SCIREA Journal of Mathematics

http://www.scirea.org/journal/Mathematics

July 15, 2019

Volume 4, Issue 4, August 2019

The Influence of Investors' Sentiments to Trading Volumes and Stock Returns with Information Uncertainty

Fang-Ming Hsu[1], Chien-Ho Liao[2,*], Qian-Feng Liu[3]

[1]Department of Information Management, National Dong Hwa University

1, Sec.2, Dahsueh Road, Shoufeng, Hualien, Taiwan

Email: fmhsu@gms.ndhu.edu.tw

[2]Department of Information Management, National Dong Hwa University, Taiwan

Email: jenhe12c@gmail.com

[3]Department of Information Management, National Dong Hwa University, Taiwan

Email: love2you2for@yahoo.com.tw

*: corresponding author

Abstract

Earning profit is always the paramount issue for investors. Many factors have been proved to affect stock returns, although information uncertainty is rarely discussed. This study examines how the effect of investors' sentiments on stock returns and trading volumes is moderated by information uncertainty. Text mining technology is used to analyze the opinions of investors on a popular financial forum to investigate their thoughts and feelings about investment. After that, structural equation modeling (SEM) is used to verify the hypotheses in weekly, monthly and quarterly intervals. Our findings demonstrate that investors' sentiments positively affect the stock returns only in the short-term (weekly) interval, and this relation is also moderated by information uncertainty in the short-term interval. Investors' sentiments negatively affect the trading volumes in the short-term and median-term intervals, and this relation is also moderated by the information uncertainty in the short-term and median-term intervals.

Keywords: Information Uncertainty; Stock Return; Sentiment Analysis; Text Mining.

1. Introduction

Because investors are afraid of information uncertainty (hereinafter referred to as "IU") before information disclosure, it appears that investors who control IU more, fare better in the market. Therefore, more and more scholars are engaged in issues related to IU. The most famous studies about IU are those authored by Jiang, Lee [1] and Zhang [2], and later studies often follow the proxies for IU these scholars suggested. Jiang, Lee [1] define IU as "value ambiguity". They propose stock markets respond slowly to information because of psychological biases such as overconfidence, and the larger the psychological bias, the slower the price response when a firm's value is more ambiguous. Zhang [2] also propose that both investor overconfidence and arbitrage costs are positively correlated with the level of IU. All of these scholars observe that greater IU causes greater price drift because of the decision biases of investors, but an evaluation of psychological factors like investor overconfidence remains incomplete.

We believe investors' behavior will be affected by their mood. To represent investors' psychology, we use posts on The Motley Fool, one of the best-known professional investment forums, to represent investment sentiment. We also exploit the opinion mining technique suggested by Serrano-Guerrero, Olivas [3] to evaluate investors' psychology. In addition, many kinds of text mining pre-processing techniques are used before extracting the sentiment.

Many previous studies have proved the relation between investors' sentiments and stock prices, as well as that between investors' sentiment and stock trading volumes, albeit without considering IU. However, some previous studies have shown there is only a weak relation between IU and stock returns. Also, the findings of some other research has provided inconsistent results regarding IU [4-6]. This study was undertaken as a way to re-think the relationships among investors' sentiment, stock returns, trading volumes, and IU. Therefore, we propose and examine the following two questions, both of which are evaluated in different intervals:

RQ1. Are stock returns affected by investors' sentiment and moderated by IU?

RQ2. Are stock trading volumes affected by investors' sentiment and moderated by IU?

2. Literature review

2.1 Relation between stock prices and investors' sentiment

Many previous works have already proved that extracting implications from financial news, annual reports, or announcements can help investors to forecast stock price movement. Tetlock [7] analyzes a popular Wall Street Journal column to predict market prices and finds that a high level of media pessimism predicts downward pressure on market prices. Butler and Kešelj [8] apply two Natural Language Processing (NLP) classification techniques in the analysis of corporate annual reports for financial forecasting over the next year. Schumaker and Chen [9] examine a predictive machine learning approach for financial news articles about S&P 500 constituent stocks on Yahoo! Finance. Li [10] examines the implications of the forward-looking statements in the Management's Discussion and Analysis (MD&A) section of corporate filings for future performance. Schumaker, Zhang [11] evaluate sentiment in financial news articles, and find subjective news articles are easier to predict in terms of price direction, especially for negative sentiment. Chan and Chong [12] propose a sentiment analysis engine to establish that sentiments expressed through financial texts are useful for predicting the trends of a stock market index. Ahmad, Han [13] explore the relation between media-expressed firm-specific tone and firm-level stock returns. Their findings show that news with a negative tone impacts firm-level stock returns briefly, with such impact being rapidly time-varying in nature.

Alongside analyzing formal information, some scholars have begun to concern themselves with the question of whether the collective emotion of investors can affect stock movement. Das and Chen [14] classify the messages of investors on Yahoo's message board according to an index of sentiment. Bollen, Mao [15] analyze the text content of daily Twitter feeds and achieve an accuracy rate of 86.7% in predicting the daily up and down changes in the closing values of the Dow Jones Industrial Average (DJIA). Yu, Duan [16] compare the effects of social media and conventional media on short term stock market performances. Li, Wang [17] propose a media-aware quantitative trading strategy utilizing sentiment information extracted from Web media to study the impact of the public mood of investors on stocks, along with the sentiment concerning firm-specific news. Other relevant works have been published by Piñeiro-Chousa, López-Cabarcos [18], Corea [19], Oliveira, Cortez [20], Guo, Sun [21] and so on. Therefore, our first hypothesis is: H1. Investors' sentiment affects stock returns.

2.2 Relation between trading volumes of stock and investors' sentiment

Some papers examine not only the relation between investors' sentiments and stock prices, but also that between investors' sentiments and trading volumes. Antweiler and Frank [22] analyze messages posted on Yahoo! Finance and Ranging Bull. They find the impact of investors' sentiments on trading volumes are strong while investors' opinions are divergent, whereas the effect of such sentiments on stock prices are weak. Tetlock [7] also proves high and low levels of pessimism induce high trading volumes. Li, Liang [23] indicate a strong correlation between sentiment and trading volume, and a similar correlation with asset price volatility. Price, Doran [24] also find that the sentiment of earnings conference calls is a significant predictor of abnormal returns and trading volumes. Liu [25] prove that a stock market is more liquid when sentiment indices rise, and market trading volume also increases when investor sentiment is higher. Kim and Kim [26] also test whether investor's sentiment has predictive power for stock returns, volatility, and trading volume. Therefore, our second hypothesis is: H2. Investors' sentiment affects trading volumes.

2.3 Related works regarding information uncertainty

Information uncertainty issues have been broadly discussed in financial fields, and research interest in the topic has grown since 2005. Jiang, Lee [1] are famous pioneers in this field. They define IU as "value ambiguity" and propose four proxies (firm age, return volatility, trading volumes, and equity duration) for it, in order to explore the relation between IU and stock returns. Their findings lead them to conclude that investing in high IU companies earns low future returns on average, but the effects of prices and earnings momentum among high IU companies are much stronger. Although the aforementioned paper has come under heavy criticism by Schultz [27], Jiang and Lee's proxies for IU are referred to by many scholars in later studies.

Zhang [2] is another famous scholar who explores relevant issues. He points out that greater IU causes higher expected returns following good news, and relatively lower expected returns following bad news. He also suggests five proxies for IU: firm age, analyst coverage, dispersion in analyst forecasts, stock volatility, and cash flow volatility.

Numerous papers regarding the impact of IU on stock momentum exist. Kim [28] explores the impact of IU on the January effect. He claims that without IU the January effect will disappear. Ciccone [29] also studies the January effect with IU. In his study, the forecast dispersion is used to proxy for IU. Berkman, Dimitrov [30] study stocks with differences of opinion (the definition is similar to IU) and short-sales constraints around earnings announcements. Francis, Lafond [31] explore explanations for post-earnings-announcement-drift (PEAD) anomalies by looking at how a rational investor responds to IU. Cheng, Dhaliwal [32] investigate whether the difficulty in assessing the extent of risk transfer affects IU of bank between securitizing banks and investors in asset-backed securities. Chang and Tsai [33] claim that the Mergers and Acquisitions of privately-held targets involve IU, which investors are more likely to overestimate. Chen, Lin [34] examine how IU surrounding IPO (initial public offering) firms influences earnings management and long-run stock performance. Gaspar and Massa [35] apply dispersion in analysts' forecasts as a proxy for IU to explore the relation between a firm's market power and the idiosyncratic volatility of its stock returns.

Besides some scholars point out IU enhances the effects focused on in their studies. Gerard [36] investigates the relation among earnings announcement abnormal returns, abnormal trading volumes, and subsequent returns for a large sample of European firms. The findings show that market surprise is positively related to future abnormal returns, especially when IU is high. Leippold and Lohre [37] find price and earnings momentum profits to be more pronounced for portfolios characterized by higher IU. Hillert, Jacobs [38] also find firms covered by the media exhibit significantly stronger momentum, and this effect is more pronounced for stocks with high uncertainty. All this would seem to suggest that IU has moderating effect. Therefore, we propose the following two hypotheses:

H3. The impact of investors' sentiment on stock returns is moderated by the level of IU.

H4. The impact of investors' sentiment on stock trading volumes is moderated by the level of IU.

In Fig1 our research model is shown.

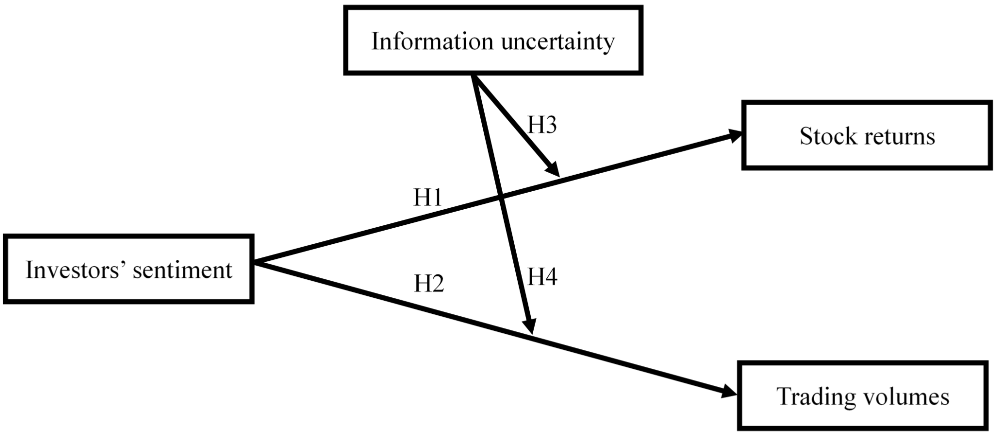

Figure 1. Research model

3. Methodology

3.1 Operationalization

In this section, we introduce the operationalization of our model. We extract investors' emotion toward articles from the forum known as The Motley Fool. We collect daily stock prices and trading volumes from the Center for Research in Security Prices (CRSP) database to calculate stock returns and the rate of change (ROC) of trading volumes. Also, to represent IU, we select the proxies of IU proposed by Jiang, Lee [1] and Zhang [2], and firm size, stock price volatility, trading volume volatility, and dispersion in analyst forecasts are used.

3.2.1. Investor's sentiment

We capture articles posted by investors from the forum of The Motley Fool by a self-developed crawler, and extract the emotion toward the articles using SentimentAnalyzer. Then we adjust the sentiment score range from 0~1 to -0.5~0.5. A score less than -0.1 is classified as a negative sentiment, and a score of more than 0.1 is classified as a positive sentiment, with the remaining scores falling outside this range being classified as neutral. Finally, we aggregate the sentiment scores (excluding the neutral sentiments) attaching to each stock in each of the intervals, respectively.

3.2.2. Stock returns

We capture the stock prices of S&P 500 constituent stocks from the CRSP database, and calculate the stock returns in different intervals. The formula of weekly interval is explained as follows, and monthly and quarterly intervals are similar.

- Stock returns in weekly interval = $(\text{closing price on the last day of this week} - \text{closing price on the last day of last week}) \,/\, \text{closing price on the last day of last week}$

3.2.3. Rate of change of trading volumes

We capture the trading volumes of S&P 500 constituent stocks from the CRSP database, and calculate the ROC of trading volumes in different intervals. The formula of weekly interval is explained as follows, and monthly and quarterly intervals are similar.

- Rate of change of trading volumes in weekly interval =

$$(\text{cumulative trading volumes of this week} - \text{cumula tivetrading volumes of last week}) \,/\, \text{cumulative trading volumes of last week}$$

3.2.4. Proxies for IU

To represent IU, this study uses the proxies for IU proposed by Jiang, Lee [1] and Zhang [2] including firm size, stock price volatility, trading volume volatility, dispersion in analyst forecasts, and classified the stocks into high IU and low IU portfolios. Each proxy is explained as follows:

3.2.4.1. Firm size

This study uses market value to represent firm size. A firm's market value is calculated by its closing price and outstanding shares in the end of the fiscal year.

$MV_i = P_i \times Q_i$

MV_i: the market value of stock i

P_i: the closing price of stock i in the end of fiscal year

Q_i: the outstanding shares of stock i in the end of fiscal year

3.2.4.2. Dispersion in analyst forecasts

This means the variance of the quarterly forecasting results of a company predicted by different analysts in the same quarter. We use the standard deviation of the EPS forecasting of a company by every analyst, as represented below:

$$S_{DISP_{i,q}} = \sqrt{\sum_{j=1}^{n} (EPS_{i,q,j} - \overline{EPS}_{i,q})^2/(n-1)}$$

$$\overline{EPS}_{i,q} = (1/n) * \sum_{j=1}^{n} EPS_{i,q,j}$$

$S_{DISP_{i,q}}$: the standard deviation of the EPS forecasting of company i in the q^{th} quarter

$\overline{EPS}_{i,q}$: the mean value of the EPS forecasting of company i in the q^{th} quarter

n: the number of analysts in that quarter

$EPS_{i,q,j}$: the EPS of company i predicted by the i^{th} analyst in the q^{th} quarter

3.2.4.3. Stock price volatility

In this paper, the stock price volatility is calculated by the standard deviation of the price of each stock in every weekly, monthly and quarterly interval. The formula of weekly interval is explained as follows, and monthly and quarterly intervals are similar.

- Weekly interval

$$S_{P_{i,w}} = \sqrt{\left(\sum_{j=1}^{n} (P_{i,w,j} - \bar{P}_{i,w})^2\right) / (n-1)}$$

$$\bar{P}_{i,w} = (1/n) * \sum_{j=1}^{n} P_{i,w,j}$$

$S_{P_{i,w}}$: the standard deviation of the price of stock i in the w^{th} week

$\bar{P}_{i,w}$: the mean value of the price of stock i in the w^{th} week

n: the number of trading days in that week

$P_{i,w,j}$: the closing price of stock i on the j^{th} day of the w^{th} week

3.2.4.4. Trading volume volatility

The trading volume volatility is calculated by the standard deviation of the trading volume of each stock in every weekly, monthly and quarterly interval. The formula of weekly interval is explained as follows, and monthly and quarterly intervals are similar.

- Weekly interval

$$S_{V_{i,w}} = \sqrt{\left(\sum_{j=1}^{n} (V_{i,w,j} - \bar{V}_{i,w})^2\right) / (n-1)}$$

$$\bar{V}_{i,w} = (1/n) * \sum_{j=1}^{n} V_{i,w,j}$$

$S_{V_{i,w}}$: the standard deviation of the trading volume of stock i in the w^{th} week

$\bar{V}_{i,w}$: the mean value of the trading volume in the w^{th} week

n: the number of trading days in that week

$V_{i,w,j}$: the trading volume of stock i on the j^{th} day of the w^{th} week

4. Results and discussion

4.1. Experimental procedure and data collection Results

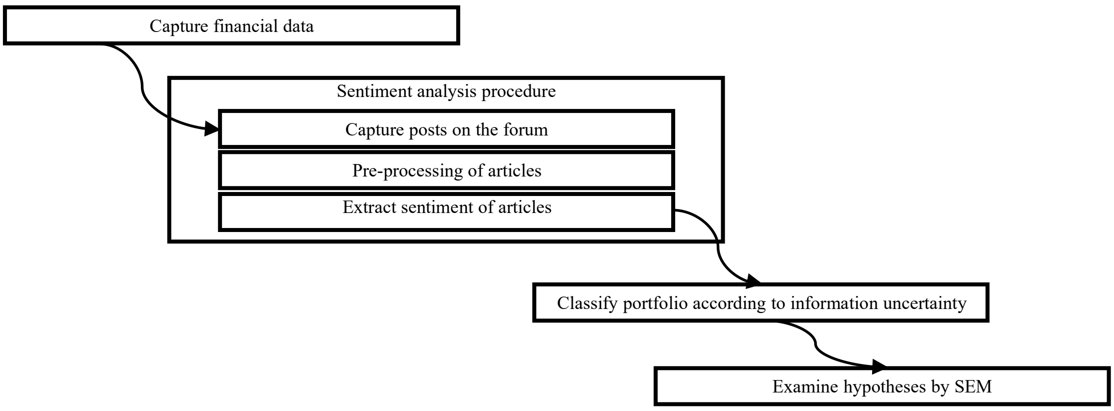

Figure 2. Experimental procedure

In this paper, we study S&P 500 index constituent stocks from Dec 29th, 2003 to Dec 31th, 2013, excluding the financial tsunami period from October 2007 to March 2009. This exclusion is because the effects of macroeconomics are larger than the effects of microeconomics in that period. Our experimental procedure is shown in Figure 2. As a starting point, we capture the daily closing prices and the daily trading volumes of S&P 500 index constituent stocks from the CRSP database, and gather the quarterly analyst forecasts from the Institutional Brokers' Estimate System (I/B/E/S). To calculate the market value of the stocks, we download the outstanding shares from the COMPUSTAT database. Our data collection results are shown in table 1.

For sentiment analysis, we download all the articles related to S&P 500 index constituent stocks posted by investors on the forum of The Motley Fool. After pre-processing process, each article is sent to SentimentAnalyzer to ask for its sentiment score. According to our proxies for IU, we classify the stocks into either high or low IU portfolios. And aggregate the sentiment scores of each portfolio by different proxies in different intervals. After all the data have been prepared, we examine our hypotheses by SEM.

Table 1. Data collection results

Data source	Data	No. of corps.	No. of records	Data duration
CRSP	Closing prices	498	1,190,827	2003.12.29~2013.12.31
	Trading volumes	497	1,188,303	2003.12.29~2013.12.31
COMPUSTAT	Outstanding shares	500	4,829	2004.12.1~2013.12.1
IBES	Analyst forecasts	315	271,680	2004.1.1~2013.12.31
The Motley Fool	Articles	457	105,081	2003.12.29~2013.12.31

Note: Our research duration excludes the financial tsunami period from October 2007 to March 2009.

4.2. Sentiment Analysis Process

Before sentiment analysis, there are still a few pre-processes that must be carried out. Four pre-processing steps of sentiment analysis are applied to these articles.

1. Remove strings with URLs, Email addresses, usernames, hash tags and HTML tags from the articles because these strings are not related to the emotion of investors.

2. Remove all the special characters, excluding the emoticons, because numerous scholars have suggested that emoticons can express the emotion of the author at the moment of posting [39-41].

3. Transform the emoticons to general expressions using the emoticon dictionary of SlangIt.com. For example, an emoticon “^_^” is transformed to the word “overjoyed”.

4. Transform instances of slang or abbreviations to general expressions using the slang word dictionary of SlangIt.com. For example, the slang term “BTW” is transformed to its meaning: “By the way”.

After the preprocessing stage of sentiment analysis, we use the opinion mining tool proposed by Serrano-Guerrero, Olivas [3] to extract the emotion of an article. For collecting scores automatically, we write a program to link to SentimentAnalyzer, submit every article sequentially, capture the response scores, and save them into our sentiment database. The range of response scores is 0 to 1.0. A sentiment score of less than 0.5 means a negative sentiment, whereas a score of more than 0.5 means a positive sentiment. To make it easy to observe positive/negative effects in SEM, we adjust the sentiment score range from 0 ~ 1.0 to -0.5 ~ 0.5 in this paper. A score of less than -0.1 means a negative sentiment, while a score of more than 0.1 means a positive sentiment. Other scores mean a sentiment is neutral. Finally, we aggregate the positive and negative sentiment scores attaching to each stock in each interval, but ignore the neutral sentiments. We also ignore some sentiment scores data on sporadic days which is less than our research interval. It is worth mentioning that not every company has an article every day because sometimes there is no relevant discussion about it for a whole week.

4.3. Results of portfolio classification by information uncertainty

Our study choose the proxies for IU proposed by Jiang, Lee [1] and Zhang [2] including firm size, stock price volatility, trading volume volatility, and dispersion in analyst forecasts. In their studies, the density of IU is divided into deciles, but there are too few articles regarding one-tenth of the S&P 500 constituent stocks discussed in the forum of The Motley Fool in one week to make this feasible. Therefore, we classify the first 100 companies as high IU portfolio, and the last 100 companies as low IU portfolio, and ignore surplus companies after sorting each proxy for IU corresponding to the companies, respectively, in each interval. The following are the results of each proxy.

4.4.1. Firm size

We use the market value of a company to represent firm size and sort the market value in ascending order every year from 2004 to 2013. Therefore, there are 10 orders in this proxy, and each order is classified into high IU or low IU portfolios.

4.4.2. Stock price volatility

We calculate the standard deviation of stock prices to express the volatility of stock prices. We sort the standard deviation in descending order in each interval. In our research period, there are 439 orders in the weekly interval, 101 orders in the monthly interval, and 33 orders in the quarterly interval. Finally, each order is classified into high IU or low IU portfolios.

4.4.3. Trading volume volatility

Similarly, we calculate the standard deviation of trading volumes to express the volatility of trading volumes, and sort the standard deviation in descending order in each interval. In the research period, there are 439 orders in the weekly interval, 101 orders in the monthly interval, and 33 orders in the quarterly interval. Finally, each order is classified into high IU or low IU portfolios.

4.4.4. Dispersion in analyst forecasts

Also, we calculate the standard deviation of analyst forecasts to express the dispersion in analyst forecasts, and sort the standard deviation in descending order in quarterly intervals. There are 33 orders in this proxy, and each order is classified into high IU or low IU portfolios.

4.4. Experiment Result and Discussion

We use AMOS 2.0 to execute SEM to test our hypotheses. After that, we discuss the global goodness of fit and the results of structural modeling, but omit measurement modeling because there is no latent variable.

4.5.1. The global goodness of fit

Three types of overall model fit measures suggested by Hair [42] are used in this study: absolute fit measures, incremental fit measures, and parsimonious fit measures.

An absolute fit measure is used to determine the degree to which the overall model (structural and measurement models) fits the sample data [42]. In this study, RMR, SRMR, RMSEA and GFI are used.

An incremental fit measure is used to compare the proposed model to a baseline model [42]. AGFI, CFI, NFI, RFI, IFI are used in this study.

A parsimonious fit measure is used to diagnose whether model over fits the data with too many coefficients [42]. A parsimonious fit measure is less important in our model because only three observed variables are used, so we ignore this measure.

The processes of testing our hypotheses are divided into two phases. In the first phase, we test hypotheses 1 and 2, and the results of the goodness of fit are shown in Table 2, and our goodness of fit of our model is acceptable.

In the second phase, we test hypotheses 3 and 4, and classify the portfolios by IU in different intervals. Therefore, 24 datasets are built and named from model 1 to 24. The results of the goodness of fit are shown in Table 3. We can see that the goodness of fit of each dataset is acceptable except for datasets 20 and 24. However, either of the results of the structural model in these two datasets are not significant, so we can ignore them.

Table 2. Results of the goodness of fit in our model in the first phase.

Intervals	Index of the goodness of fit
Weekly	RMR = 0.003, GFI = 0.992, AGFI = 0.954, and SRMR =0.0438
Monthly	RMR = 0.007, RMSEA = 0.053, GFI = 0.998, AGFI = 0.988, and SRMR =0.022
Quarterly	RMR = 0.04, GFI = 0.999, AGFI = 0.996, CMIN/DF = 4.392, SRMR = 0.0126, and RMSEA = 0.027

Table 3. Results of the goodness of fit in our model in the second phase.

Weekly interval			
Model No.	**IU proxies**	**IU degree**	**Indexes**
1	Firm size	High	RMR = 0.002, RMSEA = 0.04, GFI = 0.999, SRMR = 0.0189, AGFI = 0.991, and CMIN/DF = 3.858
2		Low	RMR = 0.005, GFI = 0.976, SRMR = 0.0776, and AGFI = 0.859
3	Dispersion in analyst forecasts	High	RMR = 0.002, RMSEA = 0.088, GFI = 0.995, SRMR = 0.0365, and AGFI = 0.968
4		Low	RMR = 0, RMSEA = 0, GFI = 1, SRMR = 0.0019, AGFI = 1, NFI = 0.981, CFI = 1, RFI = 0.942, and CMIN/DF = 0.081
5	Stock price	High	RMR = 0.012, GFI = 0.982, SRMR = 0.0671, and

	volatility		AGFI = 0.894
6		Low	RMR = 0.001, GFI = 0.988, SRMR = 0.055, and AGFI = 0.928
7	Trading volume volatility	High	RMR = 0.005, GFI = 0.988, SRMR = 0.0552, and AGFI = 0.928
8		Low	RMR = 0, RMSEA = 0.06, GFI = 0.997, SRMR = 0.0275, AGFI = 0.982, and CMIN/DF = 4.447
Monthly interval			
Model No.	**IU proxies**	**IU degree**	**Indexes**
9	Firm size	High	RMR = 0.002, RMSEA = 0, GFI = 1, SRMR = 0.0064, AGFI = 0.999, NFI = 0.849, and CMIN/DF = 0.249
10		Low	RMR = 0.013, GFI = 0.992, SRMR = 0.0461, and AGFI = 0.949
11	Dispersion in analyst forecasts	High	RMR = 0, RMSEA = 0, GFI = 1, SRMR = 0.0005, and AGFI = 1, NFI = 0.999, RFI = 0.996, CMIN/DF = 0.003
12		Low	RMR = 0.001, GFI = 0.992, SRMR = 0.0452, and AGFI = 0.951
13	Stock price volatility	High	RMR = 0.028, RMSEA = 0.054, GFI = 0.998, SRMR = 0.0238, and AGFI = 0.986
14		Low	RMR = 0, RMSEA = 0, GFI = 1, SRMR = 0.002,

			AGFI = 1, NFI = 0.991, CFI = 1, RFI = 0.972, and CMIN/DF = 0.044
15	Trading volume volatility	High	RMR = 0.013, RMSEA = 0.054, GFI = 0.998, SRMR = 0.0229, and AGFI = 0.987
16		Low	RMR = 0.001, RMSEA = 0.015, GFI = 0.999, SRMR = 0.0181, AGFI = 0.992, NFI = 0.815, CFI = 0.956, IFI = 0.973, and CMIN/DF = 1.137

Quarterly interval

Model No.	IU proxies	IU degree	Indexes
17	Firm size	High	RMR = 0.01, GFI = 0.99, SRMR = 0.0509, and AGFI = 0.938
18		Low	RMSEA = 0.099, GFI = 0.993, SRMR = 0.0412, AGFI = 0.96
19	Dispersion in analyst forecasts	High	RMR = 0.016, RMSEA = 0.08, GFI = 0.995, SRMR = 0.0351, and AGFI = 0.971
20		Low	RMR = 0.004, and GFI = 0.945
21	Stock price volatility	High	RMSEA = 0.011, GFI = 0.999, SRMR = 0.0133, AGFI = 0.996, IFI = 0.902, and CMIN/DF = 1.119
22		Low	RMR = 0.001, RMSEA = 0.078, GFI = 0.995, SRMR = 0.0342, and AGFI = 0.972
23	Trading	High	RMSEA = 0.02, GFI = 0.999, SRMR = 0.0123,

	volume		AGFI = 0.996, and CMIN/DF = 1.763
24	volatility	Low	RMR = 0.006, and GFI = 0.931

4.5.2. The results of structural model

Table 4 and 5 are the results of structural model in phase 1 and phase 2 respectively, and $\gamma 1$ and $\gamma 2$ are the codes of path coefficients which represent the impact of investors' sentiment on stock returns and on trading volumes, respectively, as shown in Figure 2.

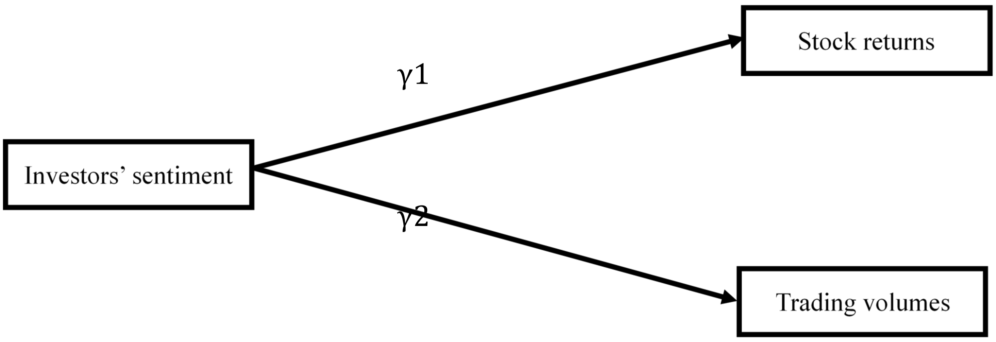

Figure 2. Codes of path coefficients

Table 4. Results of structural model in the first phase

Intervals	**Weekly**		**Monthly**		**Quarterly**	
code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**	**Path coefficients**	**p value**
$\gamma 1$	0.03	$.000^{***}$	0.002	0.818	-0.003	0.856
$\gamma 2$	-0.029	$.000^{***}$	-0.042	$.000^{***}$	-0.021	0.154

Note: *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$

In the first phase, without considering IU, we test H1: investors' sentiment affects stock returns, and H2: investors' sentiment affects trading volumes. The findings shown in Table 4 prove that stock returns can be affected by investors' sentiment ($\gamma1$) positively only in the short term (weekly). Trading volumes can also be affected by investors' sentiment ($\gamma2$) in both the short (weekly) and middle (monthly) term, but the relation is negative. However, in the long term (quarterly), neither of these relations ($\gamma1$ and $\gamma2$) are significant.

Table 5. Results of structural model in the second phase in the weekly interval

Intervals	**Weekly**			
IU proxies	**Firm size**			
Model No.	**1**		**2**	
IU Degree	**High**		**Low**	
Code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**
$\gamma1$	0.085	0.000***	0.007	0.485
$\gamma2$	0.001	0.976	-0.012	0.209
IU proxies	**Dispersion in analyst forecasts**			
Model No.	**3**		**4**	
IU Degree	**High**		**Low**	
Code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**
$\gamma1$	0.02	0.204	0.011	0.477

γ2	-0.012	0.428	-0.031	0.057

IU proxies	**Stock price volatility**			
Model No.	**5**		**6**	
IU Degree	**High**		**Low**	
Code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**
γ1	0.049	0.000***	0.001	0.942
γ2	-0.028	0.052	0.007	0.679

IU proxies	**Trading volume volatility**			
Model No.	**7**		**8**	
IU Degree	**High**		**Low**	
Code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**
γ1	0.034	0.000***	0.033	0.307
γ2	-0.032	0.001**	-0.055	0.085

Note: *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$

Table 6. Results of structural model in the second phase in the monthly interval

Intervals	**Monthly**			
IU proxies	**Firm size**			
Model No.	**9**		**10**	

IU Degree	High		Low	
Code of path coefficients	Path coefficients	p value	Path coefficients	p value
γ1	0.022	0.483	-0.005	0.759
γ2	-0.03	0.34	-0.042	0.006**
IU proxies	Dispersion in analyst forecasts			
Model No.	11		12	
IU Degree	High		Low	
Code of path coefficients	Path coefficients	p value	Path coefficients	p value
γ1	-0.001	0.954	0.021	0.355
γ2	-0.033	0.141	-0.008	0.705
IU proxies	Stock price volatility			
Model No.	13		14	
IU Degree	High		Low	
Code of path coefficients	Path coefficients	p value	Path coefficients	p value
γ1	0.001	0.964	0.039	0.09
γ2	-0.06	0.005**	-0.03	0.183
IU proxies	Trading volume volatility			

Model No.	15		16	
IU Degree	High		Low	
Code of path coefficients	Path coefficients	p value	Path coefficients	p value
γ1	0.000	0.982	0.032	0.437
γ2	-0.055	0.000***	-0.087	0.036*

Note: *: p < 0.05; **: p < 0.01; ***: p < 0.001

Table 7. Results of structural model in the second phase in the quarterly interval

Intervals	Quarterly			
IU proxies	Firm size			
Model No.	17		18	
IU Degree	High		Low	
Code of path coefficients	Path coefficients	p value	Path coefficients	p value
γ1	0.043	0.299	-0.013	0.571
γ2	0.005	0.895	-0.023	0.312
IU proxies	Dispersion in analyst forecasts			
Model No.	19		20	
IU Degree	High		Low	
Code of	Path coefficients	p value	Path coefficients	p value

path coefficients				
γ1	-0.004	0.839	0.001	0.965
γ2	-0.04	0.206	-0.007	0.833
IU proxies	**Stock price volatility**			
Model No.	**21**		**22**	
IU Degree	**High**		**Low**	
Code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**
γ1	-0.006	0.85	-0.016	0.607
γ2	-0.032	0.304	-0.028	0.369
IU proxies	**Trading volume volatility**			
Model No.	**23**		**24**	
IU Degree	**High**		**Low**	
Code of path coefficients	**Path coefficients**	**p value**	**Path coefficients**	**p value**
γ1	-0.007	0.743	0.035	0.485
γ2	-0.03	0.194	-0.026	0.607

Note: *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$

In the second phase, we test H3: the impact of investors' sentiment on stock returns is moderated by the level of IU, and H4: the impact of investors' sentiment on stock trading volumes is moderated by the level of IU. Below, we explain the results according to each interval, respectively.

4.5.2.1. Findings in the weekly interval:

As seen in Table 5, the experiment results for the weekly interval show that investors' sentiment positively significantly affects stock prices ($\gamma 1$) in high IU portfolios of firm size, stock price volatility, and trading volume volatility (models 1, 5, and 7). However, none of these relations ($\gamma 1$) in low IU portfolios are significant in each interval. It is proven that the impact of investors' sentiment on stock returns is moderated by the degree of IU only in the short term. This result is consistent with the first phase of our study. Therefore, hypothesis 3 is accepted only in the weekly interval.

Similarly, investors' sentiment significantly affects trading volumes ($\gamma 2$) negatively in high IU portfolios of trading volume volatility (model 7), but none of the relations ($\gamma 2$) in low IU portfolios are significant, either. Therefore, the impact of investors' sentiment on trading volumes is moderated by the degree of IU, and hypothesis 4 is accepted, too.

4.5.2.2. Findings in the monthly interval:

The experiment results in the monthly interval are shown in Table 6. In stock price volatility portfolios, investors' sentiment significantly affects trading volumes ($\gamma 2$) negatively in high IU portfolios (model 13), but that effect is not significant in low IU ones (model 14). So, we see that the impact of investors' sentiment on trading volumes is moderated by the degree of IU in stock price volatility in monthly interval. This result is consistent with our hypothesis 4.

In trading volume volatility portfolios (model 15 and 16), investors' sentiment significantly affects trading volumes ($\gamma 2$) negatively, both in high and low IU portfolios, and the path coefficients of low IU portfolios is higher than that of high IU ones. However, we can't say the impact of low IU on the relation between investors' sentiment and trading volumes is stronger than that of high IU. Because IU portfolios of trading volumes are classified by trading volume volatility, trading volume volatility of high IU portfolios must be larger. However, It can only be expressed that the accumulated trading volumes of high IU portfolios in the end of month are closer to that of the last month than the low IU portfolios. In other words, accumulated trading volumes of high IU portfolios could recover to original levels more quickly than those of low IU ones in a longer (monthly) interval. Investors' sentiment still triggers more trading volume variation in high IU portfolios than that in low IU ones. That means if investors remain in a good mood, accumulated trading volume could show a downward trend, but its volatility in the high IU portfolios is still larger than that of the low IU ones. If investors are brimming with negative emotion, accumulated trading volume becomes much higher, and trading volume volatility of the high IU portfolio could also be much larger than that of low IU ones. Therefore, the impact of investors' sentiment on trading volumes is still moderated by the degree of IU in this IU classification portfolio. Our hypothesis 4 is proved.

Interestingly, the result of firm size proxy in the low IU portfolio is significant (model 10), but the result in the high IU one is not. This is not in line with our hypothesis. The reason is that we calculate firm size by firm value, and some stocks with high firm value are also characterized by high stock price volatility or high trading volume volatility (e.g., AAPL, GOOG, AMZN, and FB).

4.5.2.3. Quarterly results:

As shown in Table 7, the moderating effects by the degree of IU on the impact of investors' sentiment either on trading volumes or on stock prices are not significant in the quarterly interval.

In sum, investors' sentiment affects stock prices only in the short term, and its impact is moderated by the degree of IU. Investors' sentiment affects trading volumes in the short and middle term, their relation is negative, and the impact of investors' sentiment is also moderated by the degree of IU. Finally, the degree of IU does not affect the impact of investors' sentiment either on trading volumes or on stock prices in the long term. The reasons for the above results are that most traders posting on the internet are noise traders, and they scramble to offload stocks when there is new information, especially in the case of bad news. Our results are in line with Tetlock [7].

5. Conclusion and suggestions

According to our collections and reviews of previous works, we discover IU may not directly impact on stock prices or trading volumes, but it moderates the relations between investors' sentiment and stock prices or trading volumes. Both Jiang, Lee [1] and Zhang [2] observe that greater IU will cause greater price drift because of investor decision biases, but no one has yet actually evaluated any psychological factor. Therefore, we utilize opinion mining technology to analyze the posts in the discussion board of The Motley Fool to represent investors' sentiment. First, we test the relations between investors' sentiment and stock returns or trading volumes. After that, we verify that IU will impact on the relations between investors' sentiment and stock returns or trading volumes. We also examine these hypotheses in different time intervals (weekly/monthly/quarterly).

The results show that investors' sentiment will impact on stock prices in the short term (weekly). This means investors making more positive comments about a stock will cause upward stock prices. On the other hand, more negative comments on a stock will cause downward stock prices. Moreover, IU will moderate the relations between investors' sentiment and stock prices in the short term (weekly), too. This implies higher IU stock will cause investors' sentiment to exert a more dramatic impact on stock prices. These results are consistent with Zhang [2].

Investors' sentiment will impact on trading volumes negatively in both weekly and monthly intervals. This is because investors will make extreme and unreasonable decisions to sell stocks when the market is full of pessimistic emotion. IU will moderate the relations between investors' sentiment and stock prices in both monthly and weekly intervals, but it is not significant in the quarterly interval. Besides, we also find the accumulated trading volumes of the high IU portfolios could recovery to original levels more quickly than that of low IU ones could in a longer (monthly) interval. Our results are in line with Tetlock [7]. Most investors talking on the internet are noise traders. They prefer to stop their loss and gains when any signal appears in the market.

Based on the above, we propose two suggestions. On the buyer side, people who like high risks and high returns can choose the stock with low market value, high stock volatility and high volume volatility (i.e. high IU stocks). Besides relying on general stock analysis tools, investors can choose stocks with positive sentiment in a stock forum as their next week targets, and keep away from those with negative sentiment and higher volumes. In contrast, short sellers can choose stocks with negative sentiment and higher volumes as their next week targets.

In future works, scholars can use or design better sentiment analysis tools. We believe better tools will achieve better results. To get a stronger sense of the relations between investors' sentiment and stock prices or trading volumes, shorter intervals can be used. Future scholars may be interested in the period of the so-called financial tsunami (which we omit), and other research objects can be applied and chosen for generalization purposes. Finally, we are very happy for more scholars or investors to examine our suggestions.

Acknowledgement

We would like to thank all the people who participated in this study. This research did not

receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Reference

[1] Jiang, G.H., C.M.C. Lee, and Y. Zhang, *Information uncertainty and expected returns.* Review of Accounting Studies, 2005. **10**(2-3): p. 185-221.

[2] Zhang, X.F., *Information Uncertainty and Stock Returns.* Journal of Finance, 2006. **61**(1): p. 105-136.

[3] Serrano-Guerrero, J., et al., *Sentiment analysis: A review and comparative analysis of web services.* Information Sciences, 2015. **311**: p. 18-38.

[4] Yan, Z.P. and Y. Zhao, *When Two Anomalies Meet: The Post-Earnings Announcement Drift and the Value-Glamour Anomaly.* Financial Analysts Journal, 2011. **67**(6): p. 46-60.

[5] Chen, Y.F. and H.N. Zhao, *Informed trading, information uncertainty, and price momentum.* Journal of Banking & Finance, 2012. **36**(7): p. 2095-2109.

[6] Cheema, M.A. and G.V. Nartea, *Momentum returns and information uncertainty: Evidence from China.* Pacific-Basin Finance Journal, 2014. **30**: p. 173-188.

[7] Tetlock, P.C., *Giving content to investor sentiment: The role of media in the stock market.* The Journal of Finance, 2007. **62**(3): p. 1139-1168.

[8] Butler, M. and V. Kešelj, *Financial forecasting using character n-gram analysis and readability scores of annual reports*, in *Advances in artificial intelligence*. 2009, Springer. p. 39-51.

[9] Schumaker, R.P. and H. Chen, *Textual analysis of stock market prediction using breaking financial news: The AZFin text system.* ACM Transactions on Information Systems (TOIS), 2009. **27**(2): p. 12.

[10] Li, F., *The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach.* Journal of Accounting Research, 2010. **48**(5): p. 1049-1102.

[11] Schumaker, R.P., et al., *Evaluating sentiment in financial news articles.* Decision Support Systems, 2012. **53**(3): p. 458-464.

[12] Chan, S.W.K. and M.W.C. Chong, *Sentiment analysis in financial texts.* Decision Support Systems.

[13] Ahmad, K., et al., *Media-expressed negative tone and firm-level stock returns.* Journal of Corporate Finance, 2016. **37**: p. 152-172.

[14] Das, S.R. and M.Y. Chen, *Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web.* Management Science, 2007. **53**(9): p. 1375-1388.

[15] Bollen, J., H.N. Mao, and X.J. Zeng, *Twitter mood predicts the stock market.* Journal of Computational Science, 2011. **2**(1): p. 1-8.

[16] Yu, Y., W. Duan, and Q. Cao, *The impact of social and conventional media on firm equity value: A sentiment analysis approach.* Decision Support Systems, 2013. **55**(4): p. 919-926.

[17] Li, Q., et al., *Media-aware quantitative trading based on public Web information.* Decision Support Systems, 2014. **61**: p. 93-105.

[18] Piñeiro-Chousa, J.R., M.Á. López-Cabarcos, and A.M. Pérez-Pico, *Examining the influence of stock market variables on microblogging sentiment.* Journal of Business Research, 2016. **69**(6): p. 2087-2092.

[19] Corea, F., *Can Twitter Proxy the Investors' Sentiment? The Case for the Technology Sector.* Big Data Research, 2016. **4**: p. 70-74.

[20] Oliveira, N., P. Cortez, and N. Areal, *Stock market sentiment lexicon acquisition using microblogging data and statistical measures.* Decision Support Systems, 2016. **85**: p. 62-73.

[21] Guo, K., Y. Sun, and X. Qian, *Can investor sentiment be used to predict the stock price? Dynamic analysis based on China stock market.* Physica A: Statistical Mechanics and its Applications, 2017. **469**: p. 390-396.

[22] Antweiler, W. and M.Z. Frank, *Is all that talk just noise? The information content of internet stock message boards.* The Journal of Finance, 2004. **59**(3): p. 1259-1294.

[23] Li, N., et al., *Network Environment and Financial Risk Using Machine Learning and Sentiment Analysis.* Human and Ecological Risk Assessment, 2009. **15**(2): p. 227-252.

[24] Price, S.M., et al., *Earnings conference calls and stock returns: The incremental informativeness of textual tone.* Journal of Banking & Finance, 2012. **36**(4): p. 992-1011.

[25] Liu, S.M., *Investor Sentiment and Stock Market Liquidity.* Journal of Behavioral Finance, 2015. **16**(1): p. 51-67.

[26] Kim, S.H. and D. Kim, *Investor sentiment from internet message postings and the predictability of stock returns.* Journal of Economic Behavior & Organization, 2014. **107**: p. 708-729.

[27] Schultz, P., *Discussion of "information uncertainty and expected returns".* Review of Accounting Studies, 2005. **10**(2-3): p. 223-226.

[28] Kim, D., *On the information uncertainty risk and the January effect.* Journal of Business, 2006. **79**(4): p. 2127-2162.

[29] Ciccone, S.J., *Investor Optimism, False Hopes and the January Effect.* Journal of Behavioral Finance, 2011. **12**(3): p. 158-168.

[30] Berkman, H., et al., *Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements.* Journal of Financial Economics, 2009. **92**(3): p. 376-399.

[31] Francis, J., et al., *Information uncertainty and post-earnings-announcement-drift.* Journal of Business Finance & Accounting, 2007. **34**(3-4): p. 403-433.

[32] Cheng, M., D.S. Dhaliwal, and M. Neamtiu, *Asset Securitization, Securitization*

Recourse, and Information Uncertainty. Accounting Review, 2011. **86**(2): p. 541-568.

[33] Chang, S.C. and M.T. Tsai, *Long-run performance of mergers and acquisition of privately held targets: evidence in the USA.* Applied Economics Letters, 2013. **20**(6): p. 520-524.

[34] Chen, S.S., et al., *Information Uncertainty, Earnings Management, and Long-run Stock Performance Following Initial Public Offerings.* Journal of Business Finance & Accounting, 2013. **40**(9-10): p. 1126-1154.

[35] Gaspar, J.M. and M. Massa, *Idiosyncratic volatility and product market competition.* Journal of Business, 2006. **79**(6): p. 3125-3152.

[36] Gerard, X., *Information Uncertainty and the Post-Earnings Announcement Drift in Europe.* Financial Analysts Journal, 2012. **68**(2): p. 51-69.

[37] Leippold, M. and H. Lohre, *International price and earnings momentum.* European Journal of Finance, 2012. **18**(6): p. 535-573.

[38] Hillert, A., H. Jacobs, and S. Muller, *Media Makes Momentum.* Review of Financial Studies, 2014. **27**(12): p. 3467-3501.

[39] Bifet, A., G. Holmes, and B. Pfahringer, *MOA-TweetReader: Real-Time Analysis in Twitter Streaming Data*, in *Discovery Science: 14th International Conference, DS 2011, Espoo, Finland, October 5-7, 2011. Proceedings*, T. Elomaa, J. Hollmén, and H. Mannila, Editors. 2011, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 46-60.

[40] Nagy, A. and J. Stamberger. *Crowd sentiment detection during disasters and crises*. in *Proceedings of the 9th International ISCRAM Conference*. 2012.

[41] Khan, F.H., S. Bashir, and U. Qamar, *TOM: Twitter opinion mining framework using hybrid classification scheme.* Decision Support Systems, 2014. **57**: p. 245-257.

[42] Hair, J.F., *Multivariate Data Analysis*. 1998: Prentice Hall.