Microblogging Sentiment Investor, Return and Volatility in the COVID-19 Era: Indonesian Stock Exchange

Putri FARISKA¹, Nugraha NUGRAHA², Ika PUTERA³, Mochamad Malik Akbar ROHANDI⁴, Putri FARISKA⁵

Received: November 20, 2020 Revised: January 25, 2021 Accepted: February 03, 2021

Abstract

The covid-19 pandemic scenario caused the most extensive economic shocks the world has experienced in decades. Maintaining financial performance and economic stability is essential during the pandemic period. In these conditions, where movement is severely restricted, media consumption is considered to be increasing. The social media platform is one of the media online used by the public as a source of information and also expressing their sentiment, including individual investors in the capital market as social media users. Twitter is one of the social media microblogging platforms used by individual investors to share their opinion and get information. This study aims to determine whether microblogging sentiment investors can predict the capital market during pandemics. To analyze microblogging sentiment investors, we classified sentiment using the phyton text mining algorithm and Naïve Bayesian text classification into level positive, negative, and neutral from November 2019 to November 2020. This study was on 68 listed companies on the Indonesia stock exchange. A Vector Autoregression and Impulse Response is applied to capture short and long-term impacts along with a causal relationship. We found that microblogging sentiment investor has a significant impact on stock returns and volatility and vice-versa. Also, the response due to shocks is convergent, and microblogging investors in Indonesia are categorized as a “news-watcher” investor.

Keywords: Microblogging Investor Sentiment, Volatility, Return, VAR, Naïve Bayesian

JEL Classification Code: C11, C63, G11, G41, O16

1. Introduction

World Bank has released data related to economic shocks caused by pandemic Covid-19 that has occurred for the last several decades. The imposition of restrictions on human movement to break the chain of virus spread caused economic shocks. Conditions where movement is severely restricted in the last few months, media consumption is considered to be increasing along with internet consumption worldwide. The social media platform is one of the media online used by the public as a source of valid nor hoax information and also used to express their sentiment on a daily basis. Information is considered as something that can influence stock price movements towards a new equilibrium known as the concept of market efficiency. Qian and Rasheed (2006) have indicated that news is difficult to predict, so the stock market price will follow a random walk pattern and produce predictions with no more than 50% accuracy. The development of research in the field of behavioral finance with a big data approach showed that even though news or information is unpredictable, the initial indicators can be extracted from social media, one of which is Twitter (Bollen et al., 2011). The behavioral finance research area that discusses how a person’s sentiment can predict the stock market is sentiment investor. Research conducted by Antweiler and

¹First Author. Doctoral Student in Pasca Sarjana, Universitas Pendidikan Indonesia, Indonesia. Email: putri.fariska@upi.edu
²Professor. Faculty of Economics and Business, Universitas Pendidikan Indonesia, Indonesia. Email: nugraha@upi.edu
³Lecturer. Faculty of Economics and Business, Universitas Pendidikan Indonesia, Indonesia. Email: ikaputerawaspada@upi.edu
⁴Corresponding Author. Lecturer, Faculty of Economics and Business, Universitas Islam Bandung, Indonesia (Postal Address: Purnawarman Street, No. 59, Tamansari, Bandung Wetan, Bandung, West Java, 40117, Indonesia) Email: moch.malik@gmail.com
⁵Lecturer. Faculty of Economics and Business, Telkom University, Indonesia. Email: putrifariska@telkomuniversity.ac.id

© Copyright: The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Frank (2004) has indicated that the stock message board can predict market volatility with a statistically significant stock return. Microblogging investor sentiment also has a strong prediction on market returns. The accuracy of this prediction is consistent with the behavioral finance hypothesis (Oh & Sheng, 2011). In making decisions, investors not only look at financial information (P/E, Tobin Q) but also at information from social media to reflect their sentiments besides liquidity, and VIX explains the relationship between social media and the stock market (Chousa et al., 2016).

Maintaining stock market volatility is one of the indicators of financial performance that every country must maintain during this pandemic period. High volatility indicates uncertainty in the market and tends to fluctuate, and the volatility of stock returns has a significant impact on future market movements under the impact of shocks (Nguyen & Nguyen, 2019). Baker and Jeffrey (2006) have indicated that stocks with high volatility generate low returns in subsequent periods. Investor interest is positively influenced by previous stock price performance, and investor sentiment from posting on the internet has predictive power for volatility and trading volume (Kim & Kim, 2014). It makes sentiment investors predict future stock returns either in aggregate or at the corporate and individual level. Furthermore, Zhang et al. (2011) also have indicated that opinion on Twitter shows a significant positive correlation with stock market volatility. It has indicated that the social media platform has a significant impact on a stock market’s financial feature.

The increase in the number of social media users during the Covid-19 pandemic, especially for those in developing countries, has influenced the way people look in terms of seeking out information or sharing information with the public because social influence originating from expert investors is more influential than the Book Value Per Share (Rahayu et al., 2021). This pandemic situation makes people anxious from a health perspective causes psychological instability for investors when investing in the market (Luu & Luong, 2020). From the description above, the purpose of this study is to determine whether microblogging investor sentiment can predict the stock market in terms of market volatility and return.

2. Literature Review

2.1. The Relation between Microblogging Investor Sentiment Volatility and Stock Returns

Sentiment Investor is a study that discusses the relationship between social interaction and investment (Cabarcos et al., 2019). It is not easy to measure sentiment or human emotions because they used surveys as a tool to determine investor emotions in traditional ways. With technological advancement and the increase in the number of internet users and social media, it has become making measurement easier than before (Sahana & Anuradha, 2019). Additionally, Sahana and Anuradha (2019) have indicated that the internet and social media, as a development of information technology, provide a platform to express emotions to the public and greatly influence overall public opinion. Twitter, Facebook, or Web blog, and other social media provide a form of blogging that allows users to text write short updating is called microblogging service. Through microblogging service, people easily share information and opinions by writing short updating because of the function of microblogging as a mediated social practice (Dijck, 2011) or a container of interacting activities.

Some scholars have indicated that through social media such as Twitter, blogs or forums such as Yahoo! Financial message boards or news website can capture investor sentiments, which have an impact on the capital market (Antweiler & Frank, 2004; Sahana & Anuradha, 2019; Petit et al., 2019). Sentiment investor microblogging divided into three categories: those are news media content (Tetlock, 2007), data search or a query on the internet (Da et al., 2014), and posting on social media (Antweiler & Frank, 2004; Bollen et al., 2011).

To capture investor sentiment on social media, some scholars use the text classifier method to convert and measure sentiment investor microblogging. Some of the methods used are naive bayesian classification methods (Antweiler & Frank, 2004; Sprenger et al., 2013). We use the top-down analysis in this research; investors’ sentiment is measured and its impact on the stock market. Some researchers with this approach are (Sprenger et al., 2013; Coelho, 2019), which summarizes investor sentiment as a determinant of the stock market.

2.2. Hypotheses

According to the theoretical model developed by Van Bommel (2003), investors who have information with limited transaction capacity are motivated to disseminate information about share prices or have the desire to post messages about the shares they are trading. Individual investors are market members with limited access to information. Individual investors are market members with limited access to information (Hirshleifer & Teoh, 2003). With the improvement of information technology, social media platforms can be sources of information for individual investors. Capturing every opinion that appears on social media can be measured easily and even becomes a driving force for developing research in the area of behavioral finance, especially on investor sentiment through text classification technology.

Emotions, opinions, or information conveyed by investors can be easily analyzed with the naïve Bayesian classification method (Antweiler & Frank, 2004; Sprenger et al., 2013) to
deduce the relationship between the general public’s views on stocks and changes in the stock market (Bourezk et al., 2020). Sprenger et al. (2013) use a limited rationality model approach in which individuals are subject to the persuasion bias proposed by DeMarzo et al. (2003), and assume that individual investors on social media platforms reflect the nature of the model where group opinion is not only seen from its accuracy but also seen from how well a person is connected to their social network.

We will look at the causal relationship between the variables. Recent research describing the relationship between local daily happiness sentiment extracted from Twitter and stock returns indicate an interdependence between online activities and the stock market (Zhao, 2020). Moreover, also to see the shocks caused by sentiment investor microblogging, volatility, and return on the stock market during the Covid-19 pandemic. Based on the previous researches, we propose the following hypothesis:

**H1:** Sentiment investor microblogging can predict volatility and returns on the capital market.

**H2:** Sentiment investor microblogging, volatility, and return on the capital market have a causal relationship.

**H3:** The shocks sentiment investor microblogging, volatility, and returns on the capital market are convergent.

### 3. Research Method

#### 3.1. Naïve Bayesian Text Classification

To measure opinion or information on microblogging investor sentiment, we conducted a message classification approach in line with the Naïve Bayesian classification method and research indicated by (Antweiler & Frank, 2004; Sprenger et al., 2013). Naïve Bayesian is the most widely used algorithm in text classification. Daily messages or opinions are taken from the social media platform Twitter based on #stockcode, which are listed on the exchange consisting of 68 active stock codes taken from November 2019 – November 2020. The method of taking mining data in the form of daily data uses the approach taken by (Oh & Sheng, 2011), consists of 5 (five) phases pipeline system technique (1) Downloading data, (2) Pre-processing, (3) Sentiment Analysis, (4) Prediction Classification, and (5) Evaluation and Analysis. After Data Cleansing was performed, the number of opinions about 2,840 tweets with 324 usernames. Furthermore, the data is entered into the Naïve Bayesian model, consisting of 80% training data, namely 2,280 data and 20% testing data of 560 data. The following is a table that presents sample data for tweets that were randomly selected for training data with manual labels as follows:

<table>
<thead>
<tr>
<th>Sample Tweets</th>
<th>Manual Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trimmed Profit, AKRA Soars 15.38% @ MarketMover.ID</td>
<td>Positif</td>
</tr>
<tr>
<td>Alumina Makes Antam Loss</td>
<td>Negatif</td>
</tr>
<tr>
<td>#BKSL #BlueGolden Shares #Stockpick today - Early stock pick on 11/10/19 - Currently increasing + 3.4%</td>
<td>Positif</td>
</tr>
<tr>
<td>#BBCA - still eagerly awaiting bbca - TradingView</td>
<td>Netral</td>
</tr>
<tr>
<td>Will it happen or not? #BBCA</td>
<td>Netral</td>
</tr>
<tr>
<td>Hopefully, tomorrow #bbni will rise sharply. So you can just thin again</td>
<td>Netral</td>
</tr>
</tbody>
</table>

#### 3.2. Financial Data Set, Variables

The financial data used were taken from November 2019 – November 2020 using daily data. In this study we use...
daily volatility based on intra-day data constructed Parkinson (1980), as follows:

\[
\text{VOL} = \frac{\left( \ln \left( H_t - \ln \left( L_t \right) \right) \right)^2}{4 \ln (2)} \tag{1}
\]

Where \( H_t \) and \( L_t \) show the highest and lowest daily stock prices, while the data for stock returns, the calculations used in this study are based on simple return calculations, namely as follows:

\[
\text{Return} = \frac{R_t + 1 - R_{t-1}}{R_t - R_{t-1}} \tag{2}
\]

\( R_t \) is the return in a certain period, the stock return for one period in the future, so that the stock return calculation is the quotient between the difference between the stock price next year and the current stock price divided by the stock price. Both the opinion on Twitter and the volatility and rate of return on shares are calculated based on each stock code. The definitions of the variables used in this study are as follows:

### 3.3: Model Specifications

The VAR model is a statistical approach used in this study, with several important analyses, including forecasting, Impulse Response, forecast decomposition variance, and causality test (Juanda & Junaidi, 2012). In addition, in this study, to prove the proposed hypothesis, the Causality Test is testing the causal relationship between the variables of the Vector Autoregressive (VAR) system, which is tested using the Granger Causality test. Based on the literature review, the regression model proposed is as follows:

\[
\text{Volatility} = f(\text{Sentiment Investor, Return}) \tag{3}
\]

\[
\text{Return} = f(\text{Volatility, Sentiment Investor}) \tag{4}
\]

\[
\text{Sentiment Investor} = f(\text{Volatility, Return}) \tag{5}
\]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microblogging Investor Sentiment</td>
<td>Positive, Negative, and Neutral Signals from Twitter User daily Opinions</td>
</tr>
<tr>
<td>Volatility</td>
<td>Stock movements in short-term active trader are calculated daily</td>
</tr>
<tr>
<td>Stocks Return</td>
<td>Return Shares are calculated based on 68 share codes of issuers listed on the exchange</td>
</tr>
</tbody>
</table>

### 4. Results and Discussion

We used a unit root test using the ADF (Augmented Dickey-Fuller) method to see stationary data. It can be seen in table 6 that the three variables used are considered to be stationary at the level * \( \alpha < 0.01 \), * \( \alpha < 0.05 \), and *** \( \alpha < 0.10 \) provided that the absolute value of the \( F \)-statistic is < critical value. For this reason, all data is stationary, and the next step is to create a VAR model.

Table 4 shows the VAR model based on the optimal lag. In the first VAR model, where volatility is the dependent variable, it is known that microblogging investor sentiment in the \( t-1 \) and \( t-2 \) periods has an opposite relationship with the volatility of period \( t \). According to Hoffmann and Post (2015), this is caused by a structural change in the mindset of investors related to previous investment experience or interpreting situations subjectively (Mitroi & Oproiu, 2014; Malmendier et al., 2020). This result is in line with the opinion of research conducted by (Petit et al., 2019) that argues sentiment appears as vital information and captures information on market variables related to microblogging investor sentiment as well as market volatility. Like-wise with the stock returns in period \( t-1 \) has an opposite relationship with volatility. However, the stock returns in the period \( t-2 \) have a unidirectional relationship. Meanwhile, the volatility in period \( t-1 \) and \( t-2 \) has a direct relationship with the sentiment in period \( t \). Whereas in the second VAR model, where the return is the dependent variable, it can be seen that in period \( t-1 \), microblogging investor sentiment has a direct relationship with stock returns, but in period \( t-2 \), there is an opposite relationship; this is due to a consistent reversal pattern. With sentiment errors that lead to temporary price errors (Da et al., 2014). However, the volatility in period \( t-1 \) and \( t-2 \) has a direct relationship with the return of shares, as well as return in period \( t-1 \) and period \( t-2 \) has a direct relationship with the rate of return in period \( t \).

In the third model, where microblogging investor sentiment becomes the dependent variable, the result is that the sentiment in period \( t-1 \) and period \( t-2 \) has a direct relationship with a sentiment in period \( t \). This illustrates that sentiment in the previous period still affects sentiment in period \( t \); this is due to the bias of conservatism, where once individuals form an impression, they are slow to change that impression in the face of new evidence. Investors remain skeptical about new information and only gradually update their views (Pompian, 2011). While the rate of return in period \( t-1 \) and \( t-2 \) has a direct relationship with the sentiment in period \( t \), volatility in the \( t-1 \) period has a significant negative effect. However, contrary to the \( t-2 \) volatility, which has a significant positive effect, this is due to the momentum where momentum occurs because “traders” move slowly when news appears, or
momentum appears when “trader” overreacts to previous news when other news comes. According to Hong and Stein (2007), the news will spread slowly to “news-watchers” and react gradually to the news resulting in “underreaction.” For this reason, it can be concluded that Indonesia’s microblogging investors are “news-watcher” investors. From the results of the description above, it can be concluded that microblogging investor sentiment can predict the volatility and rate of return of shares and, at the same time, answer the first hypothesis in this study.

Table 4: Var Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Return (−1)</th>
<th>Sentiment (−2)</th>
<th>Volatility (−1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return (−2)</td>
<td>0.120784</td>
<td>0.211554</td>
<td>0.024481</td>
</tr>
<tr>
<td>Sentiment (−1)</td>
<td>0.002209</td>
<td>0.264342</td>
<td>−0.000638</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.325750</td>
<td>0.126198</td>
<td>0.373203</td>
</tr>
</tbody>
</table>

| Variable          | Sentiment (−2) | Return (−1) | Volatility (−1) |
|-------------------|----------------|-------------|-----------------
| Sentiment (−2)    | −0.001058      | 0.135101    | −0.000423      |
| Volatility (−1)   | −0.113161      | −2.833515   | 0.510211       |
| R-squared         | 0.049697       | 0.386345    | 0.138707       |

Table 5: Granger Causality Test

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>F-statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentiment does not Granger Cause Return</td>
<td>4.65196</td>
<td>0.0096</td>
</tr>
<tr>
<td>Return does not Granger Cause Sentiment</td>
<td>3.21555</td>
<td>0.0403</td>
</tr>
<tr>
<td>Volatility does not Granger Cause Return</td>
<td>4.54950</td>
<td>0.0107</td>
</tr>
<tr>
<td>Return does not Granger Cause Volatility</td>
<td>4.17022</td>
<td>0.0156</td>
</tr>
<tr>
<td>Volatility does not Granger Cause Sentiment</td>
<td>5.06410</td>
<td>0.0064</td>
</tr>
<tr>
<td>Sentiment does not Granger Cause Volatility</td>
<td>2.97086</td>
<td>0.0514</td>
</tr>
</tbody>
</table>

Table 5 use to answer the second hypothesis in this study; it appears that microblogging investor sentiment has a significant effect on stock returns and vice versa at the significance level or \( \alpha < 0.01 \) and \( \alpha < 0.05 \), so it can be concluded that between microblogging investor sentiment and stock returns have a two-way causality relationship. This is in line with the opinion proposed by (Hoffmann & Post, 2015), which states that the rate of return has a strong impact on the rate of return on sentiment formation. Petit, et al. (2019) states that sentiment influences the future return rate. Volatility is also considered a two-way causal relationship with microblogging investor sentiment at the significance level or \( \alpha < 0.01 \) and \( \alpha < 0.10 \). Like Antweiler and Frank (2004), who use the online sentiment on yahoo finance to predict volatility in the market, and Da et al. (2014), investor sentiment is closely related to transitory daily volatility. This study also captures a two-way causality relationship on volatility and stock returns at a significance level or \( \alpha < 0.05 \). In other words, volatility has a significant effect on the rate of return and vice versa.

Figure 1 shows the convergent effect of each variable used in this study using the Impulse Response Function or IRF analysis approach. In the first graph, the response of return to sentiment shows a movement that is getting closer to the balance point or returning to the previous balance point. This shows that the impact of the response received by stock returns due to 10 months of investor sentiment shocks
is convergent, or the shock response will disappear over time and will not leave a permanent effect on stock returns. The response return to shocks caused by investor sentiment at the beginning of the response will be positive and move closer to the equilibrium point. The response received by volatility due to shocks to the investor is also convergent. The response to these shocks will disappear over time. It will not leave a permanent effect on volatility in the stock market with a negative initial response and move closer to zero. Like-wise, with the response received by investor sentiment due to shocks caused by stock returns and stock market volatility; as a result, these shocks will eventually disappear and leave no permanent effect. The impact of investor sentiment on volatility and stock returns or vice versa has a short term impact, and this factor is in line with the results of research on investor underreaction where they sometimes make mistakes where they do not react to financial news and over the next six months, these errors are gradually corrected because stock prices slowly move towards levels that should be (Barberis et al., 1998). This explanation also indicates that the market is inefficient.

5. Conclusions

First, with the increasing number of social media users during the Covid-19 pandemic, information, news, or opinions posted on social media, especially on the Twitter platform, have been converted into positive, negative, and neutral sentiments. This sentiment appears as a strong source of opinion or information. It impacts stock returns with a significant positive impact and a significant negative impact on market volatility. The conservatism bias is a factor in the relationship between microblogging investor sentiment and financial features, and the research concludes that microblogging investors in Indonesia are included in the “news-watcher” investor category. Second, microblogging investor sentiment, stock returns, and market volatility have a two-way causality, this is in line with the opinion of previous research built by (Hoffmann & Post, 2015) for the relationship between sentiment and stock returns, and Da et al. (2014) stated investor sentiment is closely related to transitory daily volatility. Third, shocks during the Covid-19 pandemic will not leave a permanent impact or shows a convergent effect for microblogging investor sentiment shocks on stock returns and market volatility.

References


Coelho, J., D’almeida, D., Coyne, S., Gilkerson, N., Mills, K., & Madiraju, P. (2019). Social Media and Forecasting Stock


