INVESTOR SENTIMENT BASED ON SEARCH ENGINE DATA FOR PREDICTING STOCK RETURNS IN INDONESIA INDUSTRIAL SECTOR

Yusuf Murtadlo Hidayat
Faculty of Economics and Business Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

Nugraha Nugraha
Faculty of Economics and Business Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

Disman Disman
Faculty of Economics and Business Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

Maya Sari
Faculty of Economics and Business Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

Muhammad Rayhan Bustam
Faculty of Humanities, Universitas Komputer Indonesia, Bandung, Indonesia

ABSTRACT
This study aims to investigate investor sentiment and its effect on the industrial sector in Indonesia. In the study, investor sentiment was extracted from data obtained from search engines. Then the data were used to see how the sentiment affected the stock return in each listed industrial company. Fifteen industrial sector companies listed on the stock exchange were selected; each was analyzed for their search volume and the effect on stock returns using panel data regression from March 2020 to April 2022. The result shows that investor sentiment affects the level of stock returns in the industrial sector in Indonesia. This result indicates that with the increase in search volume on search engines, which reflects positive sentiment, there will be an increase in stock trading transactions and vice versa. This study's findings will help investors make investment decisions, especially in the industrial sector.

Keywords: investor sentiment; search engine-data; industrial sector; search-based data; stock return

DOI: http://dx.doi.org/10.15549/jeeecar.v10i6.1500
INTRODUCTION

The topic of sentiment analysis is one of the active areas in today’s studies on behavioral finance. In addition to risk and return considerations (Kristanti et al., 2022), an investor’s investment decisions are influenced by emotional things, one of which is sentiment (Blajer-Goleciewska, 2018; Quang et al., 2023). Since behavioral finance has been proven to affect financial performance (Khan et al., 2017), researchers have begun to figure out how to measure those sentiments. They are measuring and trying to find this sentiment’s effect on the stock market (Kim & Kim, 2014).

Sentiment can be generally measured by two types of measurement, namely direct and indirect measurements. The direct measurement usually comes from a survey of investors, and it can be realized by asking the investors how their emotions and feelings relate to the market conditions. Besides, the indirect measurement is obtained from economic, financial, and other variables, which are expected to reflect the investors’ feelings and emotions (Brochado, 2020). Various studies on measuring sentiment have been carried out for decades, and there have even been studies combining these two measurements simultaneously (Beer et al., 2011).

Due to the rapid growth of the internet and big data in recent years, sentiment measurement methods are also developing rapidly. Several studies have focused on sentiment measurement using microblogging. Bollen et al. (2011) used Twitter feeds to represent investing moods and then examined the impact of those moods on Dow Jones’ financial performance. As a result, data from Twitter has an accuracy rate of more than 80% for predicting financial performance. Social media activity is also used to measure investor sentiment, as Siganos et al. (2017) researched. Investor sentiment studied was obtained from Facebook status update data; positive and negative sentiments were then associated with stock price volatility. In addition, news and disclosure on the internet are also used as data sources to determine the polarity of sentiment toward stock indices (Eachempati & Srivastava, 2022). Jin et al. (2020) performed stock prediction using social media websites, and they compared to find the most efficient predictive methods. Moreover, Songling et al. (2021) demonstrated that Industrial Companies’ Financialization was affected by investor sentiment, while Pham et al. (2021) used Google search results to examine the relationship between investor sentiment and the level of bank stock prices in Vietnam. Research on investor sentiment on stock returns gives different results; according to Akarsu and Süer (2022), significant results are found in developing countries.

Indonesia’s industrial sector has grown significantly over the years, as has the industrial sector in Asia’s leading nations, which has also demonstrated significant development and durability. The industrial sector of Indonesia, which has the largest economy in Southeast Asia, has aided in the expansion of the country’s economy. Government regulations, infrastructural development, technology improvements, and domestic and international economic considerations impact the industrial sector’s performance and future growth potential in both regions. Investor perceptions of the growth potential of Indonesia’s industrial sector influence investor sentiment towards the sector. Indonesia’s consumer market size, government programs to encourage industrial development, and increasing local and regional demand are examples of factors that can encourage investors. Sentiment can quickly alter in response to new information, market developments, or changes in political and monetary conditions.

However, a study on investor sentiment analysis to predict financial performance based on Google search data in the industrial sector has yet to be carried out, especially a study analysis for the last ten years, from 2012 to 2021. Therefore, this study was conducted to determine how investor sentiment from Google search data affects stock returns in Indonesia’s industrial sector. At the study stage, the stock ticker was used as a keyword to extract the Google Trends search volume for company names in the relevant industry sector. The search volume delta was then compared to the stock price of each company. The study was also conducted with the hope that it can be a reference for researchers and a source of information in investment decisions for investors.
LITERATURE REVIEW

According to Fama's Efficient Market Hypothesis, an efficient stock market reflects all accessible information (Fama, 1970). According to EMH, every investor makes logical investment decisions based on the available facts. However, a field of finance known as behavioral finance has also emerged, which holds that investors' choices are not solely based on logic but are partly impacted by psychological aspects. An inefficient market and the potential for anomalous outcomes might result from an investor bias. The existence of deviations can support the claim that market prices do not accurately reflect all information.

Behavioral theory has expanded into several fields of study outside of human psychology and sociology. Behavioral finance focuses on the behavior of financial agents in financial markets, which tend to be irrational, and how this can affect their investment decisions; this is a sociological and psychological synthesis related to economics. This theory holds that a person's behavior towards decision-making activities cannot be rationally and irrationally explained. As a result, investors' behavior when making decisions regarding investment activities, purchasing, and selling securities, and portfolio diversification is also irrational (Baker & Wurgler, 2006).

A theory called the Adaptive Market Hypothesis (AMH) aims to bridge the gap between behavioral finance theory and EMH theory. According to AMH, markets are typically efficient, but inefficiency and adaptation may arise due to shifting market conditions and investor behavior. This hypothesis acknowledges that investor behavior can impact market dynamics and that investors are not always perfectly rational (Lo, 2005).

Stock price volatility within a specific time frame might be predicted by investor sentiment. However, to measure this sentiment, many proxies can be used to see how the sentiment can affect stock price volatility. Investor sentiment is defined by Brauer (1993) as the investor's opinion, belief, or speculation, which are typically influenced by the investor's emotions regarding future cash flows and the risks associated with the investment (Brauer et al., 1993). The tendency to speculate is one aspect of investor sentiment that could be defined. Based on this definition, even if arbitrage forces are equal across stocks, sentiment drives the relative demand for speculative investment and consequently results in a cross-sectional effect.

Researchers have examined how investor sentiment affects stock market performance in a substantial and expanding body of literature. Chen (2017) investigated investor sentiment impact on global stocks return, while Nguyen and Pham (2018) examined search-based sentiment and stock market reactions in Vietnam. Generally, it has been discovered that investor sentiment is linked to cross-section and future returns using the proxies derived from surveys (Lee et al. 2002).

Proxies' investor sentiment was derived from non-survey data, including data mining and textual analysis (Trichilli, 2019; Mcgurk et al., 2020). Trichilli (2019) examined Google search terms to determine whether they can be used to forecast the dynamic connection between investor sentiment and index returns. According to their findings, googling investors' sentiment is the leading net transmitter of shocks in MENA country's market indexes. These findings demonstrate the reliable and substantial predictive power of searching investor sentiment to identify investor behavior, particularly during market instability, and to forecast market returns in MENA financial markets.

According to Da et al. (2011), the search volume gathered from Google trends can represent the behavior of investors in the broader population. Additionally, Da et al. (2011) believe that Google's aggregate search frequency is a good indicator of investor attention and evident concern since individual investors first use the search engine before searching for stocks to buy.

METHODOLOGY

The data used in this research was investor sentiment obtained from the search volume of stock tickers on Google Trends for the period March 2020 to April 2022; then, according to Takeda (2014), the increase or decrease is calculated by finding the difference between the search volume for a specific month (SV_t) and the search volume in the previous month (SV_{t-1}) as in the following equation.

\[ \Delta SV = SV_t - SV_{t-1} \]  

Detailed information about investor sentiment obtained from Google trends is described in a
previous study by Da et al. (2021) and Huang et al. (2019).

Stock return data (Rt) were obtained from 15 companies in the industrial sector listed on the Indonesia Stock Exchange (IDX) from March 2020 to April 2022. The stock return was calculated using the following equation

\[ R_t = \frac{p_t - p_{t-1}}{p_{t-1}} \]  

(2)

Then, to determine the effect of investment sentiment on stock returns, it is obtained by using panel data regression analysis with the following equation:

\[ R_{it} = \alpha + \beta_1 S_{it} + e_{it} \]  

(3)

where

- \( R_{it} \): stock return rate at the i-company and t-time
- \( S_{it} \): Google search volume delta at the i-company and t-time

**DISCUSSION**

Stock return volatility in the industrial sector from May 2020 to April 2022 is presented in Figure 1. As shown in Figure 1, several companies experienced quite varied fluctuations. This result is clarified by the descriptive statistics in Table 1, which shows a relatively large standard deviation for several companies.

The surge in stock returns for MARKJK issuers in mid-2020 may have been due to the corporate action taken in the form of dividend distribution. Apart from MARKJK, a significant return increase will also occur for MLPLJK in 2021, allegedly due to the issuance of a rights issue at that time. MLPLJK is an issuer that has the highest stock return volatility compared to other issuers in the industrial sector.

![Figure 1: Indonesia’s Industrial Companies Stock Return March 2020 – April 2022](source: author's work)

![Table 1: Descriptive statistics of Indonesia’s Industrial Companies Stock return (March 2020-April 2022)](source: author's work)
The results of each company's keyword search on Google Trends for March 2020 to April 2022 are then processed by calculating the difference in changes between the current month and the last month. These results are shown in Figure 2, and the descriptive statistics are described in Table 2. The issuer with the most significant volume of search intensity is SCCO JK; SCCOJK also has the most significant search intensity volatility compared to other industrial sector issuers.

Figure 2: Indonesia's Industrial Companies in Google Trends (March 2020 – April 2022)
Source: author’s work.

Table 2: Descriptive statistics of Indonesia’s Industrial Companies in Google Trends (March 2020 – April 2022)

<table>
<thead>
<tr>
<th>Company</th>
<th>Sample size</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABMMJK</td>
<td>24</td>
<td>-37.25</td>
<td>46.25</td>
<td>1.5938</td>
<td>15.12265</td>
</tr>
<tr>
<td>VOKSJK</td>
<td>24</td>
<td>-50.25</td>
<td>33.50</td>
<td>0.0208</td>
<td>19.95919</td>
</tr>
<tr>
<td>UNTRJK</td>
<td>24</td>
<td>-38.50</td>
<td>53.50</td>
<td>-0.3958</td>
<td>21.34129</td>
</tr>
<tr>
<td>SCCOJK</td>
<td>24</td>
<td>-53.25</td>
<td>41.75</td>
<td>-1.1979</td>
<td>26.15350</td>
</tr>
<tr>
<td>MLPLJK</td>
<td>24</td>
<td>-28.00</td>
<td>28.40</td>
<td>0.2750</td>
<td>10.18952</td>
</tr>
<tr>
<td>MARKJK</td>
<td>24</td>
<td>-27.25</td>
<td>17.90</td>
<td>.3875</td>
<td>10.49296</td>
</tr>
<tr>
<td>LIONJK</td>
<td>24</td>
<td>-19.70</td>
<td>27.40</td>
<td>.4188</td>
<td>11.82098</td>
</tr>
<tr>
<td>KOBXJK</td>
<td>24</td>
<td>-29.75</td>
<td>27.25</td>
<td>.6944</td>
<td>15.20820</td>
</tr>
<tr>
<td>KBLIJK</td>
<td>24</td>
<td>-42.15</td>
<td>38.40</td>
<td>.0799</td>
<td>20.50126</td>
</tr>
<tr>
<td>HEXAJK</td>
<td>24</td>
<td>-35.20</td>
<td>54.00</td>
<td>1.0799</td>
<td>14.82334</td>
</tr>
<tr>
<td>DYANJK</td>
<td>24</td>
<td>-31.50</td>
<td>19.80</td>
<td>.0000</td>
<td>10.71054</td>
</tr>
<tr>
<td>BMTRJK</td>
<td>24</td>
<td>-16.00</td>
<td>40.75</td>
<td>.3646</td>
<td>10.18440</td>
</tr>
<tr>
<td>BHITJK</td>
<td>24</td>
<td>-24.50</td>
<td>57.00</td>
<td>.7535</td>
<td>14.03990</td>
</tr>
<tr>
<td>ASIIJK</td>
<td>24</td>
<td>-24.00</td>
<td>23.15</td>
<td>-0.5903</td>
<td>11.44985</td>
</tr>
<tr>
<td>ASGRJK</td>
<td>24</td>
<td>-49.10</td>
<td>36.85</td>
<td>-.1354</td>
<td>19.29726</td>
</tr>
</tbody>
</table>

Source: author’s work.
Using the panel data of regression analysis procedure, results of the investor sentiment have a significant effect on stock returns; the panel data regression equation is obtained as follows:

\[ R_{it} = 0.029088 + 0.001492 S_{it} + e_{it} \]  \hspace{1cm} (4)

Further testing is used to find the best regression model to determine which model is better between the common and fixed effects; the Chow test is carried out. As shown in Table 3, the probability of the cross-section F-test is 0.000; it is smaller than the alpha value. Thus, the common effect is the best model between the expected and fixed effects.

**Table 3: Chow-test results**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>d.f.</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-section F</td>
<td>0.974027</td>
<td>0.4797</td>
</tr>
<tr>
<td>Cross-section Chi-square</td>
<td>14.011140</td>
<td>0.4489</td>
</tr>
</tbody>
</table>

Source: author's work.

The next step is determining the best regression model between the common and random effects; the Lagrange Multiplier test is carried out. The result is that the Lagrange Multiplier value of 8.171988 is smaller than the value of the chi-square table (0.0039), so the random effects model is the best model for estimating in this research.

More detailed Random Effect results can be seen in Table 4. The probability F-test of 0.004042, smaller than alpha, indicates that the model fits the data significantly.

**Table 1: Random-effect results**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>S. Error</th>
<th>t-Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>0.001492</td>
<td>0.000516</td>
<td>2.891958</td>
<td>0.0041</td>
</tr>
<tr>
<td>C</td>
<td>0.029088</td>
<td>0.008375</td>
<td>3.473256</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>Rho</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-section random</td>
<td>0.004094</td>
<td></td>
<td>0.0007</td>
<td></td>
</tr>
<tr>
<td>Idiosyncratic random</td>
<td>0.154302</td>
<td></td>
<td>0.9993</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weighted Statistics</th>
<th>Mean dependent var</th>
<th>0.029108</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.023842</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.020996</td>
<td>0.155816</td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.154171</td>
<td>8.152700</td>
</tr>
<tr>
<td>F-statistic</td>
<td>8.377607</td>
<td>1.638664</td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>0.004042</td>
<td></td>
</tr>
</tbody>
</table>

Even though the significance level is not very strong, the search intensity generally improves stock return. The current study's findings are consistent with those of Takeda and Wao (2014), who examined the relationship between Google search intensity and stock prices in Japan. Takeda and Wao (2014) discovered that while there was a strong positive association between search volume and transaction volume, there was only a marginally more vital positive link between search volume and stock returns. Because there are fewer individual investors in Japan than corporate investors, there is assumed to be a weak correlation between search volume and stock price. As a result, the search volume on Google may not accurately reflect individual investors in Japan.

The current findings are also consistent with previous studies, such as Beer et al. (2013) study, whose results show that investor sentiment significantly affects stock returns in France. According to research by Beer (2013), stock returns in the short-term market can be predicted using investor sentiment collected from Google search volume. For smaller issuers than for larger issuers, there is a far stronger correlation between search volume and stock returns. This result is allegedly due to the solid fundamental aspects of large issuers. These results are relevant to other research; Ekinci &
Bulut (2020) demonstrate that search volume was linked to stock returns, particularly in stocks with small capitalizations.

Salisu et al. (2013) study results show that investor attention based on Google trends significantly affects stock returns across eleven major sectors in The United States (Salisu et al., 2013). In addition, similar results are also found in the Khan et al. (2020) study. They proposed a study on household investors’ sentiment index based on Google search volume and US industry stock returns. The industrial sector is one of the sectors with a significant impact on stock returns, as measured by the volume of Google searches on stock return listing companies for most sectors in the US stock market.

CONCLUSION AND RECOMMENDATION

The analysis results show that stock return in Indonesia’s industrial sector is affected by investor sentiment. These results show that the increase in search volume shows positive investor sentiment significantly affects stock returns. The results of this study will aid investors in decision-making regarding investments and in predicting stock returns using information from Google Trends. Future studies related to this topic are to expand the unit of analysis for just not one sector; it is even possible to conduct studies in several countries. It is also necessary to develop more accurate keywords or search terms so that investor sentiment can be better reflected.

ACKNOWLEDGEMENT

The BPI supported this research – Puslapdik- Ministry of Education, Culture, Research, and Technology – Republic of Indonesia.

REFERENCES


ABOUT THE AUTHORS

Yusuf Murtadlo Hidayat, email: yusuf.murtadlo@upi.edu

**Yusuf Murtadlo Hidayat** is one of lecturers at the Business Education Study Program at the Universitas Pendidikan Indonesia. He holds a bachelor's degree at Padjadjaran University, majoring in Statistics, and a Masters's degree in Applied Statistics at Padjadjaran University. Currently, he is a doctoral candidate in the postgraduate program at the Indonesian University of Education.

**Nugraha** is one of the professors at the Indonesian Universitas Pendidikan Indonesia (UPI). His research area are behavioral finance and financial management. He also conducted much research, training, community service, and scientific work, especially in accounting and financial management.

**Disman** is one of the professors at the Indonesian Universitas Pendidikan Indonesia (UPI). He completed his undergraduate studies at the IKIP Bandung, majoring in General Economics Education in 1983, second strata at UNPAD, majoring in Economics in 1990, and doctoral at the Indonesian University of Education with a major in Social Studies-Economy Education in 2005. He has also conducted a lot of research, training, community service, and scientific work in the form of books and journal articles which were published and published in national and international journals.

**Maya Sari** is one of the lecturers at Universitas Pendidikan Indonesia. Was completed her doctoral degree in Universitas Pendidikan Indonesia. Many scientific works have been published in the research area of financial management and personal finance.

**Muhammad Rayhan Bustam** is one of the lecturers at the English Study Program at Universitas Komputer Indonesia. He received his master's degree at Universitas Padjadjaran, majoring in English Linguistics. Currently, he is completing his study by taking a doctoral degree in a postgraduate program at Universitas Pendidikan Indonesia, majoring in Linguistics.