ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business
Specialization: Financial Economics

The Power of Media: News Sentiment and Stock Market Returns

Author:	Thomas Dani Yoe
Student number:	595730
Thesis supervisor:	Dr. Ruben de Bliek
Second reader:	Dr. C.S. Howard
Finish date:	9 June 2024

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second reader, Erasmus School of Economics or Erasmus University Rotterdam.

ABSTRACT

In this thesis, I analysed the impact and forecasting capabilities of news sentiment on the daily returns of the Nikkei 225 index. With data related to the Nikkei 225 index from 2018 to 2024, I used linear regressions and forecasting models to analyse the relationship between news sentiment and returns. My results show that there is a positive relationship between overall news sentiment and daily returns of the Nikkei 225 index. This relationship was strengthened with more news coverage, and negative sentiment seemed to impact daily returns more than positive sentiment. These results, however, are dependent on the time period analysed as the COVID-19 pandemic may have influenced these results. Finally, my out-of-sample analysis shows that news sentiment significantly improves the information set of a benchmark model and offers predictive power beyond looking at past returns. Robustness tests, however, show that these findings are sensitive to using open-to-open prices to calculate daily returns.

TABLE OF CONTENTS

CHAPTER 1 Introduction

The efficient market hypothesis (EMH) claims that all investors have perfect information and that their actions reflect complete rationality (Hill, 2020). The advent of behavioural finance, however, challenges this theory by emphasising the role of irrational behaviours, such as emotion and personal sentiment, in the financial decision-making process (Tan & Taş, 2020). In reality, the prevalence of inefficiencies, such as biases and information asymmetry, cause the opinions of others to play a significant role in the behaviour of financial markets (Chen et al., 2014). News has been found to be a significant factor influencing the changes in financial markets (Du, 2020). Previous studies related to this topic have found disproportionate responses to positive and negative news about the market as a result of biases and overreactions. These results align less with the efficient market hypothesis and more with the realm of behavioural finance. Recent literature in behavioural finance focuses on analysing the impact of various sources of financial news on the behaviour of financial securities. A group of these papers analyse financial news article sentiment and its effect on stock market index behaviour. Similarly, I will assess the sentiment and quantity of news on a stock market index and subsequently analyse its impact and forecasting capabilities of the daily returns of that index.

Similar studies have been conducted using different sources of news sentiment on various stock indices. A group of studies have utilised the growing use of social media to analyse sentiment related to different financial securities. Tan and Taş (2020), for example, used Twitter activity and sentiment related to specific stocks within the S&P 500, S&P 350 Europe, and S&P Emerging Markets Core as metrics to track information diffusion and analyse its impact on trading volume and returns. Through regression equations, they found that Twitter activity and sentiment were associated with higher stock returns and trading volume. Bijl et al. (2016) instead used Google search volume to predict next-day stock returns of companies in the S&P 500. Using a panel data regression with fixed effects, they found that Google search volume was negatively correlated with next day stock returns. While social media provides valuable data, much of the literature focuses on standard financial news articles as sources of public sentiment. The paper by Tetlock (2007) analyses the content of the “Wall Street Journal” in order to predict subsequent prices of the Dow Jones Industrial Index (DJIA) index. Using regression equations, they found that high amounts of media pessimism were correlated with lower market prices in subsequent days. The paper also finds that extreme values of media pessimism were correlated with higher trading volume in the following days. Instead of analysing a specific news article, the paper done by Allen et al. (2019) uses a large sample of different news articles provided by Thomson Reuters News Analytics (TRNA) to create daily sentiment scores for the DJIA index. Using Ordinary Least Square (OLS) and Quantile Regressions (QR), the paper finds that the average daily news sentiment score had a significant positive correlation with DJIA returns.

Although the relationship between standard news article sentiment and stock index behaviour has been studied before, most papers analyse it in the context of the U.S. or European market. While this does provide valuable insights, the Asian market remains relatively understudied. With the Nikkei 225 being a leading Asian stock index, I will analyse the impact of financial news article sentiment and activity on the returns of the Nikkei 225 index (WSJ, n.d.). In the context of Japan, similar studies have been performed that analyse the impact of sentiment on Japanese stocks. The paper done by Ishijima et al. (2015), for example, looks at daily articles from the newspaper "Nikkei" between 2007 and 2012 and counts the number of positive and negative words used to describe the economic situation at the time. The findings of this paper indicate that these words have predictive power of up to three days of stock prices on the Tokyo Stock Exchange. My goal is to analyse sentiment and activity from various news articles using more technologically sophisticated resources and analyse its impact and forecasting capabilities of the returns of the Nikkei 225 index. After reviewing previous literature, the following research question can be formed:

"How and to what extent does news article activity and sentiment contain explanatory and forecasting capabilities for Nikkei 225 Index returns?"

To examine this question, I collected data regarding news sentiment and activity associated with the Nikkei 225 Index from the Ravenpack Global Macro Data Set (Wharton Research Data Services, n.d.-b). The daily returns, as calculated similarly to Tan and Taş (2020) using open-to-open prices, were obtained from Bloomberg. This analysis was conducted based on data between 2018 and 2024. I began by addressing problems of stationarity within my data and resolving potential econometrics violations, such as heteroskedasticity and autocorrelation. In the next phase, the explanatory power of news article sentiment and activity on the returns of the Nikkei 225 index was assessed. This was done using a simple linear regression controlling for volatility, trading volume, past returns, and the market capitalisation of the Nikkei 225 index (all obtained from Bloomberg). I then extended these baseline regression equations to test for the interaction effect between news sentiment and activity. I also extended the baseline regression model to differentiate between positive and negative sentiment using dummy variables. In the last phase, the forecasting capabilities of news sentiment on the returns of the Nikkei 225 Index was analysed using two Autoregressive Moving Average with Exogenous Variables (ARMAX) models. One of these models included lags of news sentiment values as a variable, and one did not. Initially, the suitable lag length for the ARMAX models was determined using BIC values. Afterwards, an evaluation of both models was conducted using an in-sample and out-of-sample analysis by observing the subsequent BIC, AIC, MSE, MAE, and MAPE values.

My hypothesis was that news sentiment values were positively correlated with higher returns of the Nikkei 225 Index and that these values can also be used to forecast future returns more accurately. This would be apparent from the positive and statistically significant value of the news sentiment variable's coefficient in the regression equations and the improvement of forecasting models when including lags of news sentiment values. My research provides a new and extended understanding of the relationship between news and a stock market index within a Japanese context. This research, therefore, provides new insights into the relatively understudied Asian region. However, given the uniqueness of different regions and the various methods of assessing public sentiment, there will be sufficient variance that would remain unexplained.

My results show that there is a positive relationship between overall news sentiment and daily returns of the Nikkei 225 index. This is in line with previous studies and indicates that this relationship could be present when analysing other Asian indices as well. Additionally, negative sentiment seemed to impact the daily returns of the index more than positive sentiment. This is also in line with previous studies and can be linked to the risk-averse behaviour of investors (Chen et al., 2003). Furthermore, my analysis shows that the impact of news sentiment on the returns of the Nikkei 225 index increases with news coverage, which is in line with theories of limited investor attention (Smales, 2021). Finally, my out-of-sample analysis shows that news sentiment significantly improves the information set of a benchmark model and offers predictive power beyond looking at past returns. Combined with findings from previous studies on other indices, this suggests that the forecasting capabilities of news sentiment could be present in other Asian indices.

The remainder of this research paper is structured as follows. Chapter 2 discusses the previous research that has been done surrounding topics related to my study and introduces the hypotheses that form the basis of this research. Chapter 3 describes the data I used in my research and any transformation of the variables performed. This section also provides the descriptive statistics associated with each variable. Chapter 4 explains the methodology I follow to perform my research and answer my hypotheses. Chapter 5 displays and discusses the results of my analysis and relates them to previous literature and my hypotheses. The last chapter concludes this paper and introduces the limitations of this research paper and recommendations for follow-up research. Additional supportive materials are provided in the Appendix.

CHAPTER 2 Theoretical Framework

Previous studies have investigated various measures of news sentiment and activity on different behaviours of financial securities. This section explores the different studies that have been performed on topics related to my research question and provides the foundation for my analysis. While most research papers can be split between social media and standard news analysis, Japanese markets have proved to be more reactive to standard news compared to social media (Gan et al., 2023). Therefore, my research will focus on standard news article sentiment and activity as opposed to social media. This will serve as a measure of public sentiment, which will be analysed in relation to the performance of the Nikkei 225 index.

2.1 Social Media

Due to the prevalence of social media sentiment studies, it is important to understand the results of these research papers to compare the differences to standard news articles. The majority of studies have focused on using Twitter as a source of media sentiment and analysed its impact on stock returns. The paper by Ranco et al. (2015), for example, analyses Twitter sentiment and activity related to companies in the Dow Jones Industrial Index (DJIA) index and its impact on their stock prices. Using a Granger causality test, they find that there is a significant relationship between Twitter sentiment and stock returns during periods of high Twitter activity. While most papers analyse this relationship in the context of a U.S. market, some research other geographies. The paper by Nofer and Hinz (2015) uses Twitter to assess public mood and uses this to understand its impact on German stock market returns. They find that public mood levels weighted by the number of followers the user has can predict stock returns in future periods. The paper provides the explanation that these results are due to the emotional contagion that occurs when users spread information on social media sites.

A multitude of other social media platforms have also been investigated to measure public sentiment. The paper done by Nguyen et al. (2015), for example, focuses on extracting sentiment categories on 18 different stocks from messages posted on the Yahoo Finance Message Board. Using the sentiment associated with the particular company, they find that the accuracy of stock predicting models increases compared to models that only use historical prices. Another example is the paper done by Li et al. (2019), which used data from one of the largest microblogging platforms in China called “Tencent Weibo” and assessed the correlation between sentiment and the movements of the Hushen 300 index. The paper finds that sentiment obtained from this platform can assist in explaining the short-term fluctuations of the Hushen 300 index.

2.2 Traditional News

Traditional news articles have been the cornerstone of research for many years. Research papers have utilised traditional news sentiment and activity to analyse various relationships across asset classes. A large group of papers have analysed the impact of traditional news articles on the volatility of different assets. The paper done by Jiao et al. (2016) focuses on analysing the effect of news coverage on the volatility and trading volume of stocks traded on the AMEX, NASDAQ, and the NYSE. The results of the paper show that higher news coverage results in a decrease in volatility and trading volume. The paper also compares news coverage with social media coverage and finds opposite results for social media. This paper highlights the stark differences in outcome when comparing social media to traditional news in research. Other papers have analysed news sentiment in the context of initial public offering (IPO) under-pricing. Studies performed by Bajo and Raimondo (2017) assess over 2800 IPOs in the U.S. to understand the effect of the tone of newspaper articles on the under-pricing associated with these IPOs. Through the use of over 27,000 newspaper articles, they observe that when newspaper articles present the IPO in a positive manner, the underpricing of the IPOs are more severe. These results appeared to be stronger when the newspaper is considered more reputable and when the news is reported closer to the date of the IPO. The authors claim that the tone of newspaper articles primarily affects retail investors who rely on second-hand information which alter their beliefs about the IPO.

Other research utilises news sentiment to analyse different asset classes, such as real estate and gold. Ruscheinsky et al. (2018), for example, looks at real estate related news sentiment and its impact on real estate investment trust (REIT) movements. This paper finds a positive correlation between news sentiment and REIT prices with a three-to-four-month lag. They argue that investor opinions take time to change and that their decisions are heavily influenced by their level of optimism about the future. The study done by Smales (2014), on the other hand, looks at the returns of gold futures. Similar to many studies analysing news sentiment, the findings of this paper suggest that the effect of negative news sentiment is greater than the effect of positive news sentiment on the commodity's future prices. They also find that this relationship is strengthened during recessions.

The use of news sentiment and activity to analyse the performance of stock market indices has been performed in many countries using different data sources. Some papers opt to use a single newspaper, while others utilise large data sets that aggregate multiple newspapers from various sources. The highly influential paper done by Tetlock (2007) was one of the first to analyse this relationship. Using daily content of the "Wall Street Journal", the author analyses the relationship between pessimism displayed in the newspaper and DJIA prices. Using vector autoregressions (VAR), the author finds that high levels of media pessimism were correlated with lower subsequent prices of the DJIA index. They also observe that abnormally high or low amounts of pessimism displayed in the newspaper result in higher trading volume. The underlying reason for these results can be linked to the theory of DeLong et al. (1990a), which highlights the existence of "noise" traders who have random beliefs about potential future dividend payments. With the assumption that investors display downward-sloping demand functions for financial assets, this paper posits that the beliefs of "noise" traders regarding possible future dividend payments can affect the price of these financial securities. With pessimism displayed in the news, these "noise" traders encounter a belief shock, which results in them selling these securities and hence increasing trading volume.

Other papers instead use a collection of newspaper articles to study this relationship instead of relying on one. Considering retail investors have a multitude of different sources they can rely on for information, it is plausible to assume that relying on one newspaper could be inaccurate. The paper by Allen et al. (2019b) does this by using daily aggregated news sentiment scores from multiple newspapers and analyses its impact on DJIA prices. Using a sample between 2006 and 2012, the author uses Ordinary Least Square (OLS) and Quantile Regression (QR) to assess the relationship using multi-factor models. The findings of this paper show that the prices of the DJIA are significantly affected by news sentiment scores and that financial news sentiment can be used as an additional factor in multi-factor models such as the Fama-French three-factor model. These results are in line with the paper by Tetlock (2007) by showing that news sentiment is positively correlated with prices of the DJIA index. Considering theories of "noise" traders and emotional contagion align with empirical studies showing that news sentiment is positively correlated with stock prices, it would be plausible to assume that news sentiment from a data base of various sources would similarly be positively correlated with Japanese stock market returns. Therefore, my first hypothesis for this study would be:

Hypothesis 1.1: *Daily news article sentiment scores related to the Nikkei 225 index is positively correlated with same day market returns.*

Several studies utilising both social media and traditional news have highlighted the differences between positive sentiment and negative sentiment with many proving that negative sentiment has a stronger impact on financial security returns than positive sentiment. Papers analysing Twitter sentiment on stock returns, for example, found that negative sentiment has a larger impact on Nikkei 225 index, NYSE, NASDAQ, and AMEX returns than positive sentiment. (Affuso and Lahtinen, 2018; Han et al., 2018). The study done by Smales (2014) similarly finds discrepancies between positive and negative news in the gold futures market. In the case of traditional news, similar results appear consistent. The paper done by Chen et al. (2003), for example, examines how U.S. news impacts the stock returns of major foreign indices, including the Nikkei 225 index, through a GARCH model. Using data from 1985 to 2001, they find that the effect on foreign indices is much larger when negative news is presented compared to positive news. This phenomenon can be attributed to the traditional view that investors are risk-averse and more sensitive to negative news compared to an equal degree of positive news. Based on this, I would expect similar results to hold when analysing news sentiment associated with the Nikkei 225 index. Therefore, my next hypothesis is that:

Hypothesis 1.2: *Negative news sentiment scores inflict a greater effect on daily Nikkei 225 index returns as opposed to positive news sentiment scores.*

In reality, investor attention is limited, and when investors focus more of their attention on a specific event, the price associated with that asset changes more rapidly (Smales, 2021). Therefore, the returns of an asset only changes if investors pay attention to the specific event surrounding the asset (Huberman & Regev, 2001). This can be related to the idea that the effect of sentiment on the returns of a stock market index would be heightened if more news articles were posted about that index. Therefore, I hypothesize that:

Hypothesis 1.3: *The correlation between daily news article sentiment scores and same day Nikkei 225 index returns is strengthened with more news activity associated with the index.*

The predictive capabilities of sentiment scores on stock returns is a widely researched topic. Most papers, however, analyse the relationship in the context of the S&P 500 index. The paper done by Fazlija and Harder (2022), for example, shows that news article sentiment scores can be used to predict the direction of the S&P 500 index. Mohan et al. (2019) found similar results and improved the accuracy of S&P 500 stock price prediction through the use of over 265,000 financial news articles to extract a sentiment score. Several studies have attempted to analyse this relationship within China. Xu et al. (2021), for instance, uses a large database of various newspapers in China and analyses its forecasting abilities of Chinese stocks. Using regression models and comparing both in-sample and out-of-sample data, they find that news sentiment and social media sentiment have significant predictive power of the direction of Chinese stocks. The results displayed that these metrics performed better at forecasting Chinese stock movements than macroeconomic indicators. Ishijima et al. (2015) find similar predictive capabilities of sentiment extracted from the newspaper "Nikkei" on stock returns in the Tokyo Stock Exchange. Given that investor opinions take time to change as a result of news, I forecast that news sentiment would have predictive power of Nikkei 225 index returns (Ruscheinsky et al., 2018). Considering most research papers found short-term predictive capabilities, my last hypothesis is that:

Hypothesis 2: *Daily news article sentiment scores related to the Nikkei 225 index can be used to predict next day market returns.*

My research contributes to the current scientific literature in the following ways. Firstly, I plan on researching the relatively understudied Asian region to enrich the current understanding of news sentiment on stock returns. Secondly, I plan to use a large data set of news articles over many days instead of relying on one newspaper for sentiment as done in previous research (see eg. Ishijima et al., 2015; Tetlock, 2007). This could improve the research's credibility as investors have multiple sources of information instead of just one. Lastly, I plan to add to current literature, such as Du (2020), by also assessing the forecasting capabilities of news sentiment and activity on Nikkei 225 index returns.

CHAPTER 3 Data

To analyse the effect and forecasting capabilities of news sentiment and activity on the returns of the Nikkei 225 index, I utilised different sources to obtain data for my analysis. I focused on a sample between January 2018 and February 2024 for my research and obtained 1493 trading day observations. All these observations were related to daily measures of the Nikkei 225 index as a whole.

3.1 News Sentiment Score and Activity

To get the daily news sentiment score and activity related to the Nikkei 225 index, I used a data source called Ravenpack which is provided by Wharton Research Data Services (Wharton Research Data Services, n.d.-b). This data provider extracts multiple news articles from various news sources and filters out the news related to a specific entity using algorithms that detect whether the specific entity is mentioned in the article, which could be anything such as countries, companies, places, or stock indices. The different news sources include articles from thousands of websites such as Investing.com, Yahoo! News, CNBC, and The Economic Times. After the algorithm matches the news articles to a specific entity, it breaks down the content of the article to identify the different events taking place in the story. As a news article often contains different entities, it also detects the role of the specific entity in the news article, such as the acquirer or the acquiree in an acquisition. The final stage is the consolidation of information, where different news articles on different dates are assigned to a specific entity for which valuable data can be calculated based on the content of the articles. As I am analysing the Nikkei 225 index, I filter out news articles that specifically mention this index in their articles.

Using this, I obtained multiple news articles corresponding to different dates that mention the Nikkei 225 index. For any news article that mentions the index, Ravenpack calculates a relevance score of how relevant the index is in the underlying news story. With a range of scores between 0 and 100, a score of 0 means the index was only passively mentioned in the story, while a score of 100 means that the index was the focus of the story. As values above 75 are considered to be significantly relevant, I filtered out the news articles that had a score below this threshold to remove stories that are not directly relevant to the Nikkei 225 index.

Associated with each news article, I also extract a news sentiment score called the “Composite Sentiment Score”. This score ranges from -1 to 1 and calculates the news sentiment of the story by analysing the emotionally charged words and phrases in the article that experts have identified to have a short-term impact on the stock prices of 100 large-cap stocks. Scores above 0 are considered to be positively phrased news associated with the index, while scores below 0 are considered to be negative signals related to the index. Because a single date often has multiple news articles associated with the index, I calculate the average sentiment score of all these news articles to obtain a daily average news sentiment score per date. This will represent my metric to assess the effect of news sentiment on the returns of the Nikkei 225 index. Figure 1a shows the average daily news sentiment score associated with the Nikkei 225 index over time in a column chart. We see that there is an even spread between positive and negative sentiment scores with negative daily news sentiment scores often having a larger absolute value than positive scores. The descriptive statistics related to this variable are presented in the first row of Table 1.

The news activity is calculated by simply counting the number of articles that correspond to a specific date. It is evident that this variable contains multiple outliers, as is shown in the histogram in Figure A1 in Appendix A. To resolve this, I winsorise this variable at the 1% level to minimize the effect of these outliers. Additionally, as shown by the histogram in Figure B1 in Appendix B, the variable containing news activity presents a significant right-skewed distribution and is resolved by taking the natural logarithm of the winsorised news activity values. This will represent my metric to assess the effect of news activity on the returns of the Nikkei 225 index. The descriptive statistics related to this variable are presented in the second row of Table 1. Figure 1b displays the number of news articles associated with the Nikkei 225 index over time before being winsorised and before adjusting for the natural logarithm of the values. We observe that peaks in values are often associated with key events such as the Nikkei 225 index breaking its record value on the 22nd of February in 2024 and the Japanese economy beating projected growth targets in 2021 (Rennison & Karaian, 2024; Toh, 2021).

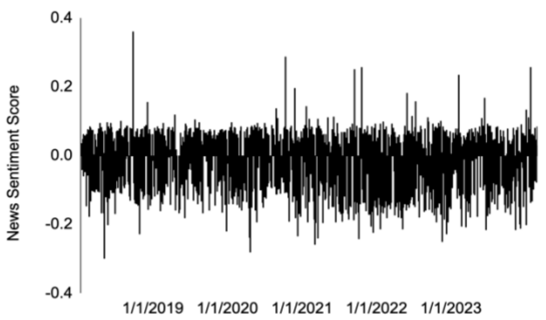

Figure 1a: News Sentiment Score. Column chart representing the daily average news sentiment score of the Nikkei 225 index.

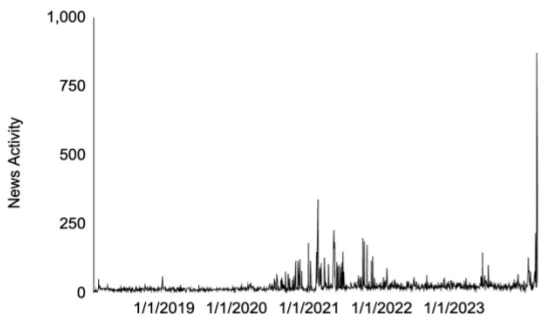

Figure 1b: News Activity. Line chart representing the number of news articles associated with the Nikkei 225 index per day.

3.2 Daily Returns

As I am calculating the returns of the Nikkei 225 index on a daily basis, I will use the open prices to determine the daily price change. The open prices will be obtained from Bloomberg, which provides the opening prices of the Nikkei 225 index every trading day. Using these prices, I follow the method by Tan and Taş (2020) to calculate the daily returns, as this paper similarly assesses the effect of sentiment on daily stock returns. The returns will be calculated by taking the percentage change of the opening prices between day t and day t + 1. This results in the following formula to calculate returns:

$$Daily\ Returns_t = \frac{Open_{t+1} - Open_t}{Open_t}$$

After calculating the daily returns, it is shown on the histogram in Figure A2 in Appendix A, that the returns variable contains outliers. I, therefore, winsorise this variable at the 1% level to minimize the effect of these outliers. These daily returns are then matched to the sentiment score and activity level of the same date. The descriptive statistics related to this variable are presented in the third row of Table 1.

3.3 Control Variables

In order to limit the effect of endogeneity, I will introduce control variables to better understand the relationship between news sentiment and activity on the daily returns of the Nikkei 225 index. Similar to the data for daily returns, I will obtain all the data required for my control variables from Bloomberg on a daily level. One common control variable used in the literature is the daily returns of previous periods (see eg. Dong and Gil-Bazo, 2020; Tetlock, 2007). I aim to follow a similar method by summing up the returns in the previous five trading days as my control variable. The returns are calculated by assessing the change in opening prices of the Nikkei 225 index on a daily level. From Figure A4 in Appendix A, however, we see that there are slight outliers and I, therefore, winsorise this variable at the 1% level. This new variable will represent my control variable for previous returns. The descriptive statistics related to this variable are presented in the fourth row of Table 1.

Another common control variable used in the literature is the volatility of the index. The paper done by Gan et al. (2023), for example, also uses volatility as a control variable when assessing the effect of news sentiment on stock index returns. Bloomberg provides a 10-day average volatility on a daily basis, for which I will calculate the average of for the previous five trading days. This average will be used as my control variable for the Nikkei 225 index. From Figure A3 in Appendix A, however, we see that there are slight outliers and I, therefore, winsorise this variable at the 1% level. Additionally, from Figure B2 in Appendix B, we see that this variable presents a significant right-skewed distribution and is resolved by taking the natural logarithm of the winsorised average volatility values. This new variable will be used as my control variable for the volatility of the Nikkei 225 index. The descriptive statistics related to this variable are presented in the fifth row of Table 1.

The third control variable commonly used in the literature is the size of the stock index. The data will be obtained from Bloomberg which provides the market capitalisation of the Nikkei 225 index on a daily level. I will then take the natural logarithm of the previous day's market capitalisation as a control variable. This will represent my control variable for the size (market capitalisation) of the Nikkei 225 index. The descriptive statistics related to this variable are presented in the sixth row of Table 1.

The last control variable I will use in my analysis is the trading volume of the index. Several papers have used the abnormal trading volume (abnormal turnover) as a control variable when analysing the impact of news sentiment on stock returns (see eg. Dong and Gil-Bazo, 2020; Gan et al., 2023; Tetlock, 2007). Bloomberg provides the trading volume of the Nikkei 225 index on a daily basis, which I will use to calculate the abnormal trading volume. I follow a method to calculate abnormal trading volume similar to Tan and Taş (2020) by subtracting the average natural logarithm of the previous five-day trading volumes from the natural logarithm of the current day's trading volume. This results in the following formula to calculate the abnormal trading volume:

$$Abnormal\ Trading\ Volume = \ Ln(Trading\ Volume_t) - \frac{\sum_{i=t-5}^{t-1} Ln(Trading\ Volume_i)}{5}$$

As my control variable I will use the abnormal trading volume of the previous day instead of the current day as similarly done in previous literature. The descriptive statistics related to this variable are presented in the seventh row of Table 1.

3.4 Testing for Stationarity

Before conducting the analysis, I made sure to test for the stationarity of the variables. This would otherwise lead to spurious regressions and incorrect standard errors. To do so, I will use a Dickey-Fuller Test for which the null hypothesis assumes a unit root. If this null hypothesis is rejected, we can assume that the variable is stationary and can be used in my regression analyses. I first plotted the time series for each variable to understand how the data behaves. This is done to understand what kind of Dickey-Fuller Test to perform on the variables. Plotting the time series, I observe that the daily returns, news sentiment score, previous returns, and the previous day's abnormal turnover variable all display a horizontal trend with a mean of 0. For these variables, I will perform a Dickey-Fuller Test with no constant and no trend. The variables representing the natural logarithm of the previous day's market capitalisation and the natural logarithm of the number of articles display a trending graph with a constant. For these variables, I will perform a Dickey-Fuller Test with a constant and a trend. The variable representing the natural logarithm of the average volatility displayed a horizontal trend with a constant other than 0. For this variable, I performed a Dickey-Fuller Test with a constant but no trend.

The results of the Dickey-Fuller Tests are displayed in Table C1 in Appendix C. We see that all variables appear to be stationary at the 1% level except the variable that represents the natural logarithm of the previous day's market capitalisation. This indicates that this variable is non-stationary, and to resolve this, I will take the first difference of the variable and use this in my regression analyses. The last row in Table C1 in Appendix C shows the Dickey-Fuller Test results of the differenced value of the natural logarithm of the previous day's market capitalisation. We now see that non-stationarity has been resolved, and I will use this new variable in my analysis. The descriptive statistics for this variable are presented in the eighth row of Table 1.

Variable Name	Number of Observations	Mean	Median	Standard Deviation	Min	Max
Sentiment	1,493	-0.0189	0.0026	0.0895	-0.3000	0.3600
Ln_Activity	1,493	2.9322	2.8904	0.5700	1.7918	4.8978
Returns	1,493	0.0003	0.0004	0.0105	-0.0312	0.0255
Previous_Returns	1,493	0.0021	0.0037	0.0253	-0.0733	0.0682
Ln_AvgVolatility	1,493	2.8000	2.8095	0.3718	1.9482	3.7901
Ln_MarketCap	1,493	33.6849	33.6706	0.1787	33.2697	34.1781
AbnTurn	1,493	0.0005	-0.0187	0.2014	-0.5929	0.8761
D_Ln_MarketCap	1,493	0.0003	0.0004	0.0112	-0.0567	0.0736

Table 1: Descriptive Statistics. This table presents the descriptive statistics of the two independent variables (news sentiment & news activity), the dependent variable (daily returns), and the control variables.

CHAPTER 4 Method

4.1 Initial Analysis

To get an initial understanding of the relationship between news sentiment and activity on the daily returns of the Nikkei 225 index, I first correlate the variables against each other to see how the independent variables interact with the dependent variable and how the variables interact with each other to check for any initial signs of multicollinearity. I also assess scatter plots between the independent variables and the dependent variable to further examine initial relationships between the variables. This will also allow me to check for the linearity of the data.

4.2 Contemporaneous Regressions

To get a thorough understanding of the impact of news sentiment and activity on the daily returns of the Nikkei 225 index, I will use linear regression models. The dependent variable representing the daily return on day t, is indicated as $Returns_t$. The independent variables, daily news sentiment and the natural logarithm of the number of articles are indicated as $Sentiment_t$ and $Ln\_Activity_t$, respectively. The four control variables representing the cumulative past returns, the natural logarithm of the average volatility, the differenced natural logarithm of the previous day's market capitalisation, and the previous day's abnormal turnover are indicated as $Previous\_Returns_t$, $Ln\_AvgVolatility_t$, $D\_Ln\_MarketCap_{t-1}$, and $AbnTurn_{t-1}$, respectively. These control variables will be grouped into $Controls_t$ in the equations for simplicity.

I will perform three linear regressions to answer my hypotheses. The first regression equation includes the daily sentiment score and the natural logarithm of the number of articles as independent variables. These variables will allow me to understand how news sentiment and activity affect the daily returns of the Nikkei 225 index. This regression aims at answering hypothesis 1.1, and the equation for this regression is shown in equation (1) below:

$$Returns_t = \beta_0 + \beta_1 Sentiment_t + \beta_2 Ln\_Activity_t + \sum_{k=1}^{K} \beta_{3k} Controls_{kt} + \varepsilon_t \tag{1}$$

The next regression equation aims to answer hypothesis 1.2 by splitting the sentiment score between positive and negative sentiment using dummy variables. After assigning a dummy variable to both positive and negative sentiment scores, I create interaction variables to see how positive and negative sentiment affect daily returns separately. The variables representing the effect of positive and negative sentiment on day t on the daily returns of the Nikkei 225 index are indicated as $Positive\_Sentiment_t$ and $Negative\_Sentiment_t$, respectively. The equation for this regression is shown in equation (2) below:

$$\begin{aligned} Returns_t = \beta_0 + \beta_1 Positive\_Sentiment_t + \beta_2 Negative\_Sentiment_t \\ + \beta_3 Ln\_Activity_t + \sum_{k=1}^{K} \beta_{4k} Controls_{kt} + \varepsilon_t \end{aligned} \tag{2}$$

The final regression equation aims to answer hypothesis 1.3 by showing the interaction effect between news sentiment and activity on the daily returns of the Nikkei 225 index. By creating an interaction variable, I can analyse whether the impact of news sentiment on the daily returns of the index increases when there is more news activity through the variable, $Interaction_t$. The equation for this regression is shown in equation (3) below:

$$\begin{aligned} Returns_t = \beta_0 + \beta_1 Sentiment_t + \beta_2 Ln\_Activity_t + \beta_3 Interaction_t \\ + \sum_{k=1}^{K} \beta_{4k} Controls_{kt} + \varepsilon_t \end{aligned} \tag{3}$$

To ensure the Ordinary Least Squares (OLS) assumptions are not violated, I will test whether all the assumptions are met before and after performing every regression. Heteroskedasticity will be tested using a White test, correlated errors will be tested using a Breusch-Godfery test, multicollinearity will be tested using a variance inflation factor (VIF) test, and the normality of errors will be tested by visually observing the histogram of the residuals.

4.3 First Robustness Test

As the dataset used in the regression analysis covers the period of the COVID-19 pandemic, it would be reasonable to analyse how the pandemic altered the results of my regressions and whether my results would still hold in the separate periods before, during, and after the pandemic. Several papers have analysed the impact of news sentiment on stock returns during the COVID-19 pandemic and found that news sentiment remains to have a significant impact on returns during this pandemic (see eg. Gherghina et al., 2023). As most papers classify the years 2020 and 2021 as the period of the pandemic, I will split my dataset into three periods and run my regression analyses within each (see eg. Huynh et al., 2021). The three periods will be before the pandemic (2018 - 2019), during the pandemic (2020 - 2021), and after the pandemic (2022 - 2024). This will further assess whether my results are robust to different time periods by observing changes in the coefficients, significance levels, and R^2 values.

4.4 Predictive Analysis

My second hypothesis assesses the forecasting capabilities of news sentiment on the daily returns of the Nikkei 225 index. In order to examine whether this is the case, I will compare the forecasting accuracy of a model including lags of news sentiment scores to the accuracy of a benchmark model that does not include lags of news sentiment scores. This will be performed in both an in-sample dataset and an out-of-sample dataset. Through this, I will be able to check whether the forecasting of future values can be done by training a model through historical data. The in-sample dataset consists of data between the start of 2018 and the end of 2021, while the out-of-sample dataset starts in 2022 and ends in February 2024. Starting with the in-sample analysis, I will train the models within the in-sample dataset and test the accuracy of both models in the same period. For the out-of-sample analysis, I will train the models in the in-sample dataset and test the forecasting capabilities of both models in the out-of-sample dataset by comparing the fitted and actual values. This will allow me to see which model performs better at predicting values after training them using historical data.

4.4.1 In-Sample Analysis

In order to create a forecasting model, I built a model that allows me to match the autocovariance and autocorrelation structure of the data. Therefore, I begin with the in-sample dataset for which the benchmark model will be an ARMAX model that contains lags of the return variable, lags of the error term, and lags of the control variables used in my analysis. This leads to equation (4) as my benchmark model:

$$Returns_t = \beta_0 + \sum_{p=1}^{P} \beta_p Returns_{t-p} + \varepsilon_t + \sum_{q=1}^{Q} \beta_q \varepsilon_{t-q} + \sum_{k=1}^{K} \sum_{l=1}^{L_k} \beta_{k,l} Controls_{k,t-l} \tag{4}$$

The appropriate lag length used for the variables will initially be determined using autocorrelation and partial autocorrelation coefficients. I then confirm the correct lag length by first testing the autoregressive component and finding the lowest BIC score of different lags of an autoregressive model. Information criteria such as the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) have been valuable sources for model selection (Drton & Plummer, 2017). I focused on the BIC score as it penalises the complexity of the model more than other criteria, such as the AIC. This will allow me to extract the simplest model that still fits the data well. Using the lag obtained for the autoregressive component, I find the most suitable ARMA model by finding the model with the lowest BIC score when different lags of the moving average component are used. Finally, this model is used to find the most suitable ARMAX model by adding lags of the control variable and finding the one with the lowest BIC score. This is all done using the in-sample dataset.

Now expanding this model to include lags of the daily news sentiment score, I will test whether this model has better predictive capabilities compared to the benchmark model which does not contain the daily news sentiment score as a variable. Through this, I can observe whether news sentiment would improve the information set of the benchmark model and offer predictive power beyond looking at past return values. By expanding the benchmark model, I obtain equation (5) below:

$$\begin{aligned} Returns_t = \beta_0 &+ \sum_{p=1}^{P} \beta_p Returns_{t-p} + \varepsilon_t + \sum_{q=1}^{Q} \beta_q \varepsilon_{t-q} + \sum_{k=1}^{K} \sum_{l=1}^{L_k} \beta_{k,l} Controls_{k,t-l} \\ &+ \sum_{r=1}^{R} \beta_r Sentiment_{t-r} \end{aligned} \tag{5}$$

The appropriate lag length of the daily news sentiment score will be found by similarly minimising the BIC values when including different lags of the daily news sentiment score in the benchmark model. This would lead to a parsimonious specification that uses the least possible terms but still fits the data well.

After constructing both models in relation to the in-sample dataset, I use these models to predict values in the same dataset. This would allow me to compare the predicted values of both models with the actual values that occurred and assess the model's accuracy. I will then calculate three metrics of accuracy, which are the Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). As these metrics assess how inaccurate the predictions of a model are, I will observe which model has the lowest value in relation to these metrics. Afterward, a formal test will be conducted to check whether these metrics significantly differ between the models using the Diebold and Mariano (1995) test. The formulas associated with the accuracy metrics are presented below, where $Returns_t$ refers to the observed return during the period t, and $E_t(Returns_t)$ refers to the model's predicted return during the period t.

$$MSE = \frac{1}{T} \sum_{t=1}^{T} (Returns_t - E_t(Returns_t))^2 \tag{6}$$

$$MAE = \frac{1}{T} \sum_{t=1}^{T} |Returns_t - E_t(Returns_t)| \tag{7}$$

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{Returns_t - E_t(Returns_t)}{Returns_t} \right| \tag{8}$$

Two common metrics that I will also look at that are used to assess the fit of a model while penalising its complexity are the BIC and AIC values. These metrics will be assessed between models in the in-sample dataset to understand which model fits the in-sample data better while keeping its simplicity a priority.

4.4.2 Out-of-Sample Analysis

Using the ARMAX models trained in the in-sample dataset, I also use these to predict the returns that would occur in the out-of-sample dataset. After predicting the returns using both models, I similarly compare these predicted returns with the observed returns that occurred during this period. This will allow me to assess how accurate these models are at forecasting returns when trained using historical data. The metrics used to assess the accuracy of the models will similarly be the Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The model with the lowest value in relation to these metrics would be the model that is best at forecasting returns in the out-of-sample dataset. Afterwards, I will use the Diebold and Mariano (1995) test to assess whether these accuracy metrics significantly differ between the models.

4.5 Second Robustness Tests

My main analysis revolves around using open prices to calculate the daily returns of the Nikkei 225 index. The paper done by Hudson and Gregoriou (2015), however, highlights the significant differences in the various methods of calculating returns, particularly between logarithmic returns and simple returns. The paper provides theoretical proof that the average logarithmic return is less than the average simple return by the variance of these returns. This has been proven to be very important when dealing with short periods such as daily returns. Therefore, there is not necessarily a one-to-one relationship between the different return calculations. As a robustness test, I will first repeat the previous steps using the difference in the natural logarithm of open prices between day t and day t + 1 as a metric for returns. This is shown in formula (9). Additionally, I will repeat my analysis using open-to-close returns, as shown in formula (10).

$$Daily\ Returns_t = Ln\ (Open_{t+1}) \text{ - } Ln\ (Open_t) \tag{9}$$

$$Daily\ Returns_t = \frac{Close_t - Open_t}{Open_t} \tag{10}$$

CHAPTER 5 Results & Discussion

5.1 Initial Results

In this section, I report the results of my analysis starting with the preliminary results. I first correlated the variables against each other to look for any signs of multicollinearity and assess the correlation between my dependent and independent variables. From Table D1 in Appendix D, I observe that there appears to be no sign of multicollinearity as none of the correlation coefficients seem to be extremely high (>0.8). There does seem to be a relatively high positive correlation between daily news sentiment and the daily returns of the Nikkei 225 index. As the correlation coefficient between these two variables is approximately 0.35, it seems to support the initial hypothesis that higher daily news sentiment values result in higher same-day returns of the Nikkei 225 index. The correlation coefficient between the natural logarithm of the daily number of news articles and the daily returns of the Nikkei 225 index, however, seems to be significantly lower at around 0.02.

I also analysed the scatter plots of the daily news sentiment variable and the natural logarithm of the daily number of news articles against the daily returns of the Nikkei 225 index to assess the linearity of the data and confirm the initial relationship found in the correlation coefficients. Figure 2a represents the scatter plot of the daily news sentiment variable against the daily returns of the Nikkei 225 index. Here we see a slightly positive and linear relationship between the two variables, however, the spread of values appears to be large. This also seems to support the initial hypothesis that higher daily news sentiment values result in higher same-day returns of the Nikkei 225 index. Figure 2b presents the scatter plot of the natural logarithm of the daily number of news articles against the daily returns of the Nikkei 225 index. From this figure we see a linear relationship, however, it appears to be horizontal with no visible trend.

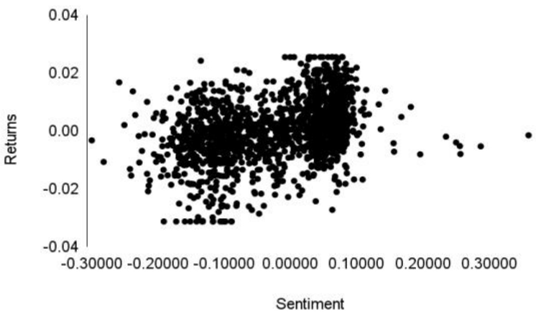

Figure 2a: Scatter Plot - News Sentiment Score. Scatter plot of the daily news sentiment score on the x-axis and the daily returns on the y-axis.

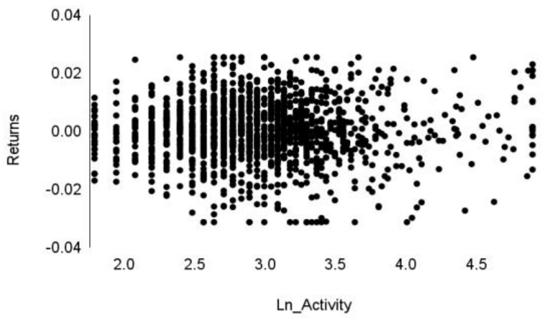

Figure 2b: Scatter Plot - News Activity. Scatter plot of the natural logarithm of daily news activity on the x-axis and the daily returns on the y-axis.

5.2 Contemporaneous Regressions

As correlation coefficients and scatter plots only represent preliminary results and are not conclusive of any relationship between the variables, a more thorough analysis will be performed. For this, I performed six OLS regressions using news sentiment, the natural logarithm of news activity, the split between positive and negative news sentiment, and the interaction between daily news sentiment and the natural logarithm of news activity as independent variables. For all the regressions, the dependent variable is the daily returns of the Nikkei 225 index. Standard errors are Newey-West (1987) adjusted up to six lags to control for heteroskedasticity and autocorrelation.

The first regression I performed used daily news sentiment values as the only independent variable without any control variables. This analysis is done to isolate the effect of news sentiment on the daily returns of the Nikkei 225 index. The results of this initial regression can be found in Table 2 under column (1). From this initial regression, it is clear that the R^2 and adjusted R^2 of this model are similar at around 12.4%. These values indicate that about 12.4% of the variance in the daily returns of the Nikkei 225 index can be explained by the daily news sentiment score. Additionally, the coefficient of the variable representing daily news sentiment scores is positive at around 0.041, which can be interpreted as follows: A 1-point increase (decrease) in the daily news sentiment score on average leads to an increase (decrease) of the Nikkei 225 index price on the same day by about 4.1%. This is because the returns variable (dependent variable) is calculated as the percentage change in open prices. The corresponding p-value is also smaller than 1%, which indicates that this result is statistically significant. These preliminary results match the results found in the scatter plot and correlation coefficients and seem to support the initial hypothesis that higher daily news sentiment values result in higher same-day returns of the Nikkei 225 index.

The second regression I performed used the natural logarithm of daily news activity as the only independent variable without any control variables to isolate the effect of news activity on daily returns. The results of this regression can be found in Table 2 under column (2). This regression, however, displays R^2 and adjusted R^2 values of 0%, implying that none of the variance in the daily returns of the Nikkei 225 index can be explained by the variable representing the natural logarithm of daily news activity. Additionally, the coefficient of the natural logarithm of daily news activity in this regression is zero and insignificant at the 10% level. Therefore, I cannot reject the null hypothesis that a change in news activity does not lead to any changes in the daily returns of the Nikkei 225 index.

The third regression combines the two variables by including the daily news sentiment and the natural logarithm of daily news activity as independent variables without including any control variables. This regression is performed to understand how the introduction of both variables would alter the previous outcome. The results of this regression can be found in Table 2 under column (3). Both the R^2 and adjusted R^2 values for this regression seem to increase compared to the first (1) model to approximately 12.6% and 12.5%, respectively. This shows that the addition of the natural logarithm of daily news activity to the first regression helps in explaining more of the variance in the daily returns of the Nikkei 225 index. The improvement in the adjusted R^2 also shows that even after being penalised for adding another variable, the model's explanatory power of the variance in the daily returns increases in relation to the first model. Compared to the first model, the coefficient of the daily news sentiment remains the same at 0.041 and can be similarly interpreted in the following way: A 1-point increase (decrease) in the daily news sentiment score on average leads to an increase (decrease) of the Nikkei 225 index price on the same day by about 4.1%. The corresponding p-value is smaller than 1%, which indicates that this result is also statistically significant. The coefficient of the natural logarithm of daily news activity in this regression is 0.001 and is insignificant at the 10% level. Therefore, I cannot reject the null hypothesis that this variable has no impact on the daily returns of the Nikkei 225 index.

The fourth regression I performed extended the third regression to include the four control variables used in my analysis. This is done to limit the effect of confounding variables that can lead to bias in my results. The results of this regression can be found in Table 2 under column (4). Although the R^2 value increased compared to the third model (3) to around 12.7%, the adjusted R^2 value decreased to 12.3%. This shows that the addition of the control variables does not significantly improve the model's explanatory power of the variance in the daily returns. Nevertheless, the coefficient of the daily news sentiment variable slightly increases to 0.042 and is significant at the 1% level. The coefficient of the natural logarithm of daily news activity, however, remains at 0.001 and is insignificant at the 10% level. When examining the coefficients of the control variables, they are all insignificant at the 10% level and, therefore, the null hypothesis that these variables have no impact on the daily returns of the Nikkei 225 index cannot be rejected.

5.2.1 Hypothesis 1.1

All the previous results indicate that there is a statistically significant positive correlation between daily news sentiment and the daily returns of the Nikkei 225 index. This is proven by the correlation coefficients, scatter plots, and the three regressions performed including daily news sentiment as the independent variable. As the fourth model (4) in Table 2 presents a statistically significant positive coefficient of 0.042 associated with daily news sentiment, **I do not reject hypothesis 1.1:** Daily news article sentiment scores related to the Nikkei 225 index is positively correlated with same-day market returns. The coefficient of 0.042 can be interpreted as follows: A 1-point increase (decrease) in the daily news sentiment score on average leads to an increase (decrease) of the Nikkei 225 index price on the same day by about 4.2%.

5.2.2 Hypothesis 1.2

In order to understand whether the effect of negative sentiment would have a larger impact on the daily returns of the Nikkei 225 index compared to positive sentiment, I split the sentiment score into two variables and assessed the difference and significance of the coefficients. This regression also includes the natural logarithm of news activity and the four control variables to limit the effect of confounding variables that can lead to potential biases in my results. The results of this regression can be found in Table 2 under the column (5). The R^2 and adjusted R^2 values associated with this model do not change compared to the fourth model and remain at 12.7% and 12.3%, respectively. This indicates that splitting daily news sentiment into positive and negative values does not significantly increase or decrease the variable's explanatory power of the variance in daily returns compared to the fourth model (4). From the coefficients of the variables, multiple results are apparent. First, the coefficient of the variable representing positive sentiment scores is positive at around 0.038. This can be interpreted as follows: A 1-point increase (decrease) in positive news sentiment scores on average leads to an increase (decrease) of the Nikkei 225 index price on the same day by about 3.8%. The corresponding p-value is smaller than 1%, which indicates that this result is statistically significant. Secondly, the coefficient of the variable representing negative sentiment scores is positive at around 0.044, which can be interpreted as follows: A 1-point increase (decrease) in negative news sentiment scores on average leads to an increase (decrease) of the Nikkei 225 index price on the same day by about 4.4%. The corresponding p-value is smaller than 1%, which indicates that this result is also statistically significant. These results seem to support the theory that negative news sentiment has a larger impact on returns compared to positive sentiment of the same magnitude. Therefore, I find substantial evidence in favour of this theory, and **I do not reject hypothesis 1.2:** Negative news sentiment scores inflict a greater effect on daily Nikkei 225 index returns as opposed to positive news sentiment scores. The variable representing the natural logarithm of news activity and the four control variables are all insignificant at the 10% level and, therefore, I cannot reject the null hypothesis that these variables have no impact on the daily returns of the Nikkei 225.

5.2.3 Hypothesis 1.3

The last regression assesses whether the effect of news sentiment on the daily returns of the Nikkei 225 index increases when more news activity is presented on that day. This is done through an interaction variable, the results of which can be found in Table 2 under column (6). Both the R^2 and adjusted R^2

values increase significantly to around 13.2% and 12.8%, respectively. This indicates that the interaction effect offers significantly more explanatory power of the variance in daily returns compared to the previous models. Additionally, as the coefficient of the interaction variable is positive at 0.016, the effect of daily news sentiment on the returns of the Nikkei 225 index is stronger when there is more news coverage on that day. The coefficient can be interpreted as follows: a 1 unit increase (decrease) in the natural logarithm of news activity heightens (reduces) the effect of a 1 unit increase in news sentiment on the daily returns of the Nikkei 225 index by approximately 1.6%. The corresponding p-value is also smaller than 1%, which indicates that this result is statistically significant. Table 3 shows the marginal effect of news sentiment scores on daily returns with different values of the natural logarithm of news activity. The consistent increases in the coefficient associated with news sentiment as news activity increases provide support for the hypothesis that the impact of news sentiment on daily returns depends on the amount of news activity presented on that day. Therefore, **I do not reject hypothesis 1.3:** The correlation between daily news article sentiment scores and same-day Nikkei 225 index returns is strengthened with more news activity associated with the index. When examining the coefficients of the control variables, they are all insignificant at the 10% level and, therefore, the null hypothesis that these variables have no impact on the daily returns of the Nikkei 225 index cannot be rejected.

Dependent Variable: $Returns_t$

Variables	(1)	(2)	(3)	(4)	(5)	(6)
$Sentiment_t$	0.041***		0.041***	0.042***		-0.006
	(0.003)		(0.003)	(0.003)		(0.018)
$Ln\_Activity_t$		0.000	0.001	0.001	0.001	0.001
		(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
$Positive\_Sentiment_t$					0.038***	
					(0.010)	
$Negative\_Sentiment_t$					0.044***	
					(0.006)	
$Interaction_t$						0.016***
						(0.006)
$Previous\_Returns_t$				-0.009	-0.009	-0.012
				(0.014)	(0.014)	(0.014)
$Ln\_AvgVolatility_t$				0.001	0.001	0.001
				(0.001)	(0.001)	(0.001)
$D\_Ln\_MarketCap_{t-1}$				0.005	0.005	0.004
				(0.032)	(0.032)	(0.032)
$AbnTurn_{t-1}$				0.000	0.000	0.000
				(0.001)	(0.001)	(0.001)
Constant	0.001***	-0.001	-0.001	-0.002	-0.002	-0.003
	(0.000)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Observations	1,493	1,493	1,493	1,493	1,493	1,493
R^2	0.124	0.000	0.126	0.127	0.127	0.132
Adjusted R^2	0.124	0.000	0.125	0.123	0.123	0.128

Table 2: Regression Results. This table presents the results of the regressions. The dependent variable is the daily returns of the Nikkei 225 index. The independent variables are the daily news sentiment scores and the natural logarithm of the number of articles per day. The control variables are the cumulative past 5-day returns of the Nikkei 225 index, the natural logarithm of the 5-day average volatility, the differenced natural logarithm of the previous day’s market capitalization, and the previous day’s abnormal turnover. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Ln_Activity_t Value	Coefficient of Sentiment_t
1	0.010
	(0.012)
1.5	0.018**
	(0.009)
2	0.026***
	(0.006)
2.5	0.034***
	(0.004)
3	0.042***
	(0.003)
3.5	0.050***
	(0.005)
4	0.058***
	(0.008)

Table 3: Change in Sentiment Coefficients. This table presents the change in the coefficient of News Sentiment as the natural logarithm of the daily number of news articles increases. The first column presents the different values of the natural logarithm of the daily number of news articles while the second column shows the coefficient of news sentiment on the daily returns of the Nikkei 225 index associated with the different values of news activity. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01.

5.2.4 First Robustness Test Results

In order to assess whether my regression results are robust to different sub-periods surrounding the COVID-19 pandemic, I split my dataset into three periods encapsulating the time before, during, and after the pandemic. After running the regressions in each period, the results in tables G1, G2, and G3 (Appendix G) present the output of the regression analyses performed in the period before, during, and after the pandemic, respectively. While the coefficients and significance levels of the news sentiment variable in the different sub-periods are similar to those in Table 2, the division between positive and negative sentiment presents different results. Before the pandemic, positive sentiment seemed to have a greater impact on daily returns compared to negative sentiment. This, however, flipped during and after the pandemic with negative sentiment having a larger impact on daily returns compared to positive sentiment. This discrepancy seemed to be highest during the pandemic, as shown in table G2 (Appendix G). A possible explanation for this might be that the fear of a recession caused by the pandemic made investors more risk-averse. The paper done by Guo et al. (2006) supports this by showing that investors are more risk-averse during economic downturns and require higher rates of return for investing in stocks. Lastly, the coefficients of the interaction effect between news sentiment and activity in the different sub-periods were all positive but only significant at the 10% level in the period before the pandemic. Therefore, while the overall news sentiment results were robust to different time periods, the results of the division between positive and negative sentiment and the interaction effect between news sentiment and activity might have been influenced by the pandemic. My results are, therefore, sensitive to different time periods, and future research should consider different time periods when analysing news sentiment and index returns.

5.3 Predictive Analysis

Before assessing the forecasting capabilities of news sentiment on the daily returns of the Nikkei 225 index, two ARMAX models were created (one with lags of daily news sentiment values as a variable and one without) with the appropriate number of lags. These two models were used to forecast daily returns in both the in-sample dataset and the out-of-sample dataset. The in-sample dataset consists of data between the start of 2018 and the end of 2021, while the out-of-sample dataset starts in 2022 and ends in February 2024.

The appropriate lag length of the benchmark model was determined using the in-sample dataset for which the lowest BIC values would determine the most parsimonious specification. Beginning with the autoregressive component, the first three columns in Table E1 (Appendix E) present the different BIC values associated with different lags of the daily returns of the Nikkei 225 index. As the BIC values consistently increased with more lags, the AR (1) model was used, and one lag was the most appropriate lag length of the daily returns variable. The last three columns in Table E1 (Appendix E) analyse the moving average component and present the different BIC values associated with different lags of the error term of daily returns. This was done using an AR (1) model. As the BIC values increased with more lags of the error term, the preferred model was an ARMA (1,1) model. Using this specification, Table E2 (Appendix E) displays an ARMA (1,1) model with different lags of the four control variables used in my analysis. As the BIC values of these models increased with more lags, the preferred benchmark model was the ARMAX (1,1) model with one lag of the four control variables.

After constructing the benchmark model using the in-sample dataset, I continue to expand this benchmark model using lags of daily news sentiment scores. This extended model will then be compared to the benchmark model's forecasts of both in-sample and out-of-sample values. Table E3 (Appendix E) presents the different BIC values associated with different lags of daily news sentiment scores. This was done using an ARMAX (1,1) model with one lag of the four control variables. As the BIC values consistently increased with more lags, the preferred extended model was an ARMAX (1,1) model with one lag of the four control variables and one lag of daily news sentiment scores.

5.3.1 In-Sample Performance

I began by forecasting the daily returns in the in-sample dataset using both models and computing the Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) values. As these metrics measure how inaccurate a model's forecasts are compared to the actual values that occurred, the model with the lowest value indicates the preferred model. The differences between the calculation of these metrics, however, could imply conflicting results. The MSE often punishes large outliers more than the other two metrics, as it is the only one that squares the errors instead of taking the absolute value. The MAPE and MAE are less sensitive to outliers, with the MAPE being scale-independent as it takes relative error measurements instead of absolute measurements. This also makes this metric easier to interpret due to its percentage-based calculations. The results of these metrics associated with the benchmark model and the extended model in the in-sample dataset are presented in Table 4 below.

Performance Metric	Benchmark Model	Extended Model	Diebold-Mariano Test
MSE	0.000112091	**0.000112063**	0.252
MAE	**0.008064605**	0.008068961	-0.792
MAPE	**1.443387269**	1.503961818	-1.895*
AIC	**-6029.536**	-6027.776	---
BIC	**-5990.551**	-5983.917	---

Table 4: Forecasting Accuracy of the Benchmark Model and Extended Model (In-Sample). This table presents the Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), AIC, and BIC values associated with the benchmark model and the extended model based on the in-sample data set. The value in bold indicates the preferred model by performance metric. The last column displays the test statistic of the Diebold and Mariano (1995) Test to see whether these metrics significantly differ from each other. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results indicate that the MSE is lower for the extended model including news sentiment as a variable compared to the benchmark model. The MAE and MAPE, however, are both lower for the benchmark model compared to the extended model. This could indicate that in the in-sample dataset, the extended model produces forecasts that contain fewer outliers in relation to the observed values compared to the benchmark model. On average, however, the benchmark model produced forecasts that were more in line with the actual returns. This could mean that using the extended model might produce more inaccurate forecasts, however, these forecasts contain fewer outliers, which could be viewed as a safer forecasting model. Table 4 also presents the AIC and BIC values associated with each model. From this, we can see that the benchmark provides a better fit to the data compared to the extended model, as the AIC and BIC values are both lower for the benchmark model.

A Diebold and Mariano (1995) test was also performed to assess whether the difference in the MSE, MAE, and MAPE values between the models was statistically significant. From the last column in Table 4, we see that only the difference between the MAPE values of the models was statistically significant at the 10% level. Therefore, the difference in forecasting accuracy was mostly insignificant. This could be the result of the models being trained and tested in the same data set and producing forecasts in line with actual observations. From the results in Table 4, there is not enough evidence that news sentiment improves the information set of the benchmark model and offers predictive power beyond looking at past return values. Therefore, from the in-sample analysis, there is not enough evidence to support hypothesis 2: Daily news article sentiment scores related to the Nikkei 225 index can be used to predict next-day market returns.

5.3.2 Out-of-Sample Performance

As we are interested in the forecasting capabilities of different models after training them using historical data, the forecasting accuracy in the out-of-sample dataset would more reliably reflect whether news sentiment offers predictive power beyond looking at past values of the daily returns of the Nikkei 225 index. This helps with the problem of overfitting as we assess the accuracy of the forecasts in relation to data that the model has not been trained on. This would more accurately reflect the real-world application of using news sentiment to predict Nikkei 225 index returns as we do not have access to future index returns.

I similarly used the benchmark and extended model trained with the in-sample dataset; however, I used them to forecast the returns in the out-of-sample dataset. From this, I also calculated the Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) values. The results of these metrics associated with the benchmark model and the extended model in the out-of-sample dataset are presented in Table 5 below.

Performance Metric	Benchmark Model	Extended Model	Diebold-Mariano Test
MSE	0.000103065	**0.000102702**	2.582***
MAE	0.008021229	**0.008001648**	2.729***
MAPE	1.312243212	**1.281003758**	1.682*

Table 5: Forecasting Accuracy of Benchmark Model and Extended Model (Out-of-Sample). This table presents the Mean Square Error (MSE), Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE) values associated with the benchmark model and the extended model based on the out-of-sample dataset. The value in bold indicates the preferred model by performance metric. The last column displays the test statistic of the Diebold and Mariano (1995) Test to see whether these metrics significantly differ from each other. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results of the out-of-sample analysis show that the extended model provides more accurate forecasts compared to the benchmark model in all three performance metrics. This is proven by the lower MSE, MAE, and MAPE values, which indicates that the extended model that includes lags of news sentiment values as a variable provides forecasts that are more in line with observed values and also contain fewer outliers in relation to the actual returns. A Diebold and Mariano (1995) test was similarly performed to assess whether the differences in the three metrics between the models were statistically significant. The last column in Table 5 presents the test statistic of the Diebold and Mariano test and shows that the difference in MSE and MAE are significant at the 1% level, while the difference in MAPE is significant at the 10% level. This provides evidence that the inclusion of lags of news sentiment values improves the information set of the benchmark model and offers predictive power beyond looking at past returns. As this is done in the out-of-sample data set, this analysis more accurately reflects the real-world application of training a model using historical data and using news sentiment to improve the forecasting accuracy of the Nikkei 225 index. Therefore, given the consistent differences in performance metrics between the models and the significance of these differences, **I do not reject hypothesis 2**: Daily news article sentiment scores related to the Nikkei 225 index can be used to predict next-day market returns.

5.4 Second Robustness Test Results

As a robustness test, I repeated my analysis using two different methods of calculating returns. The results using the difference in the natural logarithm of open prices between day t and day t + 1 as daily returns remain consistent. As the coefficients and significance levels in the regressions of this robustness test in Table F4 (Appendix F) are similar to those in Table 2, the regression results are robust to a change in the calculation of returns when using the difference in the natural logarithm of open prices. Tables F5 and F6 (Appendix F) display the in-sample and out-of-sample analyses, respectively. The results in these tables (values and significance levels) also show similar results to Tables 4 & 5 and, therefore, the in-sample and out-of-sample analyses are robust to changes in the calculation of returns when using the difference in the natural logarithm of open prices.

When calculating returns using open-to-close prices, the results are slightly different. From the regression results of this robustness test in Table F1 (Appendix F), the coefficients and significance levels are similar to those in Table 2 and are, therefore, robust to a change in the calculation of returns when using open-to-close prices. Tables F2 and F3 (Appendix F), similarly present the in-sample and out-of-sample analyses, respectively. While the results of the in-sample analysis are similar to the results in Table 4, the out-of-sample analysis presents insignificant Diebold and Mariano (1995) test statistics (unlike the results in Table 5). These variations suggest that my initial findings related to the predictive power of news sentiment are sensitive to using open-to-open prices to calculate daily returns. These changes might be the result of changes in the price of the Nikkei 225 index during non-trading hours between the closing time of day t and the opening time of day t + 1. Future research should, therefore, consider alternative calculations of daily returns when dealing with stock indices that have significant price changes during non-trading hours.

5.5 Discussion of Results

My results show that overall news sentiment is positively correlated with the daily returns of the Nikkei 225 index. This finding is similar to previous studies that examine different indices (mostly the DJIA index), such as Tetlock (2007) and Allen et al. (2019b). This shows that the positive relationship between news sentiment and daily returns could be present in other Asian indices as well. My overall analysis also suggested that negative sentiment has a greater impact on the daily returns of the Nikkei 225 index compared to positive sentiment. This is similar to previous studies that analyse different financial securities and use other sources of sentiment, such as Affuso and Lahtinen (2018) and Smales (2014). This finding is also in line with the traditional view that investors are generally more risk-averse and sensitive to negative news (Chen et al., 2003). My robustness test, however, shows that this risk-averse behaviour might have been the result of the COVID-19 pandemic, as the period before the pandemic seemed to indicate that positive sentiment had a greater impact on the returns of the Nikkei 225 index compared to negative sentiment. As this switch in behaviour occurred during the pandemic, it is possible that risk-averse behaviour depends on the business cycle and that it cannot be assumed for every time period. My overall analysis also indicated that the impact of news sentiment on the returns of the Nikkei 225 index gets strengthened when more news coverage is present on that day. This is in line with the theory that investor attention is limited and that the rate of change of an asset's price increases when investors focus more of their attention on the events surrounding that asset (Smales, 2021). My robustness tests show that while this relationship is consistent in every time period, its significance changes drastically between periods. This indicates that the time period plays a crucial part in assessing the importance of investor attention in the relationship between news sentiment and daily returns.

As for the forecasting capabilities of news sentiment, my out-of-sample analysis shows that lags of news sentiment values significantly improves the information set of the benchmark model and offers predictive power beyond looking at past returns. This indicates that models that include news sentiment as a variable can be trained using historical data to forecast future changes in the price of the Nikkei 225 index more accurately. This finding is similar to previous studies that examine different indices (mostly the S&P 500 index), such as Mohan et al. (2019) and Fazlija and Harder (2022). This shows that the forecasting capabilities of news sentiment could be present in other Asian indices as well. My robustness tests, however, show that this finding is sensitive to using open-to-open prices to calculate daily returns, as there may be significant price changes during non-trading hours.

CHAPTER 6 Conclusion

In this thesis, I analysed the impact and forecasting capabilities of news sentiment and activity on the daily returns of the Nikkei 225 index. Previous studies find that news sentiment impacts stock market returns and can be used to forecast price changes. While these results have been found in the context of stock indices in the U.S. and Europe, the Asian region remains relatively understudied. To my knowledge, until this study, no research has analysed the impact and forecasting capabilities of news sentiment on the Nikkei 225 index. My research also used a large database of various news articles instead of relying on one news article provider like many papers do. Therefore, the research question that followed was: *"How and to what extent does news article activity and sentiment contain explanatory and forecasting capabilities for Nikkei 225 Index returns?"*

In order to analyse this relationship, data regarding daily news sentiment and activity was obtained from Ravenpack, which is provided by Wharton Research Data Services (Wharton Research Data Services, n.d.-b). Data on the daily returns of the Nikkei 225 index and the control variables used in my analysis, on the other hand, were obtained from Bloomberg. I focused my research on data from 2018 until 2024. I then performed several regressions to analyse the impact of news sentiment on the daily returns of the Nikkei 225 index, the difference in the impact of negative and positive news sentiment, and how the impact of news sentiment on daily returns changes with more news coverage. I also used ARMAX models to assess the forecasting capabilities of news sentiment on the daily returns of the Nikkei 225 index. This was done by training two ARMAX models, one with lags of daily news sentiment values as a variable and one without, in the in-sample dataset. The forecasting accuracy of both models was then assessed in both the in-sample dataset and the out-of-sample dataset.

My results show that there is a positive relationship between overall news sentiment and daily returns of the Nikkei 225 index. This is in line with previous studies and indicates that this relationship could be present when analysing other Asian indices as well. Additionally, negative sentiment seemed to impact the daily returns of the index more than positive sentiment. This is also in line with previous studies and can be linked to the risk-averse behaviour of investors (Chen et al., 2003). My robustness test, however, shows that this behaviour might be the result of the pandemic and that the business cycle could affect the risk-averse behaviour of investors. Furthermore, my analysis shows that the impact of news sentiment on the returns of the Nikkei 225 index increases with news coverage, which is in line with theories of limited investor attention (Smales, 2021). The significance of these results, however, is dependent on the time period analysed. Finally, my out-of-sample analysis shows that news sentiment significantly improves the information set of the benchmark model and offers predictive power beyond looking at past returns. Combined with findings from previous studies on other indices, this suggests that the forecasting capabilities of news sentiment could be present in other Asian indices as well. My robustness tests, however, show that this finding is sensitive to using open-to-open prices to calculate daily returns, as there may be significant price changes during non-trading hours.

While this research does have its limitations, these can be potential stepping stones for future research. Firstly, I decided to analyse news sentiment and returns related to the entire Nikkei 225 index as this is usually discussed more often in the news compared to the individual companies within the index. Additionally, there were significant data restrictions regarding news sentiment scores related to individual companies within the Nikkei 225 index. Future research could investigate the relationship between news sentiment and stock returns of the individual companies within the Nikkei 225 index. This would allow for a more nuanced understanding and assess whether news sentiment can be used as an additional factor in multi-factor models such as the Fama-French three-factor model. Furthermore, as the Nikkei 225 index is not the only major index in Japan, data restrictions prevented me from analysing other indices in Japan, such as the TOPIX index. Future research could obtain sentiment values related to other indices through other sources and assess whether the results found in my analysis would still hold. Lastly, my analysis utilised short-term data such as daily returns which did not allow me to understand whether there were any long-term relationships between news sentiment and index returns. Future research could, therefore, analyse the long-term impacts of news sentiment on index returns by looking at weekly or monthly data.

REFERENCES

Affuso, E., & Lahtinen, K. D. (2018). Social media sentiment and market behavior. Empirical Economics, 57(1), 105–127. https://doi.org/10.1007/s00181-018-1430-y

Allen, D. E., McAleer, M., & Singh, A. K. (2019). Daily market news sentiment and stock prices. Applied Economics, 51(30), 3212–3235. https://doi.org/10.1080/00036846.2018.1564115

Allen, D. E., McAleer, M., & Singh, A. K. (2019b). Daily market news sentiment and stock prices. Applied Economics, 51(30), 3212–3235. https://doi.org/10.1080/00036846.2018.1564115

Bajo, E., & Raimondo, C. (2017b). Media sentiment and IPO underpricing. Journal of Corporate Finance, 46, 139–153. https://doi.org/10.1016/j.jcorpfin.2017.06.003

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. International Review of Financial Analysis (Online)/International Review of Financial Analysis, 45, 150–156. https://doi.org/10.1016/j.irfa.2016.03.015

Chen, C. W. S., Chiang, T. C., & So, M. K. P. (2003). Asymmetrical reaction to US stock-return news: evidence from major stock markets based on a double-threshold model. Journal of Economics and Business, 55(5–6), 487–502. https://doi.org/10.1016/s0148-6195(03)00051-1

Chen, H., De, P., Hu, Y. J., & Hwang, B. (2014). Wisdom of Crowds: the value of stock opinions transmitted through social media. the Review of Financial Studies, 27(5), 1367–1403. https://doi.org/10.1093/rfs/hhu001

DeLong, J., Shleifer, A., Summers, L. H., & Waldmann, R. J. (n.d.). Positive feedback investment strategies and destabilizing rational speculation. Andrei Shleifer. https://scholar.harvard.edu/shleifer/publications/positive-feedback-investment-strategies-and-destabilizing-rational-speculation

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Dong, H., & Gil-Bazo, J. (2020). Sentiment stocks. International Review of Financial Analysis (Online)/International Review of Financial Analysis, 72, 101573. https://doi.org/10.1016/j.irfa.2020.101573

Drton, M., & Plummer, M. (2017). A Bayesian information criterion for singular models. Journal of the Royal Statistical Society. Series B, Statistical Methodology, 79(2), 323–380. https://doi.org/10.1111/rssb.12187

Du, W. (2020). News and market efficiency in the Japanese stock market. the Journal of Behavioral Finance, 22(3), 306–319. https://doi.org/10.1080/15427560.2020.1774886

Du, W. (2020). News and market efficiency in the Japanese stock market. ˜the œJournal of Behavioral Finance, 22(3), 306–319. https://doi.org/10.1080/15427560.2020.1774886

Fazlija, B., & Harder, P. (2022). Using financial news sentiment for stock price direction prediction. Mathematics, 10(13), 2156. https://doi.org/10.3390/math10132156

Gan, B., Alexeev, V., & Yeung, D. (2023). Tweets versus broadsheets: Sentiment impact on stock markets around the world. the Journal of Financial Research/Journal of Financial Research. https://doi.org/10.1111/jfir.12380

Gherghina, Ș. C., Mehdian, S., & Stoica, O. (2023). INVESTORS' SENTIMENT AND EQUITY MARKETS DURING COVID-19 PERIOD: A QUANTILE REGRESSION APPROACH AND WAVELET ANALYSIS. Journal of Business Economics and Management, 24(3), 551–575. https://doi.org/10.3846/jbem.2023.19814

Guo, H., Campbell, J. Y., Cochrane, J. H., Lettau, & Ludvigson. (2006). Are investors more Risk-Averse during recessions? https://web.archive.org/web/20190428030955id_/https://files.stlouisfed.org/files/htdocs/publications/es/06/ES0623.pdf

Han, L., Li, Z., & Yin, L. (2018). Investor attention and stock returns: international evidence. Emerging Markets Finance & Trade, 54(14), 3168–3188. https://doi.org/10.1080/1540496x.2017.1413980

Hill, J. (2020). Financial markets: equities. In Elsevier eBooks (pp. 97–116). https://doi.org/10.1016/b978-0-12-818692-3.00006-2

Huberman, G., & Regev, T. (2001). Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar. ˜the œJournal of Finance/˜the œJournal of Finance, 56(1), 387–396. https://doi.org/10.1111/0022-1082.00330

Hudson, R. S., & Gregoriou, A. (2015). Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns. International Review of Financial Analysis (Online)/International Review of Financial Analysis, 38, 151–162. https://doi.org/10.1016/j.irfa.2014.10.008

Huynh, T. L. D., Foglia, M., Nasir, M. A., & Angelini, E. (2021). Feverish sentiment and global equity markets during the COVID-19 pandemic. Journal of Economic Behavior & Organization, 188, 1088–1108. https://doi.org/10.1016/j.jebo.2021.06.016

Ishijima, H., Kazumi, T., & Maeda, A. (2015). Sentiment analysis for the Japanese stock market. Global Business and Economics Review, 17(3), 237. https://doi.org/10.1504/gber.2015.070303

Jiao, P., Veiga, A., & Walther, A. (2016). Social media, news media and the stock market. Social Science Research Network. https://doi.org/10.2139/ssrn.2755933

Journal, W. S. (n.d.). Asia Stocks. WSJ. https://www.wsj.com/market-data/stocks/asia

Li, D., Wang, Y., Madden, A. D., Ding, Y., Tang, J., Sun, G., Zhang, N., & Zhou, E. (2019). Analyzing stock market trends using social media user moods and social influence. Journal of the Association for Information Science and Technology, 70(9), 1000–1013. https://doi.org/10.1002/asi.24173

Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., & Anastasiu, D. C. (2019). Stock Price Prediction Using News Sentiment Analysis. 2019 IEEE Fifth International Conference on Big

Data Computing Service and Applications (BigDataService). https://doi.org/10.1109/bigdataservice.2019.00035

Newey, W. K., & West, K. D. (1987). A simple, positive Semi-Definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703. https://doi.org/10.2307/1913610

Nguyen, T. H., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. Expert Systems With Applications, 42(24), 9603–9611. https://doi.org/10.1016/j.eswa.2015.07.052

Nofer, M., & Hinz, O. (2015). Using Twitter to predict the stock market. Business & Information Systems Engineering, 57(4), 229–242. https://doi.org/10.1007/s12599-015-0390-4

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PloS One*, *10*(9), e0138441. https://doi.org/10.1371/journal.pone.0138441

Rennison, J., & Karaian, J. (2024, February 22). Stocks in Japan hit record high, surpassing 1989 peak. The New York Times. https://www.nytimes.com/2024/02/22/business/japan-stocks-record.html

Ruscheinsky, J., Lang, M., & Schäfers, W. (2018). Real estate media sentiment through textual analysis. Journal of Property Investment & Finance/Journal of Property Investment and Finance, 36(5), 410–428. https://doi.org/10.1108/jpif-07-2017-0050

Smales, L. A. (2014). News sentiment in the gold futures market. Journal of Banking & Finance, 49, 275–286. https://doi.org/10.1016/j.jbankfin.2014.09.006

Smales, L. A. (2021). Investor attention and global market returns during the COVID-19 crisis. International Review of Financial Analysis (Online)/International Review of Financial Analysis, 73, 101616. https://doi.org/10.1016/j.irfa.2020.101616

Tan, S. D., & Taş, O. (2020). Social media sentiment in international stock returns and trading activity. the Journal of Behavioral Finance, 22(2), 221–234. https://doi.org/10.1080/15427560.2020.1772261

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. the Journal of Finance/the Journal of Finance, 62(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Toh, M. (2021, February 15). Nikkei hits 30-year high as Japan's economy bounces back better than expected. WRAL.com. https://www.wral.com/story/nikkei-hits-30-year-high-as-japans-economy-bounces-back-better-than-expected/19527831/

Wharton Research Data Services. (n.d.-b). Ravenpack. © 1993-2024 Wharton Research Data Services. Retrieved April 30, 2024, from https://wrds-www.wharton.upenn.edu/pages/get-data/ravenpack-news-analytics-40/

Xu, Y., Wang, J., Chen, Z., & Liang, C. (2021). Sentiment indices and stock returns: Evidence from China. International Journal of Finance & Economics/International Journal of Finance and Economics, 28(1), 1063–1080. https://doi.org/10.1002/ijfe.2463

APPENDIX A: Histograms With Outliers

These figures represent the histograms of the number of articles, daily returns, 5-day average 10-day volatility, and the returns control variable. These figures seem to have outliers and I, therefore, winsorise these variables at the 1% level.

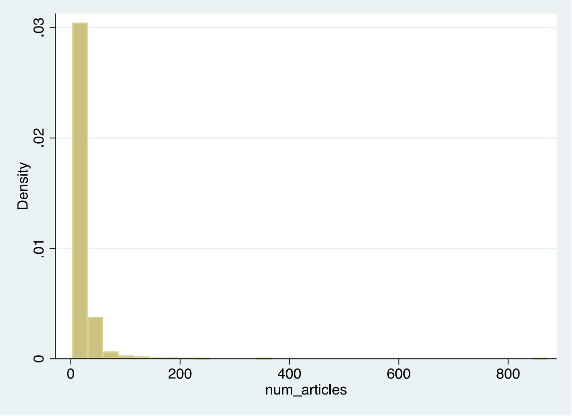

Figure A1: Number of articles. Histogram of the number of articles per day

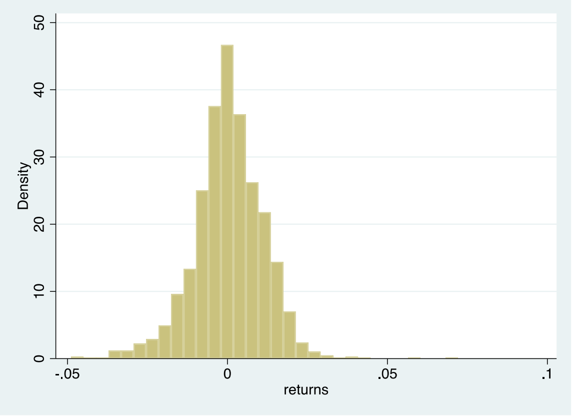

Figure A2: Daily Returns. Histogram of the daily returns of the Nikkei 225 index

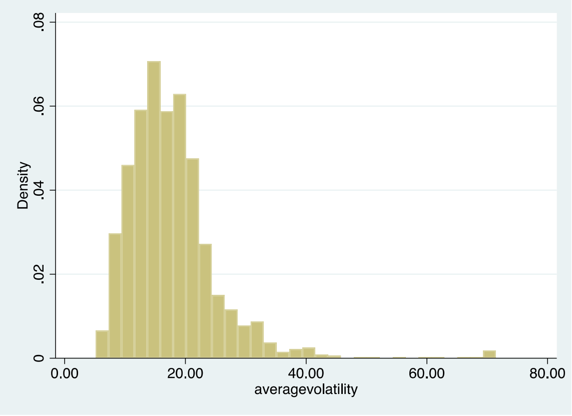

Figure A3: Average Volatility. Histogram of the 5-day average 10-day volatility of the Nikkei 225 index

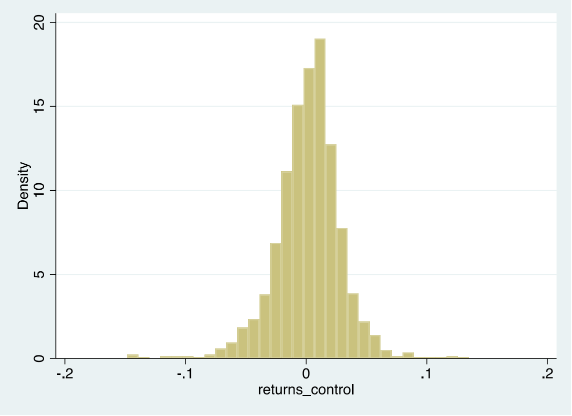

Figure A4: Returns Control. Histogram of the 5-day cumulative returns of the Nikkei 225 index.

APPENDIX B: Histograms With Skewed Distributions

These figures represent the histograms of the winsorised daily number of articles and 5-day average 10-day volatility. These histograms present skewed distributions and I, therefore, take the natural logarithm of these values.

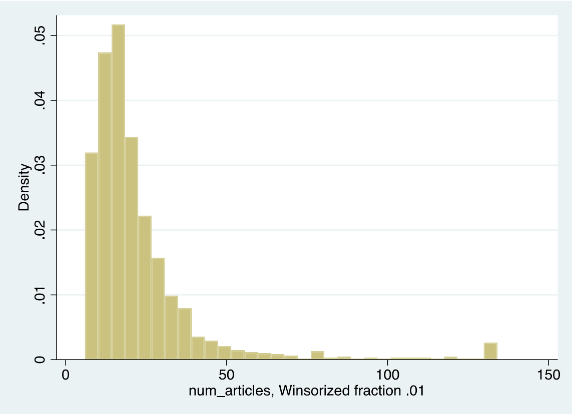

Figure B1: Number of articles. Histogram of the winsorised number of articles per day

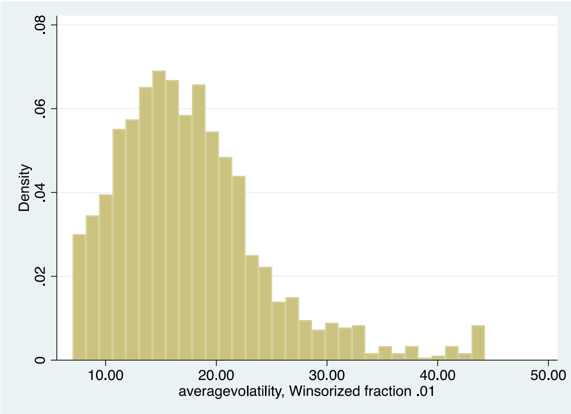

Figure B2: Average Volatility. Histogram of the winsorised 5-day average 10-day volatility of the Nikkei 225 index

APPENDIX C: Dickey Fuller Test

Variable Name	Test Statistic
Sentiment	-35.106***
Ln_Activity	-29.193***
Returns	-35.028***
Previous_Returns	-11.788***
Ln_AvgVolatility	-3.769***
Ln_MarketCap	-2.347
AbnTurn	-26.774***
D_Ln_MarketCap	-37.886***

Table C1: Dickey Fuller Test Results. This table displays the Dickey Fuller Test results of each variable. The variable is associated with test statistics. Different versions of the Dickey Fuller Test were performed depending on the behavior of the time series graph of each variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

APPENDIX D: Correlation Table

	Returns	Sentiment	Ln_Activity	Previous_Returns	Ln_AvgVolatility	D_Ln_MarketCap	AbnTurn
Returns	1.000						
Sentiment	0.3525	1.000					
Ln_Activity	0.0190	-0.0525	1.000				
Previous_Returns	0.0573	0.2170	0.0049	1.000			
Ln_AvgVolatility	0.0213	-0.0116	0.1202	0.0237	1.000		
D_Ln_MarketCap	0.0027	0.0187	0.0103	0.4132	0.0339	1.000	
AbnTurn	0.0123	0.0039	0.0905	-0.1382	-0.1299	-0.1111	1.000

Table D1: Correlation Table. This table represents the correlation coefficients of the different variables included in my regression analysis. This includes the dependent variable (Returns), the independent variables (Daily News Sentiment & The Natural Logarithm of The Daily Number of Articles), and the four control variables used in my analysis.

APPENDIX E: Lag Length Determination

Variables	(1)	(2)	(3)	(4)	(5)	(6)
L1.AR	0.075***	0.072***	0.071***	0.592***	0.565*	0.492
	(0.025)	(0.025)	(0.025)	(0.193)	(0.317)	(0.513)
L2.AR		0.040	0.038			
		(0.025)	(0.025)			
L3.AR			0.022			
			(0.025)			
L1.MA				-0.519**	-0.494	-0.422
				(0.208)	(0.322)	(0.517)
L2.MA					0.006	0.008
					(0.041)	(0.045)
L3.MA						0.014
						(0.041)
Constant	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	967	967	967	967	967	967
BIC	-6025.951	-6020.617	-6014.204	-6021.409	-6014.557	-6007.799

Table E1: ARMA Models and BIC Values. This table presents the different lags of the ARMA model, and the BIC value associated with the different models. The dependent variable is the daily returns of the Nikkei 225 index. Standard errors are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Variables	(1)	(2)	(3)
L1.AR	0.434	-0.209	-0.135
	(0.609)	(0.445)	(0.517)
L1.MA	-0.365	0.277	0.207
	(0.614)	(0.441)	(0.514)
$\text{L1. Previous\_Returns}_t$	0.008	0.052*	0.041
	(0.021)	(0.031)	(0.033)
$\text{L1. Ln\_AvgVolatility}_t$	0.000	0.003	-0.001
	(0.001)	(0.005)	(0.009)
$\text{L1. D\_Ln\_MarketCap}_{t-1}$	0.027	0.031	0.029
	(0.032)	(0.030)	(0.032)
L1. AbnTurn_{t-1}	0.003*	0.003	0.003
	(0.002)	(0.002)	(0.002)
$\text{L2. Previous\_Returns}_t$		-0.041	-0.030
		(0.028)	(0.037)
$\text{L2. Ln\_AvgVolatility}_t$		-0.003	0.006
		(0.005)	(0.017)
$\text{L2. D\_Ln\_MarketCap}_{t-1}$		0.006	0.017
		(0.029)	(0.032)
L2. AbnTurn_{t-1}		0.002	0.002
		(0.002)	(0.002)
$\text{L3. Previous\_Returns}_t$			-0.016
			(0.025)
$\text{L3. Ln\_AvgVolatility}_t$			-0.005
			(0.009)
$\text{L3. D\_Ln\_MarketCap}_{t-1}$			0.047
			(0.029)
L3. AbnTurn_{t-1}			-0.001
			(0.002)
Constant	-0.001	-0.001	0.000
	(0.003)	(0.003)	(0.003)
Observations	967	967	967
BIC	-5990.551	-5959.408	-5929.178

Table E2: ARMAX Models and BIC Values. This table presents the different lags of the control variables of the ARMAX model, and the BIC value associated with the different models. The dependent variable is the daily returns of the Nikkei 225 index. Standard errors are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Variables	(1)	(2)	(3)
L1.AR	0.466	0.476	0.534
	(0.617)	(0.623)	(0.628)
L1.MA	-0.391	-0.401	-0.458
	(0.617)	(0.624)	0.630
$\mathrm{L1.Previous\_Returns}_t$	0.008	0.007	0.016
	(0.025)	(0.026)	(0.031)
$\mathrm{L1.Ln\_AvgVolatility}_t$	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
$\mathrm{L1.D\_Ln\_MarketCap}_{t-1}$	0.023	0.018	0.007
	(0.034)	(0.038)	(0.037)
$\mathrm{L1.AbnTurn}_{t-1}$	0.003*	0.003*	0.003
	(0.005)	(0.002)	(0.002)
$\mathrm{L1.Sentiment}_t$	-0.002	-0.002	-0.003
	(0.005)	(0.005)	(0.005)
$\mathrm{L2.Sentiment}_t$		0.001	0.001
		(0.005)	(0.005)
$\mathrm{L3.Sentiment}_t$			-0.006
			(0.004)
Constant	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)
Observations	967	967	967
BIC	-5983.917	-5969.891	-5959.717

Table E3: ARMAX Models and BIC Values. This table presents the different lags of the variable representing the daily news sentiment of the ARMAX model, and the BIC value associated with the different models. The dependent variable is the daily returns of the Nikkei 225 index. Standard errors are presented in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01.

APPENDIX F: Alternative Specification Results

Dependent Variable: Returns_t (Open − to − Close)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Sentiment_t	0.041***		0.041***	0.041***		-0.013
	(0.003)		(0.003)	(0.003)		(0.015)
Ln_Activity_t		0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
$\text{Positive\_Sentiment}_t$					0.033***	
					(0.007)	
$\text{Negative\_Sentiment}_t$					0.046***	
					(0.005)	
Interaction_t						0.018***
						(0.005)
$\text{Previous\_Returns}_t$				-0.167	-0.171	-0.020
				(0.014)	(0.014)	(0.014)
$\text{Ln\_AvgVolatility}_t$				0.000	0.000	0.000
				(0.001)	(0.001)	(0.001)
$\text{D\_Ln\_MarketCap}_{t-1}$				-0.005	-0.004	-0.007
				(0.020)	(0.020)	(0.020)
AbnTurn_{t-1}				0.000	0.000	0.000
				(0.001)	(0.001)	(0.001)
Constant	0.001	0.001	0.001	0.001	0.001	0.000
	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Observations	1,493	1,493	1,493	1,493	1,493	1,493
R^2	0.230	0.001	0.230	0.231	0.232	0.245
Adjusted R^2	0.229	0.000	0.229	0.228	0.229	0.242

Table F1: Regression Results (Open-to-Close Returns). This table presents the results of the regressions. The dependent variable is the daily returns of the Nikkei 225 index calculated using open-to-close prices. The independent variables are the daily news sentiment scores and the natural logarithm of the number of articles per day. The control variables are the cumulative past 5-day returns of the Nikkei 225 index, the natural logarithm of the 5-day average volatility, the differenced natural logarithm of the previous day's market capitalization, and the previous day's abnormal turnover. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01.

Performance Metric	Benchmark Model	Extended Model	Diebold-Mariano Test
MSE	0.000060575	**0.000060528**	0.492
MAE	0.005700579	**0.005699685**	0.143
MAPE	**1.382099496**	1.43543981	-0.743
AIC	**-6623.427**	-6622.174	---
BIC	**-6584.441**	-6578.316	---

Table F2: Forecasting Accuracy of Benchmark Model and Extended Model (In-Sample & Open-to-Close Returns). This table presents the Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), AIC, and BIC values associated with the benchmark model and the extended model based on the in-sample data set. This analysis is based on the returns calculated using open-to-close prices. The value in bold indicates the preferred model by performance metric. The last column displays the test statistic of the Diebold and Mariano (1995) Test to see whether these metrics significantly differ from each other. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Performance Metric	Benchmark Model	Extended Model	Diebold-Mariano Test
MSE	**0.000054246**	0.000054290	-0.384
MAE	0.005693524	**0.005690096**	0.503
MAPE	**1.324709818**	1.332112405	-0.245

Table F3: Forecasting Accuracy of Benchmark Model and Extended Model (Out-of-Sample & Open-to-Close Returns). This table presents the Mean Square Error (MSE), Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE) values associated with the benchmark model and the extended model based on the out-of-sample data set. This analysis is based on the returns calculated using open-to-close prices. The value in bold indicates the preferred model by performance metric. The last column displays the test statistic of the Diebold and Mariano (1995) Test to see whether these metrics significantly differ from each other. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Dependent Variable: $Returns_t$ (Ln)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
$Sentiment_t$	0.041***		0.042***	0.042***		-0.006
	(0.003)		(0.003)	(0.003)		(0.018)
$Ln\_Activity_t$		0.000	0.001	0.001	0.001	0.001
		(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
$Positive\_Sentiment_t$					0.038***	
					(0.010)	
$Negative\_Sentiment_t$					0.044***	
					(0.006)	
$Interaction_t$						0.016***
						(0.006)
$Previous\_Returns_t$				-0.008	-0.009	-0.011
				(0.014)	(0.014)	(0.014)
$Ln\_AvgVolatility_t$				0.001	0.001	0.001
				(0.001)	(0.001)	(0.001)
$D\_Ln\_MarketCap_{t-1}$				0.005	0.005	0.004
				(0.032)	(0.032)	(0.032)
$AbnTurn_{t-1}$				0.000	0.000	0.000
				(0.001)	(0.001)	(0.001)
Constant	0.001***	-0.001	-0.001	-0.002	-0.002	-0.003
	(0.000)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Observations	1,493	1,493	1,493	1,493	1,493	1,493
R^2	0.124	0.000	0.126	0.127	0.127	0.132
Adjusted R^2	0.124	0.000	0.125	0.123	0.123	0.128

Table F4: Regression Results (Change in Ln (Open Prices)). This table presents the results of the regressions. The dependent variable is the daily returns of the Nikkei 225 index calculated using the change in the natural logarithm of open prices of the index between day t + 1 and day t. The independent variables are the daily news sentiment scores and the natural logarithm of the number of articles per day. The control variables are the cumulative past 5-day returns of the Nikkei 225 index, the natural logarithm of the 5-day average volatility, the differenced natural logarithm of the previous day's market capitalization, and the previous day's abnormal turnover. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Performance Metric	Benchmark Model	Extended Model	Diebold-Mariano Test
MSE	0.000112355	**0.000112325**	0.260
MAE	**0.008066326**	0.008070607	-0.757
MAPE	**1.458337781**	1.520011735	-1.848*
AIC	**-6026.444**	-6024.697	---
BIC	**-5987.458**	-5980.839	---

Table F5: Forecasting Accuracy of Benchmark Model and Extended Model (In-Sample & Change in Ln (Open Prices)). This table presents the Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), AIC, and BIC values associated with the benchmark model and the extended model based on the in-sample data set. This analysis is based on the returns calculated using the change in the natural logarithm of open prices of the index between day t + 1 and day t. The value in bold indicates the preferred model by performance metric. The last column displays the test statistic of the Diebold and Mariano (1995) Test to see whether these metrics significantly differ from each other. * p < 0.1, ** p < 0.05, *** p <0.01.

Performance Metric	Benchmark Model	Extended Model	Diebold-Mariano Test
MSE	0.000103195	**0.000102815**	2.576***
MAE	0.008020825	**0.008000568**	2.672***
MAPE	1.302642528	**1.270941027**	1.665*

Table F6: Forecasting Accuracy of Benchmark Model and Extended Model (Out-of-Sample & Change in Ln (Open Prices)). This table presents the Mean Square Error (MSE), Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE) values associated with the benchmark model and the extended model based on the out-of-sample data set. This analysis is based on the returns calculated using the change in the natural logarithm of open prices of the index between day t + 1 and day t. The value in bold indicates the preferred model by performance metric. The last column displays the test statistic of the Diebold and Mariano (1995) Test to see whether these metrics significantly differ from each other. * p < 0.1, ** p < 0.05, *** p <0.01.

APPENDIX G: Sub-Period Regression Results

	Dependent Variable: Returns_t					
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Sentiment_t	0.039***		0.039***	0.040***		-0.073***
	(0.005)		(0.005)	(0.005)		(0.028)
Ln_Activity_t		0.000	0.001	0.001	0.001	0.002
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\text{Positive\_Sentiment}_t$					0.044**	
					(0.020)	
$\text{Negative\_Sentiment}_t$					0.038***	
					(0.013)	
Interaction_t						0.044***
						(0.011)
$\text{Previous\_Returns}_t$				-0.015	-0.015	-0.022
				(0.023)	(0.023)	(0.023)
$\text{Ln\_AvgVolatility}_t$				0.000	0.000	0.000
				(0.001)	(0.001)	(0.001)
$\text{D\_Ln\_MarketCap}_{t-1}$				-0.023	-0.024	-0.016
				(0.055)	(0.055)	(0.057)
AbnTurn_{t-1}				-0.002	-0.002	-0.002
				(0.002)	(0.003)	(0.002)
Constant	0.000***	0.001	-0.002	-0.003	-0.003	-0.005
	(0.000)	(0.003)	(0.003)	(0.004)	(0.004)	0.004
Observations	480	480	480	480	480	480
R^2	0.109	0.000	0.110	0.114	0.114	0.133
Adjusted R^2	0.107	-0.002	0.107	0.103	0.101	0.120

Table G1: Regression Results (Pre-Covid Period). This table presents the results of the regressions during the pre-covid period before 2020. The dependent variable is the daily returns of the Nikkei 225 index (open-open). The independent variables are the daily news sentiment scores and the natural logarithm of the number of articles per day. The control variables are the cumulative past 5-day returns of the Nikkei 225 index, the natural logarithm of the 5-day average volatility, the differenced natural logarithm of the previous day's market capitalization, and the previous day's abnormal turnover. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Dependent Variable: Returns_t

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Sentiment_t	0.044***		0.044***	0.044***		0.010
	(0.007)		(0.007)	(0.006)		(0.028)
Ln_Activity_t		0.000	0.000	0.000	0.000	0.000
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\text{Positive\_Sentiment}_t$					0.035*	
					(0.018)	
$\text{Negative\_Sentiment}_t$					0.048***	
					(0.012)	
Interaction_t						0.011
						(0.009)
$\text{Previous\_Returns}_t$				0.005	0.005	0.002
				(0.027)	(0.027)	(0.027)
$\text{Ln\_AvgVolatility}_t$				0.000	0.000	0.000
				(0.002)	(0.002)	(0.002)
$\text{D\_Ln\_MarketCap}_{t-1}$				0.018	0.019	0.018
				(0.054)	(0.054)	(0.054)
AbnTurn_{t-1}				0.003	0.003	0.003
				(0.002)	(0.002)	(0.002)
Constant	0.001***	0.001	0.002	0.001	0.001	0.001
	(0.000)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)
Observations	487	487	487	487	487	487
R^2	0.126	0.000	0.126	0.129	0.130	0.132
Adjusted R^2	0.125	-0.002	0.123	0.118	0.117	0.120

Table G2: Regression Results (Covid Period). This table presents the results of the regressions during the covid period from 2020 to the end of 2021. The dependent variable is the daily returns of the Nikkei 225 index (open-open). The independent variables are the daily news sentiment scores and the natural logarithm of the number of articles per day. The control variables are the cumulative past 5-day returns of the Nikkei 225 index, the natural logarithm of the 5-day average volatility, the differenced natural logarithm of the previous day's market capitalization, and the previous day's abnormal turnover. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01.

Dependent Variable: Returns_t

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Sentiment_t	0.042***		0.042***	0.043***		-0.029
	(0.005)		(0.005)	(0.005)		(0.043)
Ln_Activity_t		0.001	0.001	0.001	0.001	0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\text{Positive\_Sentiment}_t$					0.040***	
					(0.015)	
$\text{Negative\_Sentiment}_t$					0.045***	
					(0.008)	
Interaction_t						0.022
						(0.014)
$\text{Previous\_Returns}_t$				-0.021	-0.021	-0.025
				(0.018)	(0.018)	(0.018)
$\text{Ln\_AvgVolatility}_t$				0.001	0.001	0.001
				(0.001)	(0.001)	(0.001)
$\text{D\_Ln\_MarketCap}_{t-1}$				0.003	0.004	-0.001
				(0.045)	(0.045)	(0.044)
AbnTurn_{t-1}				0.000	0.000	0.000
				(0.002)	(0.002)	(0.002)
Constant	0.002***	-0.003	0.000	-0.002	-0.002	-0.004
	(0.000)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Observations	526	526	526	526	526	526
R^2	0.141	0.001	0.142	0.145	0.145	0.150
Adjusted R^2	0.140	0.000	0.139	0.135	0.133	0.139

Table G3: Regression Results (Post-Covid Period). This table presents the results of the regressions during the post-covid period after 2021. The dependent variable is the daily returns of the Nikkei 225 index (open-open). The independent variables are the daily news sentiment scores and the natural logarithm of the number of articles per day. The control variables are the cumulative past 5-day returns of the Nikkei 225 index, the natural logarithm of the 5-day average volatility, the differenced natural logarithm of the previous day's market capitalization, and the previous day's abnormal turnover. Standard errors are Newey-West adjusted up to six lags for heteroskedasticity and autocorrelation and are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.