Received: January 26, 2021
Revised: February 25, 2021
Accepted: March 9, 2021



# อารมณ์ความรู้สึกของตลาดจากข่าวส่งผลกระทบต่อความผันผวนในตลาดหลักทรัพย์แห่งประเทศไทยอย่างไร?

# How does News-Based Sentiment Affect Stock Market Volatility in the Stock Exchange of Thailand?

สรรพศักดิ์ ชัชวาลย์[1*]

Sapphasak Chatchawan

## บทคัดย่อ

ดัชนีอารมณ์ความรู้สึกของตลาดถูกสร้างขึ้นมาจากหัวข้อข่าวที่เกี่ยวข้องกับตลาดหลักทรัพย์แห่งประเทศไทย และนำมาใช้ศึกษาถึงความสามารถในการพยากรณ์และผลของการเปลี่ยนแปลงของอารมณ์ความรู้สึกที่มีต่อความผันผวนในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้ข้อมูลรายวันระหว่าง 9/11/2560 – 8/2/2562

ผลการศึกษาพบว่า ดัชนีอารมณ์ความรู้สึกของตลาดจากพาดหัวข้อข่าวทางการเงินมีความสามารถในการพยากรณ์อัตราผลตอบแทนจากดัชนี SET, SET50 และ SET100 อย่างมีนัยสำคัญ นอกจากนั้น อารมณ์ความรู้สึกเชิงบวกจากข่าวของตลาดมีผลต่อความผันผวนในอัตราผลตอบแทนของ SET50 และ SET100 อย่างมีนัยสำคัญในช่วงที่ทำการศึกษา เมื่ออารมณ์ความรู้สึกเชิงบวกเพิ่มขึ้น ความผันผวนใน SET50 และ SET100 จะลดลง

**คำสำคัญ**: อารมณ์ความรู้สึกของตลาดจากข่าว ตลาดทุน การทำเหมืองข้อความ

## Abstract

The objective of the study is to investigate the effects of positive and negative sentiment from the financial news provider on volatility of equity returns during 9/11/2017- 8/2/2019. The stock market sentiment is extracted from a corpus of the financial news that is related to the Stock Exchange of Thailand through the computational linguistics.

[1]อาจารย์ประจำคณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์

[*]Corresponding Author e-mail: sapphasak@tbs.tu.ac.th

We find that the positive market sentiment significantly affects the volatility in SET50 and SET100 index returns during the period even after controlling for the trading volume. As the positive sentiment increases, the volatility of SET50 and SET100 returns decreases. However, the effects of the negative sentiment are not significant during this period.

**Keywords:** News-Based Market Sentiment Capital Market Text Mining

## Introduction

Financial firms enjoy the benefits of real-time newswire services from international news agencies, such as Thomson Reuters Eikon and Bloomberg. Newswires electronically distribute financial news, headlines and stories in the form of textual information and help market participants monitor real-time financial market conditions. Newswires spread up-to-the-minute news in every second and naturally create a voluminous amount of textual data. This paper asks the following empirical questions: How does the positive and negative market sentiment from financial news affect stock market volatility?

A growing body of literature empirically explores causal relations between financial news and stock market movements. The sentiment extracted from financial newspapers, such as *The Wall Street Journal*, *The New York Times* and *The Guardian*, also has the predictive power over stock market movements and corporate earnings (Tetlock, 2007; Tetlock, Saar-tsechansky & Macskassy, 2008; Ferguson et al., 2014). The content of news influences investors' decision making. The effect of the unfavorable tone of financial news on stock market prices is more pronounced during the economic recession and has more predictive power than the favorable tone (Garcia, 2013). In addition to the study of the printed newspapers, there exists a number of papers explore impacts of newswires on financial markets (Boudoukh, Feldman, Kogan & Richardson, 2012). The sentiment constructed from newswires has the predictive over forecasts the stock market changes better than macroeconomic variables (Uhl ,2014). The daily sentiment predicts stock market returns for very short period. However, the weekly is able to forecast returns for a longer period. Moreover, it seems that the sentiment predicts the direction of volatility better than stock prices (Allen, McAleer & Singh, 2013; Heston & Sinha, 2017; Atkins, Niranjan & Gerding, 2018). In addition to the stock market, news-based sentiment can also predict and explain movements in the term structure of interest rates (Gotthelf & Uhl, 2018) and in foreign exchange market (Uhl, 2017).

The paper attempts to quantify the daily market sentiment from Reuters Eikon's newswires by employing the technique in computational linguistics. The sentiment index is computationally extracted from a collection of more than 1,000 news headlines. The whole corpus consists of more than 20,000 words. This paper empirically analyzes the effects of the sentiment on market volatility. Apart from the pertinent literature, the paper is the first that constructs positive and negative sentiment from the newsfeed of professional media newswires, which provides a great amount of information in real time rather than using newspapers as a source of market sentiment.

## Literature Review

The paper contributes to the existing literature in three aspects. First, the study employs the user-defined dictionary method as presented in Loughran & McDonald (2011) to characterize the market sentiment into positive and negative sentiment. The technique is automated and replicable. Unlike Heston & Sinha (2017); Uhl (2014); Uhl (2017); Gotthelf & Uhl (2018) that use Thomson Reuters sentiment index, the market sentiment for SET is constructed from a collection of all news headlines which are directly linked to SET from Thomson Reuters Eikon's newswires. News headlines are retrieved from the news database by using Reuter Instrument Code. Thus, the approach is relatively flexible than Reuters sentiment index, which was released until 2012. Second, the study investigates the predictive power of news-based sentiment index on both stock market returns and the volatility. This is to confirm the findings in Allen, McAleer & Singh (2013); Heston & Sinha (2017); Atkins, Niranjan & Gerding (2018), that explore the same aspect, but in the U.S. markets. Third, to my knowledge, the paper is the first attempt to explore the effects of news-based sentiment on volatility, returns under the models that are augmented by the trading volume.

## Research Methodology

### Sources and Data Description

All financial data are collected from Thomson Reuters Eikon. The sample begins from 9/11/2017 to 8/2/2019, which is the longest trading period in which news headlines are available for scraping from Thomson Reuters Eikon's newswires. All financial news headlines were compiled by Routers Instrument Code (RIC): "BKstmad.BK". The major reason why the paper does not use Thomson Reuters Eikon sentiment score is that the score had been published until 2012. It seems that the service is no longer available during the period of this study. Thus, there are 308 observations of both financial and news-based sentiment variables that are particularly related to SET.

### Dependent Variables

There are three dependent variables, which are stock market returns, which are the returns from SET, SET50 and SET100 index. All dependent variables are computed from equation 1.

$$R_{i,t} = \ln\left(\frac{\mathrm{Y_{i,t}}}{Y_{i,t-1}}\right) \times 100, i = 1,2,3 \qquad (1)$$

$$Y_{1,t} = SET, Y_{2,t} = SET50, Y_{3,t} = SET100$$

In table 1, SET_RET, SET50_RET, and SET100_RET are stock market return variables and their descriptive statistics are provided in table 2. The mean values of three series are concentrated around 0.00. Considering other aspects of descriptive statistics, the data exhibit very similar statistical characteristics. This is not surprising when considering SET, SET50 and SET100.

**Table 1** Dependent variables in the study.

| Variable | Description |
|---|---|
| SET_RET | SET returns |
| SET50_RET | SET50 returns |
| SET100_RET | SET100 returns |

**Table 2** Descriptive statistics of dependent variables

| | SET_RET | SET100_RET | SET50_RET |
|---|---|---|---|
| Mean | -0.01 | 0.00 | 0.00 |
| Median | 0.05 | 0.06 | 0.06 |
| Maximum | 2.27 | 2.71 | 2.81 |
| Minimum | -2.42 | -2.65 | -2.60 |
| Std. Dev. | 0.71 | 0.80 | 0.80 |
| Skewness | -0.32 | -0.22 | -0.17 |
| Kurtosis | 4.12 | 4.26 | 4.31 |
| Jarque-Bera | 21.51 | 23.07 | 23.53 |
| Probability | 0.00 | 0.00 | 0.00 |
| Sum | -3.74 | -1.21 | 1.14 |
| Observations | 308 | 308 | 308 |

### Independent Variables

In the study, independent variables are classified into 2 groups, which are stock market data and market sentiment data, which are textual data from online financial news headlines. Financial data are trading volume in SET, SET50 and SET100. Trading volume data enter into EGARCH (1,1) in the form of percentage changes. The textual data in the study are the news-based market sentiment, the positive sentiment and the negative sentiment. All of which is extracted from the corpus of headline news scraped from Thomson-Reuters newswire. Table 3 gives the variable description and table 4 and 5 show descriptive statistics of each independent variable.

**Table 3** Independent variables

| Variable | Description |
|---|---|
| SET_VOLM | Percentage change of trading volume in SET |
| SET50_VOLM | Percentage change of trading volume in SET50 |
| SET100_VOLM | Percentage change of trading volume in SET100 |
| POS | Positive sentiment |
| NEG | Negative sentiment |

**Table 4** Descriptive statistics of financial market variables

| | SET_VOLM | SET50_VOLM | SET50_VOLM |
|---|---|---|---|
| Mean | 0.14 | 0.17 | -0.01 |
| Median | -0.89 | -1.10 | -2.12 |
| Maximum | 70.22 | 123.78 | 85.29 |
| Minimum | -65.66 | -89.31 | -72.07 |
| Std. Dev. | 20.35 | 32.13 | 27.77 |
| Skewness | 0.22 | 0.20 | 0.33 |
| Kurtosis | 3.83 | 3.41 | 3.10 |
| Jarque-Bera | 11.42 | 4.35 | 5.88 |
| Probability | 0.00 | 0.11 | 0.05 |
| Sum | 43.71 | 52.98 | -4.47 |
| Observations | 308 | 308 | 308 |

**Table 5** Descriptive statistics of news-based variables

| | POS | NEG |
|---|---|---|
| Mean | 0.03 | 0.05 |
| Median | 0.00 | 0.04 |
| Maximum | 0.36 | 0.29 |
| Minimum | 0.00 | 0.00 |
| Std. Dev. | 0.05 | 0.06 |
| Skewness | 2.22 | 1.05 |
| Kurtosis | 11.03 | 3.66 |
| Jarque-Bera | 1080.08 | 62.06 |
| Probability | 0.00 | 0.00 |
| Sum | 9.84 | 15.63 |
| Observations | 308 | 308 |

## Content

### Analytical preprocessing

Financial news headlines are textual data, which are unstructured in nature. That is, they contain numbers, punctuations, alphanumeric and non-alphanumeric characters together with or without ordering. Unstructured can be transformed to structured data, which are quantifiable, through the techniques in the textual analysis. In the natural language processing, textual pre-processing is the procedure of cleaning text and transformation of textual characters into the bag-of-words model. In this paper, the corpus is a daily collection of financial headline news from Reuters Eikon newswires. One day is equivalent to one document in the corpus. For example, in figure 1, there are two documents. One is 8-Feb-19 and another is 7-Feb-19. Text data pertaining to each trading day will be used in textual analysis. Figure 2 graphically illustrates the number of words in each headline.

| 21-Jun-18 | | | |
|---|---|---|---|
| 17:42:02 | RTRS | .SETI .VNI | SE Asia Stocks-Down; Philippines hits 1-1/2 year closing low |
| 17:25:32 | RTRS | .SETI 3938.T | Thai bourse targets 40 pct market cap growth by 2023, fueled by IPOs |
| 11:55:40 | RTRS | .SETI .VNI | SE Asia Stocks-Philippine shares slump 2 pct, move further into bear territory |

**Figure 1** A sample of financial news headlines from Reuters Eikon's newswire.

The ultimate goal of the analytical preprocessing is to reduce the dimensionality in the bag-of-words model. In this paper, the standard procedure of textual preprocessing adopted includes word concatenation, removing white space, numbers, punctuation marks, non-alpha numeric (symbols), unicode characters and a list of 'English' stopwords, stemming words with Porter stemming and, finally, tokenization. The financial corpus contains more than 20,000 terms before analytical preprocessing.

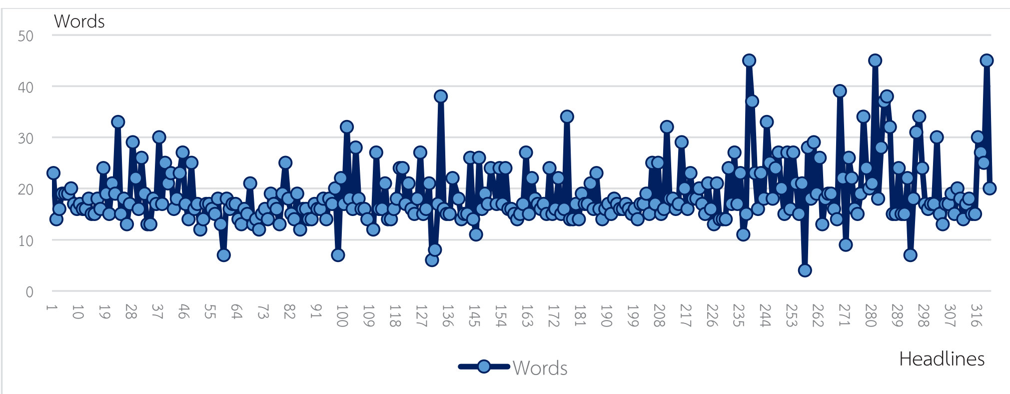


**Figure 2** The number of words in the headline news from Thomson Reuters Eikon newswire.

Some terms in financial headlines occur at a high frequency rate but they do not contribute much information. According to the concept of Shannon entropy, terms with high frequency carry little information with them. Therefore, I employ the term-frequency inverse document frequency (TF-IDF) weighting scheme as the means to drop those terms out of the document. TF-IDF is the weighting scheme, which is alternative to term-frequency (TF), based on how important words explain the document. Words that rarely occur in the document will be assigned high TF-IDF scores because they contain meaningful information. I will drop the terms that their TF-IDF values are zero out of the document and keep the terms that their values are greater than zero. Figure 3 illustrates Word Cloud of high frequency terms after trimming the corpus with the TF-IDF score. The size of the term in the word cloud represents the frequency of the occurrence of the terms in the corpus. In this corpus, the term, 'SET', occurs with the highest frequency. This is also not surprising because we collect the news that is relevant to the Stock Exchange of Thailand.

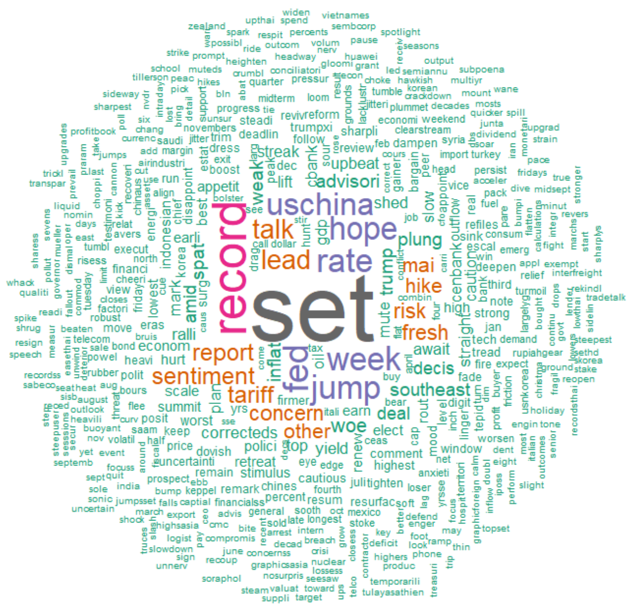


**Figure 3** Word Cloud of the corpus of headline news after TF-IDF scheme

### Dictionary method and market sentiment

The dictionary technique ensures the replicability and the consistency of the results throughout the process. The dictionary method is also used in Tetlock (2007); Tetlock, Saar-Tsechansky & Macskassy (2008); Loughran & McDonald (2011). Therefore, this study employs such techniques to extract the sentiment from the corpus of financial news headlines. In order to construct the market variables, a dictionary, which is a list of predetermined words in finance used by Loughran & McDonald (2011), is adopted as the researcher-defined dictionary. Wordlists of positive and negative terms are collected from Loughran-McDonald Sentiment Word Lists. There are 354 positive terms and 2,355 negative terms that make up the list. Words, such as 'success', 'transparency', 'tremendous', 'versatile', are regarded as positive words. As for the list of negative words, 'abandon', 'volatile', 'weaken' are examples. Positive sentiment and negative sentiment variables are constructed from these lists through the dictionary technique. To put it simply, the dictionary technique is a word counting technique.

Let $S_t \in \{POS_t, NEG_t\}$ be the market sentiment variable, which is classified as the positive and the negative sentiment. In EGARCH section, positive and negative sentiment variables are calculated according to equation (2) and (3).

$$POS_t = \frac{n_{h,t}^{POS}}{n_{h,t}} \tag{2}$$

$$NEG_t = \frac{n_{h,t}^{NEG}}{n_{h,t}} \tag{3}$$

"$t$" denotes a trading day. $POS_t$ and $NEG_t$ are the positive and the negative sentiment variables respectively. $n_{h,t}^{POS}$ is the number of the positive terms in headline $h$ on date "$t$". $n_{h,t}^{NEG}$ is the number of the negative terms in headline $h$ on date "$t$". $n_{h,t}$ is the total number of all terms in headline $h$. To illustrate the idea of how one can obtain the sentiment index from the headline, I will take the statement on 28/01/2019 as an example.

"SE Asia Stocks-Most rise; Malaysia leads gains SE Asia Stocks-Rise, Singapore leads gains on positive earnings."

There are 16 words in total. 'gains' and 'positive' are listed as positive words on the list. There are no negative words in this sample sentence. From equation (2), $n_{h,t}$ = 16 and $n_{h,t}^{POS}$ = 13.

$$POS_t = \frac{3}{16} = 0.1875$$

In this sentence, the positive sentiment is equal to 0.1875. Also, the negative sentiment can be obtained from equation 3. Figure 4 and 5 plot the time series of the negative sentiment and the positive sentiment. Thanks to the formula in equation (2) and (3), both variables cannot take the negative values.

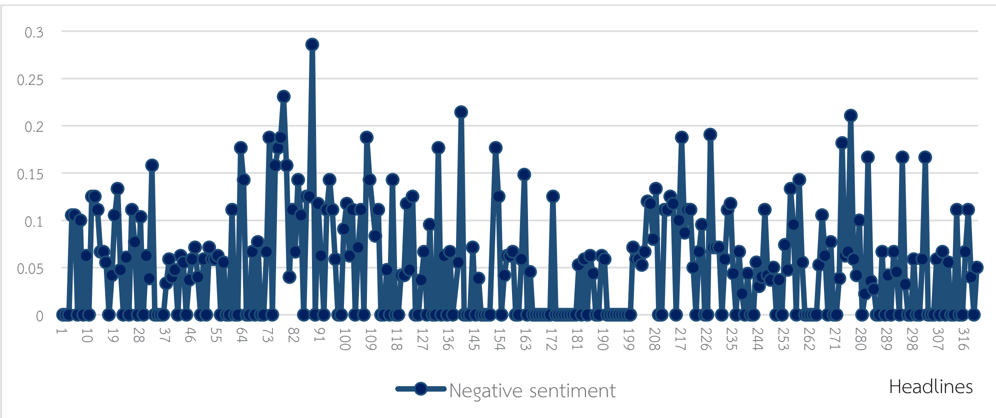


**Figure 4** Negative Sentiment index from headline news.

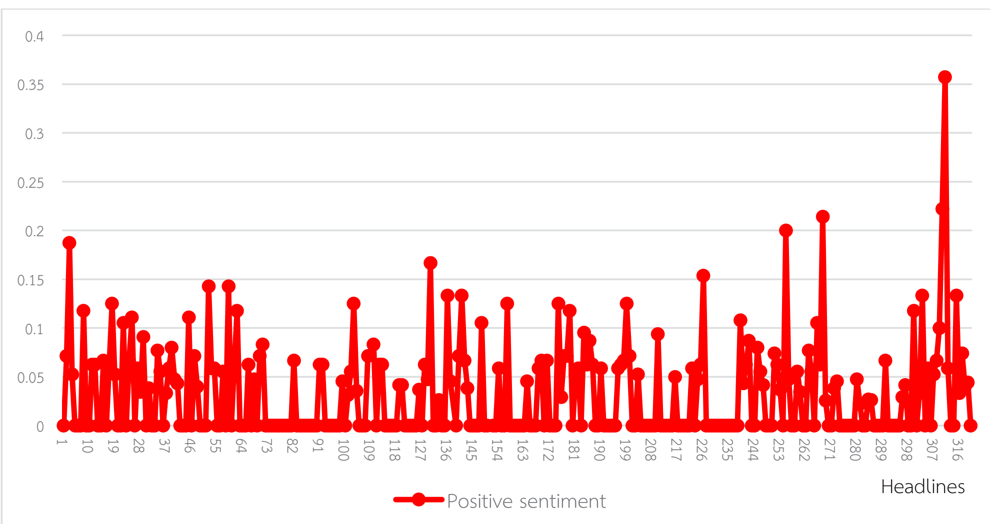


**Figure 5** Positive sentiment index from headline news.

**Table 6** Top 10 high frequency sentiment words.

| Negative words | Frequency | Positive words | Frequency |
|---|---|---|---|
| fall | 183 | gain | 118 |
| close | 52 | lead | 20 |
| low | 44 | rebound | 13 |
| fear | 28 | best | 6 |
| worry | 28 | highest | 4 |
| loss | 25 | strong | 4 |
| subdue | 24 | boost | 3 |
| drop | 23 | posit | 3 |
| lower | 19 | profit | 2 |
| concern | 14 | progress | 2 |

Top 10 high frequency negative and positive words have been shown in table 6. The term ‘fall’ is the highest frequency word for the negative terms while the term ‘gain’ occurs with the highest frequency for the positive terms. When counting words in the corpus of financial news headlines, I count and match the stemmed words of positive and negative terms that occur in the sentences.

Data are daily time series and pass the unit root test at 1% level of significance. That is, empirical results drawn from the data rule out the possibility of spurious relationship among variables. The unit root tests are presented in table 6.

**Table 7** Unit root tests

| Variables | t-statistic | P-value |
|---|---|---|
| POS | -14.54 | 0.00 |
| NEG | -14.53 | 0.00 |
| SET_RET | -16.90 | 0.00 |
| SET50_RET | -17.32 | 0.00 |
| SET100_RET | -17.06 | 0.00 |
| SET_VOLM | -15.53 | 0.00 |
| SET50_VOLM | -13.54 | 0.00 |
| SET100_VOLM | -14.43 | 0.00 |

**Empirical Strategy**

**Exponential General Autoregressive Conditional Heteroscedastic (EGARCH)**

Volatility cluster is ubiquitous in high-frequency financial data. In order to explore the effects of the news-based sentiment index, I employ EGARCH because of the three reasons. First, GARCH assumes that only the magnitude of unanticipated excess returns, not the positivity or negativity excess returns, influences the conditional variance. This assumption rules out the evidence that the volatility rises in response to bad news and falls in response to good news. Second, the non-negativity constraints imposed on GARCH parameters to ensure that the conditional variance is positive. Third, it is difficult to interpret the persistence in the conditional variance. Thus, Nelson (1991) proposes Exponential GARCH model (EGARCH) as an alternative to GARCH.

In this study, EGARCH model may have an advantage over GARCH because it ensures that the conditional variance is positive and allows for the asymmetric response of the volatility to good and bad news. EGARCH also incorporate the asymmetric leverage, which is consistent to the response of the market to the good and the bad news. This paper therefore employs Nelson's EGARCH as a work horse to analyze the effects of market sentiment.

EGARCH (1,1) Specification

Benchmark EGARCH (1,1)

$$y_t = \beta_0 + \beta_1 VOL_t + \omega_t \tag{4}$$

$$\omega_t \sim (0, h_t)$$

$$\log(h_t) = \gamma_0 + \gamma_1 \left( \frac{\omega_{t-1}}{\sqrt{h_{t-1}}} \right) + \gamma_2 \left( \left| \frac{\omega_{t-1}}{\sqrt{h_{t-1}}} \right| - \sqrt{\frac{2}{\pi}} \right) + \gamma_3 \log(h_{t-1}) + \alpha S_t$$

Trading-volume augmented EGARCH (1,1)

$$y_t = \beta_0 + \beta_1 VOL_t + \omega_t \tag{5}$$
$$\omega_t \sim (0, h_t)$$
$$\log(h_t) = \gamma_0 + \alpha VOL_t + \gamma_1 \left(\frac{\omega_{t-1}}{\sqrt{h_{t-1}}}\right) + \gamma_2 \left(\left|\frac{\omega_{t-1}}{\sqrt{h_{t-1}}}\right| - \sqrt{\frac{2}{\pi}}\right) + \gamma_3 \log(h_{t-1}) + \alpha S_t$$

I employ a parsimonious EGARCH (1,1) because the model is sufficient to capture the non-normality of the financial market data in this paper. In this study, two versions of the models are estimated as the benchmark models, which are equation (4) and (5). The difference between the models is that, in equation (5), the model is augmented with the trading volume in the variance equation. I control the trading volume in the models because the trading volume serves as the proxy of the arrival of information (Lamoureux & Lastrapes,1990) to the market. Thus, this may help examine the pure effect of the market sentiment.

In order to analyze the effect of market sentiment on stock market volatility, the sentiment enters into equation (4) and (5) in the variance equation through $S_t$ and $S_t \in \{POS, NEG\}$. Most importantly, the expected signs of parameters are $\alpha_{POS} < 0$ and $\alpha_{NEG} > 0$. This can be interpreted that the negative market sentiment induces high volatility. On the contrary, the positive market sentiment reduces the market volatility.

## Research Finding

### EGARCH (1,1)

Table 8, 9 and 10 show all estimated outputs of EGARCH (1,1). Table 8 presents the benchmark models without the positive and negative market sentiment. While table 9 presents the variants of the models that take into account the news-based market sentiment. Robustness checks have been shown in table 10. I choose to estimate the volume-augmented EGARCH (1,1) without the trading volume in the mean equation.

The main results discussed in the paper is in table 9, the coefficients of the negative market sentiment are insignificant in three equations. On the contrary, the coefficients of the positive market sentiment are statistically significant in all three equations with meaningful negative signs. The absolute effect of the positive sentiment is mostly pronounced in SET50 returns compared SET and SET100 returns. As the variance represents the volatility of market returns, from table 9, we can see that when the sentiment is increasingly positive, the market volatility tends to decrease significantly. Therefore, from November 2017 – February 2018, as the good news arrives, the overall volatility of market returns decreases.

**Robustness Checks**

Volume-augmented EGARCH (1,1) is employed for the robustness check in table 10. Under the different set up, some results exist and some are not. On the left panel in table 10, only the coefficient of the negative sentiment in SET100 returns is statistically significant and shows a positive sign. This can give rise to an explanation that bad news generates higher volatility in SET100 returns. However, the result is not that robust after comparing with the estimated results from EGARCH (1,1).

On the right panel in table 10, the coefficients of the positive sentiment in volume-augmented EGARCH (1,1) are statistically significant and display negative signs for SET and SET50 returns. The effect is not robust in SET100 returns.

KKBS JOURNAL of Business Administration and Accountancy
วารสารบริหารธุรกิจและการบัญชี มหาวิทยาลัยขอนแก่น
Vol. 5 No. 3 September - December 2021

**Table 8** Estimated results of benchmark EGARCH (1,1) and volume-augmented EGARCH (1,1) without the market sentiment

| | | Benchmark EGARCH (1,1) | | | Benchmark Volume augmented EGARCH (1,1) | |
|---|---|---|---|---|---|---|
| | $SET\_RET_t$ | $SET50\_RET_t$ | $SET100\_RET_t$ | $SET\_RET_t$ | $SET50\_RET_t$ | $SET100\_RET_t$ |
| **Mean equation** | | | | | | |
| Change in | -0.002 | -0.001 | -0.001 | - | - | - |
| Trading Volume | (0.002) | (0.001) | (0.001) | | | |
| Constant | 0.002 | 0.007 | 0.001 | 0.01 | 0.01 | 0.02 |
| | (0.04) | (0.048) | (0.05) | (0.04) | (0.04) | (0.04) |
| **Variance equation** | | | | | | |
| Trading volume | - | - | - | 2.15E-10*** | 1.08E-09*** | 6.56E-10*** |
| | | | | (0.00) | (0.00) | (0.00) |
| Constant | -0.10* | -0.05 | -0.07 | -3.5*** | -2.34*** | -2.49*** |
| | (0.06) | (0.05) | (0.05) | (0.48) | (0.28) | (0.28) |
| $\gamma_1$ | 0.04 | 0.02 | 0.03 | -0.02 | 0.05 | 0.09 |
| | (0.06) | (0.05) | (0.06) | (0.10) | (0.09) | (0.09) |
| $\gamma_2$ | -0.16*** | -0.13*** | -0.14*** | -0.02 | 0.05 | 0.03 |
| | (0.04) | (0.03) | (0.03) | (0.07) | (0.05) | (0.05) |
| $\gamma_3$ | 0.91*** | 0.92*** | 0.92*** | -0.08 | -0.57*** | -0.64*** |
| | (0.04) | (0.03) | (0.03) | (0.14) | (0.10) | (0.11) |
| Log likelihood | -317.81 | -354.93 | -353.081 | -314.510 | -348.12 | -348.95 |

*/**/*** denotes the level of significance at 10%,5% and 1%

**Table 9** The estimated results of effects of the negative and positive sentiment on market volatility.

| | Negative sentiment | | | Positive sentiment | | |
|---|---|---|---|---|---|---|
| | EGARCH (1,1) | | | EGARCH (1,1) | | |
| | ***SET_RET*** | ***SET50_RET*** | ***SET100_RET*** | ***SET_RET*** | ***SET50_RET*** | ***SET100_RET*** |
| **Mean equation** | | | | | | |
| Change in Trading Volume | -0.002 | -0.001 | -0.001 | -0.001 | 0.001 | -0.0002 |
| | (0.002) | (0.001) | (0.001) | (0.002) | (0.001) | (0.001) |
| Constant | -0.004 | 0.002 | -0.003 | 0.04 | 0.05 | 0.04 |
| | (0.04) | (0.05) | (0.05) | (0.04) | (0.04) | (0.04) |
| **Variance equation** | | | | | | |
| Positive Sentiment | - | - | - | -2.86** | -3.72*** | -2.92** |
| | | | | (1.13) | (1.19) | (1.07) |
| Negative Sentiment | -0.34 | -0.20 | -0.22 | - | - | - |
| | (0.44) | (0.39) | (0.41) | | | |
| Constant | -0.07 | -0.03 | -0.04 | -0.11 | -0.07 | -0.07 |
| | (0.06) | (0.05) | (0.05) | (0.09) | (0.08) | (0.08) |
| $\gamma_1$ | 0.02 | -0.03 | 0.01 | 0.10 | 0.12 | 0.10 |
| | (0.06) | (0.05) | (0.05) | (0.09) | (0.09) | (0.09) |
| $\gamma_2$ | -0.16*** | -0.14*** | -0.15*** | -0.13** | -0.12** | -0.12** |
| | (0.04) | (0.04) | (0.04) | (0.05) | (0.05) | (0.05) |
| $\gamma_3$ | 0.91*** | 0.93*** | 0.92*** | 0.84*** | 0.83*** | 0.86*** |
| | (0.03) | (0.03) | (0.03) | (0.06) | (0.06) | (0.06) |
| Log likelihood | -317.57 | -354.82 | -352.96 | -314.510 | -350.084 | -349.45 |

*/**/*** denotes the level of significance at 10%,5% and 1%

KKBS JOURNAL of Business Administration and Accountancy
วารสารบริหารธุรกิจและการบัญชี มหาวิทยาลัยขอนแก่น
Vol. 5 No. 3 September - December 2021

**Table 10** The estimated results of effects of the negative and positive sentiment on market volatility in volume augmented EGARCH (1,1).

| | Negative sentiment Volume Augmented EGARCH (1,1) | | | Positive sentiment Volume Augmented EGARCH (1,1) | | |
|---|---|---|---|---|---|---|
| | ***SET_RET*** | ***SET50_RET*** | ***SET100_RET*** | ***SET_RET*** | ***SET50_RET*** | ***SET100_RET*** |
| **Mean equation** | | | | | | |
| Constant | 0.01 | 0.03 | 0.02 | 0.03 | 0.05 | 0.03 |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) |
| **Variance equation** | | | | | | |
| Positive Sentiment | - | - | - | -5.46** | -5.88*** | -3.76 |
| | | | | (2.21) | (2.09) | (2.40) |
| Negative Sentiment | 1.48 | 1.97 | 2.13* | - | - | - |
| | (1.26) | (1.21) | (1.13) | | | |
| Trading Volume | 2.21E-10*** | 2.28E-10*** | 6.91E-10*** | 2.09E-10*** | 2.15E-10*** | 6.66E-10*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (129E-10) |
| Constant | -3.67*** | -3.56*** | -2.74*** | -3.32*** | -3.14*** | -2.37*** |
| | (0.53) | (0.50) | (0.28) | (0.51) | (0.48) | (0.31) |
| $\gamma_1$ | -0.003 | 0.01 | 0.11 | 0.03 | 0.03 | 0.09 |
| | (0.07) | (0.11) | (0.09) | (0.11) | (0.11) | (0.1) |
| $\gamma_2$ | -0.01 | 0.04 | 0.02 | 0.01 | 0.07 | -0.02 |
| | (0.07) | (0.07) | (0.05) | (0.07) | (0.07) | (0.06) |
| $\gamma_3$ | -0.09 | -0.10 | -0.67*** | -0.09 | -0.09 | -0.55*** |
| | (0.13) | (0.11) | (0.12) | (0.13) | (0.12) | (0.13) |
| Log likelihood | -310.39 | -343.98 | -347.32 | -307.64 | -341.29 | -347.21 |

*/**/*** denotes the level of significance at 10%,5% and 1%

## Discussion

The study constructs the news-based sentiment index from the corpus of financial news specific to the Stock Exchange of Thailand from Thomson-Reuters Eikon's media newswire from 2017- 2019. Dictionary technique is a tool that is used to extract the sentiment from the corpus. Not only the news-based sentiment, the study constructs the negative and the positive sentiment as well. The news-based sentiment consists of the positive sentiment and the negative sentiment. Thus, this allows the researcher to examine the effects of the positive and negative sentiment on market returns in Thailand during 2017-2019.

**Effects of the positive and negative sentiment on market volatility.**

The positive and the negative sentiment indices are what we obtain from the dictionary method when calculating the news-based sentiment index. In other words, in calculating the news-based sentiment index, we also obtain the positive and the negative sentiment variables. In this study, the effects of positive and negative sentiment index on volatility of stock market returns in SET, SET50 and SET100 are investigated. The parsimonious EGARCH (1,1) models with the trading volume are estimated and employed to explore the effect of positive sentiment (good news) and negative sentiment (bad news) on the volatility.

Results show that the coefficients of the positive sentiment in the variance equations are statistically significant in EGARCH (1,1) for SET, SET50 and SET100 returns. The positive sentiment tends to have larger effects on the volatility of SET50 returns than SET's and SET100's. Each coefficient exhibits the negative sign. However, this gives rise to some insightful interpretation. That is, when the positive sentiment increases, it lowers the volatility in stock markets. The findings seem to be contrast to the case of the negative sentiment in EGARCH (1,1). Coefficients of the negative sentiment in the variance equation are insignificant.

In order to provide further evidence of the findings in EGARCH (1,1), the volume augmented EGARCH (1,1) is employed as robustness checks. The robustness test confirms the effect of the positive sentiment on the volatility of SET50 and SET100 returns. The volatility of SET50 returns tends to be the most vulnerable to the positive sentiment from the financial newswire. However, with regard to the negative sentiments, even though signs of coefficients are positive, they are insignificant in the variance equation.

From table 6, we can observe that the number of negative terms is greater than the number of the positive terms. Thus, bad news is prevalent during this period of the study and when the good news arrives at the market, the investor may response more to the good news rather than the bad news. The results of the study may be dependent to the sample.

Even though the financial newswires are not publicly accessible, retail investors may benefit from the financial newswires through other local news outlets and analyst reports because local news providers, such as local financial newspapers, and financial analysts may collect the information that is related to the overall market from the newswires and spread the information to retail investors. That is, retail investors may access to the financial newswires indirectly.

**Policy recommendation**

Regulators and policy makers may be aware that news-based sentiment matters for stock market volatility. From what we have found in the study, the positive sentiment from the financial news headlines significantly reduces the volatility of stock market returns. This may be because the financial news headlines from the media newswire are one of the sources of information that investors use for their investment decisions.

## Conclusion

The market sentiment is extracted from headline news about the Stock Exchange of Thailand in Thomson Reuters newswire. The textual analysis, which is the dictionary method, quantifies the market sentiment from the corpus of headline news in the study. The news-based sentiment covers the period from 2017-2019 and the positive and the negative market sentiment variables are also calculated. The news-based sentiment index, together with the positive and the negative sentiment, contains the critical information for the market volatility.

In addition, the positive sentiment index has significant impact on the volatility of market returns. The more positive the sentiment, the lower the volatility in stock markets.

## Acknowledgement

The author would like to acknowledge the ample financial was awarded from the Stock Exchange of Thailand. The previous version of the paper is awarded the research grant from 2019 CM research innovation contest. The views expressed herein are those of the author and do not reflect the views of Thammasat Business School and the Stock Exchange of Thailand.

## References

Allan, D.E., McAleer, M.J. & Singh, A.K. (2015). Machine news and volatility: The Dow Jones industrial average and the TRNA real-time high-frequency sentiment series. In Gregoriou, G.N. (Ed.). *The handbook of high frequency trading. (pp.327-44).* Amsterdam: Elsevier.

Atkins A., Niranjan, M. & Gerding, E. (2018). Financial news predicts stock market volatility better than close price. *The Journal of Finance and Data Science*, *4*(2),120-137.

Boudoukh, F., Feldman, R. Kogan, S. & Richardson, M. (2012). *Which news moves stock prices? a textual analysis*. Cambridge: National Bureau of Economic Research.

Ferguson, N.J.et al. (2015). Media content and stock returns: the predictive power of press. *Multinational Finance Journal*, *19*(1), 1-31.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance, 68*(3), 1267-1300.

Gotthelf, N. & Uhl, M.W. (2018). News sentiment - a new yield curve factor. *Journal of Behavioral Finance*, *19*(3),31-41.

Heston, S.L. & Sinha, N.R. (2017). News versus sentiment: predicting stock returns from news stories. *Financial Analysts Journal, 73*(3),67-83.

Lamoureux, C. & Lastrapes, W. (1990). Heteroskedasticity in stock return data: volume versus GARCH effects. *The Journal of Finance, 45*(1), 221-229.

Loughran, T. & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65.

McLaren, N. & Shanbhogue, R. (2011). *Using internet search data as economic indicators*. Cambridge: National Bureau of Economic Research.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, *59*(2),347-370.

Tetlock, P. (2007). Giving content to investor sentiment: the role of media in the stock market. *Journal of Finance*, *62*(3), 1139–1168.

Tetlock, P., Saar-Tsechansky M. & Macskassy S. (2008). More than words: quantifying language to measure firms' fundamentals. *Journal of Finance*, *63*(3), 1437–1467.

Uhl, M.W. (2014). Reuters sentiment and stock returns, *Journal of Behavioral Finance*, *15*(4), 287-298.

Uhl, M.W. (2017). Emotions matter: sentiment and momentum in foreign exchange, *Journal of Behavioral Finance, 18*(3), 249-257.