.500***

Note(s): This table shows the regression results of the panel data set with TPI as the dependent variable and an interaction term with the lagged TPI in the interaction term. The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.8

Results of the Wikipedia Search Volume Panel Regression Analysis with ILLIQ as the Dependent Variable

Variables	*Specifications*					
	(1)	(2)	(3)	(4)	(5)	(6)
$ILLIQ_{i,t-1}$	0.0898***	0.0896***	0.0934***	0.0578***	0.0578***	0.0741***
	(0.0239)	(0.0236)	(0.0247)	(0.0213)	(0.0213)	(0.0245)
$TPI_{i,t-1}$						
$Wiki_{i,t-1}$	0.0019	0.0025	0.0024	0.0147	0.0164	0.0135
	(0.0122)	(0.0120)	(0.0119)	(0.0112)	(0.0118)	(0.0120)
$lnMV_{i,t-1}$		−0.0346	−0.0313	0.0661*	0.0663*	−0.0525
		(0.0444)	(0.0438)	(0.0360)	(0.0362)	(0.0418)
$R_{i,t-1}$			−0.0250*	−0.0260**	−0.0258**	−0.0349***
			(0.0145)	(0.0131)	(0.0131)	(0.0130)
$TV_{i,t-1}$				−0.2621***	−0.2622***	
				(0.0171)	(0.0170)	
$TR_{i,t-1}$						−0.1516***
						(0.0121)
$lnMV_{i,t-1} * Wiki_{i,t-1}$					−0.0065	
					(0.0111)	
$Wiki_{i,t-1} * B2C$	−0.0315	−0.0315	−0.0321	−0.0137	−0.0137	−0.0189
	(0.0203)	(0.0203)	(0.0204)	(0.0199)	(0.0199)	(0.0213)
Observations	8,083	8,083	8,083	8,083	8,083	8,083
R^2	0.0084	0.0089	0.0094	0.0508	0.0509	0.0328
Adjusted R^2	−0.0242	−0.0238	−0.0234	0.0193	0.0192	0.0006
F Statistic	22.069***	17.583***	14.855***	69.808***	59.872***	44.176***

Note(s): This table shows the regression results of the panel data set with ILLIQ in the specifications (1) to (7) and TPI in the specifications (8) to (14) as the dependent variables. The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Specifications							
(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
0.0741***							
(0.0246)							
	0.0811***	0.0801***	0.0847***	0.0549***	0.0547***	0.0578***	0.0575***
	(0.0194)	(0.0195)	(0.0206)	(0.0185)	(0.0185)	(0.0176)	(0.0176)
0.0157	0.0119	0.0114	0.0114	0.0232*	0.0275**	0.0248*	0.0299**
(0.0130)	(0.0130)	(0.0130)	(0.0129)	(0.0128)	(0.0132)	(0.0133)	(0.0139)
−0.0523		0.0338	0.0374	0.1320***	0.1325***	0.0143	0.0147
(0.0419)		(0.0383)	(0.0379)	(0.0324)	(0.0325)	(0.0347)	(0.0347)
−0.0345***			−0.0300**	−0.0316***	−0.0309***	−0.0408***	−0.0400***
(0.0131)			(0.0137)	(0.0120)	(0.0120)	(0.0116)	(0.0115)
				−0.2481***	−0.2485***		
				(0.0162)	(0.0162)		
−0.1519***						−0.1792***	−0.1798***
(0.0122)						(0.0147)	(0.0148)
−0.0083					−0.0162		−0.0192*
(0.0123)					(0.0099)		(0.0107)
−0.0189	−0.0401**	−0.0402**	−0.0409**	−0.0240	−0.0240	−0.0258	−0.0258
(0.0214)	(0.0199)	(0.0199)	(0.0201)	(0.0198)	(0.0199)	(0.0216)	(0.0218)
8,083	8,083	8,083	8,083	8,083	8,083	8,083	8,083
0.0328	0.0071	0.0075	0.0082	0.0451	0.0454	0.0403	0.0406
0.0006	−0.0256	−0.0252	−0.0246	0.0134	0.0135	0.0084	0.0086
37.929***	18.526***	14.865***	13.01***	61.624***	53.081***	54.703***	47.254***

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.9

Results of the News Volume Panel Regression Analysis with ILLIQ as the Dependent Variable

Variables	*Specifications*					
	(1)	(2)	(3)	(4)	(5)	(6)
$ILLIQ_{i,t-1}$	0.0877***	0.0877***	0.0922***	0.0558***	0.0558***	0.0714***
	(0.0226)	(0.0225)	(0.0236)	(0.0204)	(0.0203)	(0.0239)
$TPI_{i,t-1}$						
$News_{i,t-1}$	0.0097	0.0082	0.0121	−0.1122	−0.0829	0.0521
	(0.1334)	(0.1332)	(0.1325)	(0.1296)	(0.1263)	(0.1360)
$lnMV_{i,t-1}$		−0.0277	−0.0240	0.0629**	0.0556**	−0.0427
		(0.0324)	(0.0318)	(0.0256)	(0.0278)	(0.0304)
$R_{i,t-1}$			−0.0301**	−0.0301**	−0.0296**	−0.0396***
			(0.0148)	(0.0132)	(0.0131)	(0.0132)
$TV_{i,t-1}$				−0.2550***	−0.2556***	
				(0.0161)	(0.0161)	
$TR_{i,t-1}$						−0.1607***
						(0.0130)
$lnMV_{i,t-1} * News_{i,t-1}$					0.0214	
					(0.0172)	
$News_{i,t-1} * B2C$	0.0640	0.0622	0.0624	0.0823*	0.0811*	0.0635
	(0.0480)	(0.0474)	(0.0474)	(0.0435)	(0.0433)	(0.0438)
Observations	7,827	7,827	7,827	7,827	7,827	7,827
R^2	0.0092	0.0097	0.0103	0.0541	0.0543	0.0365
Adjusted R^2	−0.0248	−0.0245	−0.0239	0.0212	0.0213	0.0030
F Statistic	23.504***	18.432***	15.800***	72.047***	62.045***	47.714***

Note(s): This table shows the regression results of the panel data set with ILLIQ in the specifications (1) to (7) and TPI in the specifications (8) to (14) as the dependent variables. The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Specifications							
(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
0.0714***							
(0.0238)							
	0.0818***	0.0810***	0.0864***	0.0561***	0.0561***	0.0568***	0.0568***
	(0.0189)	(0.0190)	(0.0200)	(0.0181)	(0.0181)	(0.0172)	(0.0172)
0.0803	–0.4190	–0.4178	–0.4106	–0.5430*	–0.5419	–0.3759	–0.3736
(0.1260)	(0.3138)	(0.3132)	(0.3124)	(0.3301)	(0.3304)	(0.2722)	(0.2716)
–0.0499		0.0255	0.0296	0.1129***	0.1126***	0.0090	0.0084
(0.0332)		(0.0294)	(0.0289)	(0.0248)	(0.0275)	(0.0255)	(0.0279)
–0.0391***			–0.0355***	–0.0363***	–0.0362***	–0.0458***	–0.0458***
(0.0131)			(0.0137)	(0.0118)	(0.0118)	(0.0115)	(0.0115)
				–0.2389***	–0.2389***		
				(0.0159)	(0.0160)		
–0.1612***						–0.1908***	–0.1909***
(0.0127)						(0.0134)	(0.0133)
0.0202					0.0008		0.0017
(0.0180)					(0.0189)		(0.0165)
0.0622	0.0251	0.0268	0.0272	0.0445	0.0444	0.0276	0.0275
(0.0435)	(0.0475)	(0.0474)	(0.0471)	(0.0426)	(0.0429)	(0.0385)	(0.0387)
7,827	7,827	7,827	7,827	7,827	7,827	7,827	7,827
0.0367	0.0077	0.0080	0.0090	0.0475	0.0475	0.0455	0.0455
0.0031	–0.0264	–0.0262	–0.0254	0.0144	0.0142	0.0123	0.0121
41.150***	19.468***	15.277***	13.688***	62.862***	53.875***	60.039***	51.458***

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table A.10

Results of the Negative News Volume Panel Regression Analysis with TPI as the Dependent Variable

Variables	*Specifications*					
	(1)	(2)	(3)	(4)	(5)	(6)
$ILLIQ_{i,t-1}$	0.0886***	0.0886***	0.0931***	0.0571***	0.0571***	0.0724***
	(0.0228)	(0.0226)	(0.0237)	(0.0208)	(0.0207)	(0.0241)
$TPI_{i,t-1}$						
$Neg\ News_{i,t-1}$	−0.0597	−0.0534	−0.0546	−0.0779	−0.0908	0.0061
	(0.1515)	(0.1488)	(0.1504)	(0.1479)	(0.1635)	(0.1363)
$lnMV_{i,t-1}$		−0.0292	−0.0255	0.0601**	0.0590**	−0.0444
		(0.0316)	(0.0310)	(0.0246)	(0.0258)	(0.0294)
$R_{i,t-1}$			−0.0297**	−0.0298**	−0.0296**	−0.0392***
			(0.0148)	(0.0133)	(0.0133)	(0.0133)
$TV_{i,t-1}$				−0.2525***	−0.2527***	
				(0.0161)	(0.0161)	
$TR_{i,t-1}$						−0.1603***
						(0.0130)
$lnMV_{i,t-1}$* Neg $News_{i,t-1}$					0.0064	
					(0.0226)	
Neg $News_{i,t-1}*B2C$	0.0466	0.0456	0.0453	0.0502	0.0501	0.0427
	(0.0428)	(0.0422)	(0.0421)	(0.0373)	(0.0371)	(0.0389)
Observations	7,827	7,827	7,827	7,827	7,827	7,827
R^2	0.0085	0.0090	0.0097	0.0527	0.0528	0.0357
Adjusted R^2	−0.0255	−0.0252	−0.0246	0.0198	0.0197	0.0021
F Statistic	21.7187***	17.1827***	14.7760***	70.1607***	60.1587***	46.6371***

Note(s): This table shows the regression results of the panel data set with ILLIQ in the specifications (1) to (7) and TPI in the specifications (8) to (14) as the dependent variables. The variables are standardized to the mean of 0 and the variance of 1. The significance levels of 10 %, 5 % and 1 % are represented by *, ** and ***.

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Specifications							
(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
0.0724***							
(0.0240)							
	0.0811***	0.0803***	0.0857***	0.0559***	0.0558***	0.0563***	0.0562***
	(0.0191)	(0.0191)	(0.0201)	(0.0182)	(0.0182)	(0.0173)	(0.0173)
0.0007	–0.4735	–0.4795	–0.4791	–0.5104*	–0.4916	–0.4175	–0.3926
(0.1591)	(0.2957)	(0.2948)	(0.2928)	(0.3084)	(0.3302)	(0.2573)	(0.2809)
–0.0449		0.0260	0.0301	0.1122***	0.1138***	0.0093	0.0115
(0.0308)		(0.0284)	(0.0279)	(0.0234)	(0.0247)	(0.0244)	(0.0254)
–0.0391***			–0.0359***	–0.0366***	–0.0368***	–0.0461***	–0.0463***
(0.0133)			(0.0137)	(0.0119)	(0.0119)	(0.0116)	(0.0116)
				–0.2366***	–0.2364***		
				(0.0161)	(0.0161)		
–0.1603***						–0.1904***	–0.1904***
(0.0129)						(0.0135)	(0.0134)
0.0027					–0.0093		–0.0123
(0.0237)					(0.0225)		(0.0218)
0.0427	0.0113	0.0122	0.0120	0.0154	0.0155	0.0082	0.0082
(0.0388)	(0.0416)	(0.0417)	(0.0414)	(0.0364)	(0.0365)	(0.0339)	(0.0340)
7,827	7,827	7,827	7,827	7,827	7,827	7,827	7,827
0.0357	0.0079	0.0083	0.0092	0.0471	0.0472	0.0456	0.0457
0.0020	–0.0262	–0.0260	–0.0251	0.0140	0.0139	0.0124	0.0123
39.9742***	20.0582***	15.7490***	14.0929***	62.3664***	53.5100***	60.1741***	51.6761***

DOI https://doi.org/10.3790/ccm.54.4.589 | Generated on 2025-07-26 12:55:29
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/