iCOBEST
International Conference on Business Economics, Social Sciences & Humanities
UNIKOM

News Sentiment, News Intensity, and Price Movement of Indonesia's 45 Most Liquid Stock Index

M Nagib [1*], Z A Husodo [2]

[1,2] Graduate Program in Management, Faculty of Economics and Business,
University of Indonesia, Depok, Indonesia

E-mail: s.n.alatas@gmail.com

Abstract. This study analyzed the sentiment of the constituent news of the LQ45 index on its index movement by considering the news intensity. Two Sentiment analysis methods are used in this study. The Loughran and McDonald's dictionary as a measure of financial sentiment, then TextBlob sentiment analysis as a measure of general sentiment. We found that news intensity decreased during the financial crisis. Interestingly, news intensity increased significantly after the Covid-19 crisis. Then the constituent news sentiment has a relatively strong correlation with the index movement. Through linear regression analysis, the evidence suggests that the effect of sentiment on index returns is more significant when news intensity increases. Derived from these findings, we concluded that news sentiment is one of the unignorably price-forming factors in further analysis.

Keywords : Sentiment Analysis, News Intensity, Price Movement, Correlation Analysis, Linear Regression

1. Introduction

One of the crucial prices forming instruments in the capital market is the information received by the investors [9]. Empirical studies have shown that current market prices reflect the latest published information [3]. In order to achieve a reasonable price, the interaction between buyers and sellers in the capital market influenced the stock price movement. Supply and demand are strongly influenced by circulating information such as news. In order to transform news information, we can calculate the sentiments or nuances contained in it.

Many studies have proven that sentiment affects stock price movements [5]. Then [14] and [12] concluded that sentiment in the news is positively related to stock price movements. However, the impact of sentiment on changes in stock prices frequently does not occur immediately [1]. Moreover, through news exploitation from a broader data source, the accumulation of sentiment can serve as an early warning sign of a financial crisis [10]. In wider use, news sentiment analysis even can be used to predict stock price movements [13]. With the help of topic detection, [4] could improve prediction accuracy. Surprisingly, [8] found that news sentiment from other related companies can predict the prices of other companies.

However, these mentioned studies did not consider the intensity of the news. As is well known in the current digitalization era, the speed of dissemination and the volume of information growth has increased significantly. Especially text form information is verily easy accessible so that investor responses will be faster and prices will adjust immediately [3]. Therefore, it is crucial to research whether news intensity influences the impact of news sentiment on price movements. Due to limited news data on the Indonesian capital market, this study constructed the constituent news of the LQ45 index. As far as our observation, this study is the first to consider news intensity to analyze the impact of constituent index news sentiment on index price movements.

Based on the data we obtained, we found a significant increase in the intensity of LQ45 constituent news. This intensity comes mostly from large-capitalization enterprises. Then in terms of sectoral, finance and mining were the two sectors with the highest intensity. Furthermore, this study evaluates the impact of LQ45 constituent news sentiment on LQ45 index price movements and vice versa in two different intensity conditions. The study found that sentiment had a more significant and longer-lasting impact when news intensity increased through correlation and linear regression analysis.

2. Method

2.1. Data

There are many textual data sources such as online news portals, blogs, and even social media to perform sentiment analysis. However, we limit our search in this study to the Refinitiv Eikon only. Assuming that the news listed on the site has good credibility, thus it provides data reliability. Then, we also limit our sample only to English news, considering the sentiment analysis commonly constructed for English texts. We used all the news available on the site, which can be accessed directly on Refinitiv Eikon or need to be traced to third-party websites.

In this study, we analyze the news constituents of the LQ45 index. We consider the LQ45 index because it is an index that contains 45 groups of stocks with the highest capitalization and liquidity. Since there is an index rebalancing made by Indonesia Stock Exchange (IDX) every six months, we used 61 enterprises' news data that have been in the LQ45 index for five years, from January 2017 to October 2021 (see appendix 1). We cannot start the sample before 2017 because of Refinitiv Eikon's data unavailability. From the 61 enterprises news data we obtained, 30 enterprises have been in the index for at least four years, 20 enterprises have been in the index for at least two years, and 11 enterprises have been in the index for at least one year. Then, we manually deleted unrelated news to the specific enterprise in the news collection process. Occasionally, unrelated news occurred because of the similarity of the ticker or enterprise names. We obtained a total of 2964 news. Subsequently, we perform preprocessing procedure to remove all punctuation marks, numbers, and non-English words.

After that, we build the LQ45 index constituent news by combining each enterprise news published when the enterprise was in the LQ45 index. Due to the limited data we obtained, we conducted a monthly time frame analysis by combining all constituent news published in the same month. Lastly, we obtained LQ45 index data from yahoo finance.

The discussion of this object is interesting to design the media, because the community needs this media in order to provide information to the surrounding community. Many activities have been carried out with the surrounding community, but after seeing the results of the research, the community did not know that the Jatiwangi Art Factory community held the event. The local community is expected to get to know this community because basically this community discusses the Jatiwangi region so that the surrounding community can appreciate the regional culture and can also join the community.

2.2. Methodology

The empirical analysis in this study analyses the influence of news intensity on the relationship between LQ45 index constituent news sentiment and its index movement. Sentiment analysis intends to extract opinions and nuances carried in a text. In conducting sentiment analysis, there are two approaches, the first is to use lexicon-based (dictionaries), and the second is to use machine learning [3]. In this study, we use two lexicon-based sentiment analysis approaches. Firstly, updated financial-specific sentiment dictionary from [7]. The establishment of this dictionary is based on the fact that [6] found that 73.8% of negative words from the Harvard Sentiment Dictionary are not negative in a financial context. The sentiment dictionary is formed based on 10-K company reports on financial performance. Using this dictionary is assumed to describe sentiments in standard financial official documents. This dictionary consists of 354 positive words and 2396 negative words. In this study, the sentiment score based on this method is named financial sentiment, denoted by LMT (Loughran and McDonald tone). LMT value can be calculated as shown in equation (1). We tend to divide the relative number of sentiment words by the total words in the news like [10] to consider the size of the news instead of merely dividing with the number of positive and negative words.

$$LMT = \frac{number\ of\ positif\ words - number\ of\ negatif\ words}{total\ word\ in\ the\ news} \quad (1)$$

However, [16] showed that Loughran and Mcdoland's dictionaries were less effective in handling informal texts. Therefore, this study also considers using TextBlob sentiment analysis, a more general sentiment analysis. For example, [2] used TextBlob to analyze public sentiment on Covid-19, [11] on the Covid-19 vaccine, and [15] analyzed the sentiment of public reactions to scientific research. Then, the sentiment score based on this method is named general sentiment, denoted by TBP (TextBlob Polarity) which the value will range between [-1, 1].

Furthermore, we found that the news's intensity increased after the covid-19 case based on our obtained data. Thus we conduct two intensity condition analyses, pre-Covid-19 case, and post-Covid-19 case. We used Pearson correlation to find the relationship between news sentiment and the LQ45 index movement like [10].The usage of Pearson correlation is intended to measure the linear relationship between two data, which the following equation can calculate its coefficient:

$$\rho(x,y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (2)$$

The ρ value will range between [-1, 1]. When it is positive (negative), the two data move in the same direction (opposite direction)—the greater the coefficient value, the stronger the linear relationship. Meanwhile, there is no linear relationship between two data when the value is zero. Therefore, we use Pearson correlation to determine how strong the relationship between two intensity conditions.

Finally, we use a simple linear regression equation with OLS estimation to determine the direct relationship between news sentiment and the LQ45 index movement. In this study, we do not assume any relationship direction. Thus, we alternately estimated news sentiment and LQ45 index price as endogenous and exogenous variables in two intensity conditions. Then, we also estimate each lagged variable value to determine the effect lasts. Each estimation is done independently. The linear regression equation can be as follows:

$$y = c + \beta x + \varepsilon \quad (3)$$

Where y is an endogenous variable, x is an exogenous variable, and β is the coefficient of the relationship between x towards y, then ε is the term error of the linear regression.

3. Results and Discussion

This section tries to analyze the sources of increasing news intensity. Then we see a linear relationship between news sentiment and the price of the LQ45 index. Finally, we try to see whether sentiment impacts the LQ45 index and vice versa.

3.1. News intensity and its source

Figure 1 shows the news intensity from January 2017 to October 2021. It shows that news intensity has decreased twice in Q3 2017 and Q1 2020. This phenomenon coincided with two crises that occurred.

The first is the Turkish crisis, and the second is the crisis caused by the Covid-19 case. However, what is interesting is that post-Covid-19, the intensity of news has increased significantly.

Figure 2 shows the news intensity of the LQ45 index constituents after classification based on capitalization size. We divided into three classes of capitalization: an enterprise with high, medium, and low capitalization, which 33%, 33%, and 34% proportion respectively from 61 enterprises. In Figure 2, it can be seen that the high capitalization contributed to the highest news intensity, followed by medium and low capitalization. In comparison, the intensity of news on medium and low capitalization is not too different. Nevertheless, it can be concluded that the higher the market capitalization, the higher the intensity of the news.

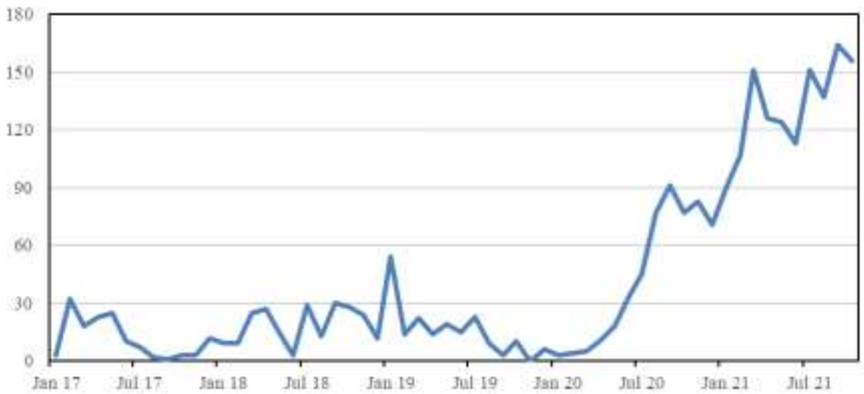

Figure 1. Monthly number of news in constituent LQ45

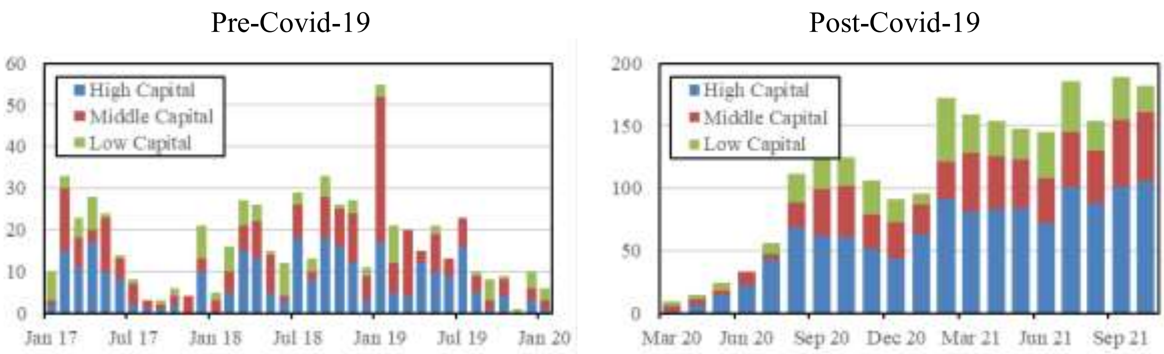

Figure 2. Monthly constituent LQ45 news intensity based on the capital classification

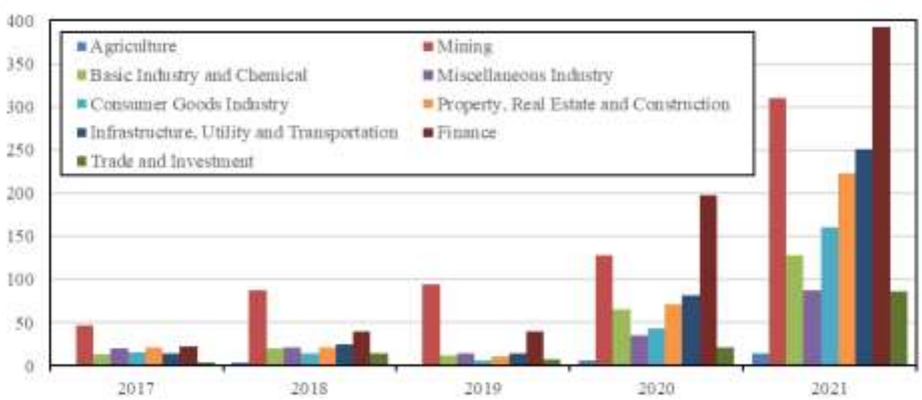

Figure 3. Constituent LQ45 news intensity based on sector classification

ISSN: 2830-0637

Furthermore, figure 3 shows the news intensity of the LQ45 index constituents classified by sector. In order to classify sectors, this study refers to the Jakarta Stock Industrial Classification (JASICA) by IDX. In figure 3, it can be seen that the news intensity of all sectors generally increases except in 2019. In addition, if we look in more detail, the pattern of increasing intensity of each issuer is approximately the same for each year. However, there is an attractive phenomenon where the mining sector had the highest intensity from 2017 to 2019. Then, from 2020 to 2021, the financial sector has the highest intensity. These intensity shifts indicate that investors' attention to the financial sector increased due to the emergence of digital banking issues.

3.2. Constituent news sentiment and price movement

We calculated the sentiment score according to the methodology described in section 3. This study uses two sentiment measurements that are financial sentiment (LMT) and general sentiment (TBP). Figure 4 shows the movement of news sentiment value from January 2017 to October 2021, which is juxtaposed with the price movement of the LQ45 index. This figure shows that LMT and TBP can be assumed to have joint movement with the LQ45 index price.

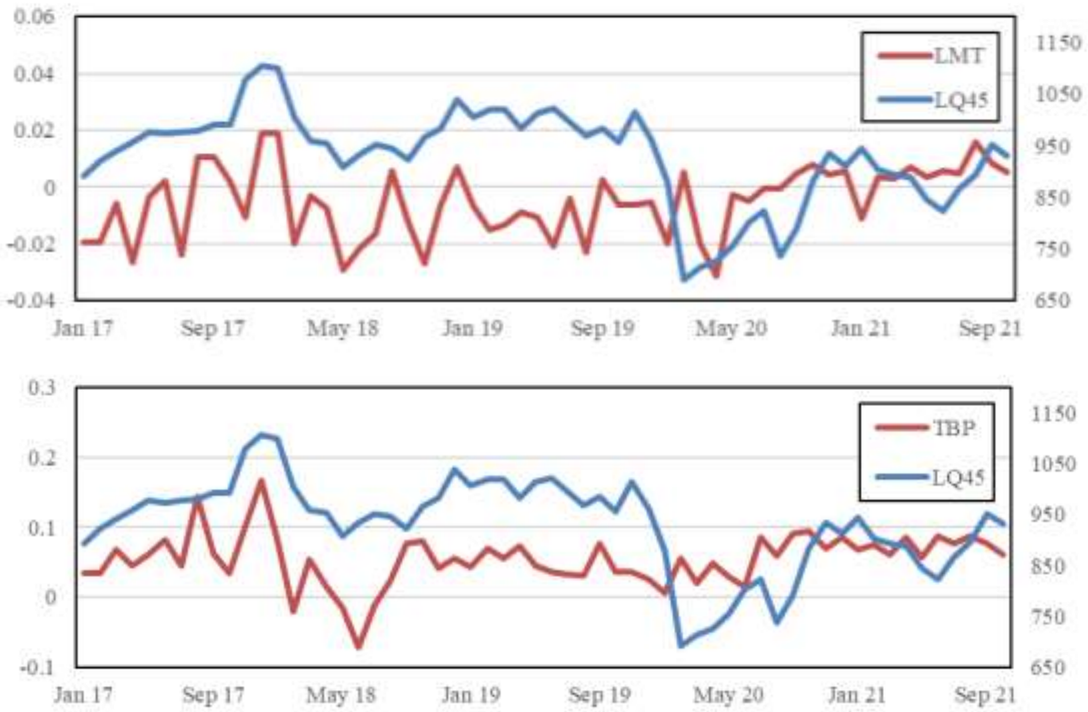

Figure 4. Financial sentiment, general sentiment, and index LQ45 price movement

3.2.1. Correlation between LQ45 index and constituent news sentiment

Table 1 shows the correlation between the LQ45 index and the return of the LQ45 index (r) towards news sentiment and its lag in the pre-Covid-19 period, which is known to have less news intensity. The table shows that the correlation between LQ45 and sentiment is quite significant, indicating that the sentiment indicator can be an indicator to predict index price movements. The correlation between LQ45 and sentiment is generally more significant at $t-1$ than at $t+1$. Thus, it indicates that sentiment can be an indicator that explains changes in index prices and determines future index price movements. However, not vice versa, the LQ45 index does not significantly influence determining sentiment in the future. Furthermore, if we see that return has a more significant correlation with sentiment at $t+1$ than sentiment at $t-1$. It means the return value can explain or influence changes in sentiment in the future.

Table 1. Correlation between the index and return LQ45 to the sentiment along with its lag (pre-Covid-19)

	LMT(-1)	LMT	LMT(+1)	TBP(-1)	TBP	TBP(+1)
LQ45	0.338	0.498	0.382	0.532	0.514	0.270
r	-0.276	0.087	0.349	0.008	0.288	0.308

Note: notation (-1) means $t-1$, and (+1) means $t+1$

Table 2 shows the correlation between the LQ45 index and the return of the LQ45 index (r) towards news sentiment and its lag in the post-Covid-19 period when the news was more intense. The table shows that the correlation value between LQ45 and sentiment is more significant than in the pre-Covid-19 period. In addition, the correlation value of LQ45 with sentiment at $t-1$ appears to be greater than at t and $t+1$. The correlation value shows that the sentiment value can be helpful as a leading indicator of LQ45. The high correlation between LQ45 and sentiment at time t means that sentiment can explain changes in the price of LQ45. However, since the correlation between LQ45 and $t+1$ sentiment is not too significant, LQ45 does not significantly affect sentiment in the future.

Table 2. Correlation between the index and return LQ45 to the sentiment along with its lag (post-Covid-19)

	LMT(-1)	LMT	LMT(+1)	TBP(-1)	TBP	TBP(+1)
LQ45	0.678	0.506	0.209	0.596	0.458	0.266
r	0.068	0.043	0.146	-0.042	0.064	-0.101

Note: notation (-1) means $t-1$, and (+1) means $t+1$

Furthermore, the correlation value between return and sentiment at time $t-1$ and t is relatively small but quite significant at time $t+1$, which means that the sentiment value at the next time has a linear relationship between return sentiment. However, this relationship is less significant than in the pre-Covid-19 case.

3.2.2. Impact constituent news sentiment toward the LQ45 index

In the previous section, it was shown that the LQ45 index correlates with the sentiment. In this section, we employed a linear regression model to determine the impact of sentiment on the LQ45 index price movement. We use the logarithm of the LQ45 index as the endogenous variable (y) since its value is relatively large compared to the sentiment value. Then the exogenous variable (x) is the sentiment value. In addition, we also estimated lagged sentiment individually to find out to what extent sentiment still influences changes in the price of the LQ45 index.

Table 3. Financial sentiments coefficient on Linear regression with LQ45 index as an endogenous variable for pre-Covid-19 case

LMT	LMT(-1)	LMT(-2)
2.145***	1.15*	0.907

Note: We estimated each lag of financial sentiment individually. Notation (-1) means $t-1$, and (+1) means $t+1$. Then *, **, *** denote 10%, 5%, 1% level of significance respectively.

Table 4. Financial sentiments coefficient on Linear regression with LQ45 index as an endogenous variable for post-Covid-19 case

LMT	LMT(-1)	LMT(-2)	LMT(-3)	LMT(-4)	LMT(-5)	LMT(-6)
5.372***	4.932***	4.499**	3.95**	5.396***	5.043***	1.046

Note: We estimated each lag of financial sentiment individually. Notation (-1) means $t-1$, and (+1) means $t+1$. Then *, **, *** denote 10%, 5%, 1% level of significance respectively.

Table 3 shows the financial sentiment coefficient towards the LQ45 index in the pre-Covid-19 period. The table shows that financial sentiment has a significant positive impact on the LQ45 index price up to lag $t-1$. Then, table 4 shows the financial sentiment coefficient towards the LQ45 index in the post-Covid-19 period. The table shows that financial sentiment has a significant positive impact on the movement of the LQ45 index up to lag $t-5$. Besides that, the post-Covid-19 estimated coefficient is more significant than the pre-Covid-19 case. The estimation results indicate that when news intensity is higher in the post-Covid-19 period, news sentiment has a more significant and longer-lasting impact on the LQ45 index.

Table 5 shows the general sentiment coefficient toward the LQ45 index in the pre-Covid-19 period. The table shows that general sentiment has a significant positive impact on the LQ45 index price up to lag $t-3$. Then, table 6 shows the financial sentiment coefficient towards the LQ45 index in the post-Covid-19 period. The table shows that general sentiment has a significant positive impact on the movement of the LQ45 index up to lag $t-5$. Besides that, the post-Covid-19 estimated coefficient is more significant than the pre-Covid-19 case. The estimation results indicate that when news intensity is higher in the post-Covid-19 period, news sentiment has a more significant and longer-lasting impact on the LQ45 index.

Table 5. General sentiments coefficient on Linear regression with LQ45 index as an endogenous variable for pre-Covid-19 case

TBP	TBP(-1)	TBP(-2)	TBP(-3)	TBP(-4)
0.617***	0.602***	0.468***	0.452**	0.218

Note: We estimated each lag of general sentiment individually. Notation (-1) means $t-1$, and (+1) means $t+1$. Then *, **, *** denote 10%, 5%, 1% level of significance respectively.

Table 6. General sentiment coefficients on Linear regression with LQ45 index as an endogenous variable for post-Covid-19 case

TBP	TBP(-1)	TBP(-2)	TBP(-3)	TBP(-4)	TBP(-5)	TBP(-6)
2.143**	2.242***	2.561***	2.733***	2.182***	1.743**	0.848

Note: We estimated each lag of general sentiment individually. Notation (-1) means $t-1$, and (+1) means $t+1$. Then *, **, *** denote 10%, 5%, 1% level of significance respectively.

4.2.3. Impact of LQ45 Index on constituent news sentiment

After analyzing the impact of sentiment on the price movement of the LQ45 index, in this section, we try to explore the impact of the price movement of the LQ45 index on the value of news sentiment. In order to measure this impact, we use a linear regression equation with sentiment as the endogenous variable (y) and the logarithm of the LQ45 index as the exogenous equation (x). Furthermore, we also regress the lagged of the LQ45 index individually to find out to what extent the LQ45 index price movement still affects sentiment changes.

ISSN: 2830-0637

Table 7. Index LQ45 coefficient on Linear regression with general sentiment as an endogenous variable

Pre Covid-19			Post Covid-19		
LQ45	LQ45(-1)	LQ45(-2)	LQ45	LQ45(-1)	LQ45(-2)
0.13***	0.10664**	0.028	0.072399***	0.061987***	0.035662

Note: We estimated each lag of Index LQ45 individually. *, **, *** denote 10%, 5%, 1% level of significance respectively. Notation (-1) means $t-1$, and (+1) means $t+1$

Table 7 shows the Index LQ45 coefficient toward the general sentiment. It shows that both before and after the Covid-19 case, the LQ45 index movement had a significant positive impact only until the $t-1$ lag. However, the coefficient value post-Covid19 case is smaller than the pre-Covid-19 case, indicating that the impact is less significant. Furthermore, Table 8 shows the Index LQ45 coefficient toward the financial sentiment. The table shows that in the case of pre-Covid-19, the LQ45 index only had a significant impact on changes in financial sentiment up to time t. However, the LQ45 index has a significant positive impact up to a lag $t-1$ in the post-Covid-19 period. So, the coefficient value post-Covid-19 case is smaller than the pre-Covid-19 case, indicating that the impact is getting smaller.

Table 8. Index LQ45 coefficient on Linear regression with general sentiment as an endogenous variable

Pre Covid-19		Post Covid-19		
LQ45	LQ45(-1)	LQ45	LQ45(-1)	LQ45(-2)
0.435***	0.25	0.139**	0.114**	0.066856

Note: We estimated each lag of Index LQ45 individually. *, **, *** denote 10%, 5%, 1% level of significance respectively. Notation (-1) means $t-1$, and (+1) means $t+1$

4. Conclusion

This study analyzes the relationship between LQ45 index constituent news sentiment and its index movement on two intensity conditions since we found that the intensity of the news increased drastically after Covid-19. Regarding the news intensity source, we found that enterprises with large capitalization accounted for the most significant LQ45 index constituent's news components. The mining sector had the highest intensity until 2019, then the intensity of the financial sector increased drastically from 2020 beyond the mining sector. This intensity switch could indicate an investor interest shift. Furthermore, based on correlation analysis, we found that both financial and general sentiment could be a leading indicator for the price movement of the LQ45 index in two intensity conditions. Then, the return value is a better leading indicator of the news sentiment in the pre-Covid-19 case. Moreover, based on the results of linear regression analysis, we found that news sentiment has a significant positive impact on the price movement of the LQ45 index. It has a more significant and longer-lasting effect when the intensity of the news is higher (post-Covid-19) than when the intensity of the news is less (pre-Covid-19). However, the effect is not the other way around. The LQ45 index had a less significant impact on news sentiment before and after Covid-19.

ISSN: 2830-0637

Appendix

Appendix 1

list of Enterprises in the LQ45 index from January 2017 until October 2021

4 Years			2 years		1 year
ADRO	GGRM	PWON	ACES	LPKR	AALI
AKRA	ICBP	SCMA	ADHI	LPPF	BJBR
ANTM	INCO	SMGR	BRPT	MEDC	BMTR
ASII	INDF	SRIL	BTPS	SMRA	BUMI
BBCA	INTP	TLKM	CPIN	SSMS	ELSA
BBNI	JSMR	UNTR	CTRA	TBIG	LSIP
BBRI	KLBF	UNVR	ERAA	TKIM	MDKA
BBTN	MNCN	WIKA	INDY	TOWR	MIKA
BMRI	PGAS		INKP	TPIA	PPRO
BSDE	PTBA		ITMG	WSKT	WSBP
EXCL	PTPP		JPFA		

References

[1] Allen, D. E., McAleer, M., & Singh, A. K. (2019). Daily market news sentiment and stock prices. *Applied Economics*, *51*(30), 3212-3235.
[2] Chandrasekaran, G., & Hemanth, J. (2022). Deep learning and TextBlob based sentiment analysis for coronavirus (COVID-19) using twitter data. *International Journal on Artificial Intelligence Tools*, *31*(01), 2250011.
[3] Gupta, A., Dengre, V., Kheruwala, H. A., & Shah, M. (2020). Comprehensive review of text-mining applications in finance. *Financial Innovation*, *6*(1), 1-25.
[4] Hajek, P., & Barushka, A. (2018, September). Integrating sentiment analysis and topic detection in financial news for stock movement prediction. In *Proceedings of the 2nd International Conference on Business and Information Management* (pp. 158-162).
[5] Ho, K. Y., & Wang, W. W. (2016). Predicting stock price movements with news sentiment: An artificial neural network approach. In *Artificial Neural Network Modelling* (pp. 395-403). Springer, Cham.
[6] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10- Ks. *The Journal of finance*, *66*(1), 35-65.
[7] Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187-1230.
[8] Nam, K., & Seong, N. (2019). Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. *Decision Support Systems*, *117*, 100-112.
[9] Nofsinger, J. R. (2005). Social mood and financial economics. *The Journal of Behavioral Finance*, *6*(3), 144-160.
[10] Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, *127*, 104119.
[11] Paliwal, S., Parveen, S., Afshar Alam, M., & Ahmed, J. (2022). Sentiment Analysis of COVID-19 Vaccine Rollout in India. In *ICT Systems and Sustainability* (pp. 21-33). Springer, Singapore.
[12] Pyo, D. J., & Kim, J. (2021). News media sentiment and asset prices in Korea: text-mining

ISSN: 2830-0637

approach. *Asia-Pacific Journal of Accounting & Economics*, *28*(2), 183-205.

[13] Schumaker, R. P., Zhang, Y., Huang, C. N., & Chen, H. (2012). Financial fraud detection using vocal, linguistic and financial cues. *Decis Support Syst*, *53*, 458-464.

[14] Seng, J. L., Huang, K. Y., & Yang, H. F. (2017, April). Financial Reports and Financial News—An Information Content Gap Analysis. In *Asian Conference on Intelligent Information and Database Systems* (pp. 333-342). Springer, Cham.

[15] Shahzad, M., & Alhoori, H. (2022). Public Reaction to Scientific Research via Twitter Sentiment Prediction. *Journal of Data and Information Science*, *7*(1), 97-124.

[16] Tabari, N., & Hadzikadic, M. (2019). Context sensitive sentiment analysis of financial tweets: a new dictionary. In *Intelligent Methods and Big Data in Industrial Applications* (pp. 367-374). Springer, Cham.