Research Article

The Influence of Research Reports on Stock Returns: The Mediating Effect of Machine-Learning-Based Investor Sentiment

Xiaohong Shen [ID],[1] **Gaoshan Wang** [ID],[2,3] **and Yue Wang** [ID][2]

[1]*School of Computer Science and Technology, Shandong University of Finance and Economics, 7366 Erhuan Dong Road, Jinan 250014, China*
[2]*School of Management Science and Engineering, Shandong University of Finance and Economics, 7366 Erhuan Dong Road, Jinan 250014, China*
[3]*Institute of Digital Economy, Shandong University of Finance and Economics, 7366 Erhuan Dong Road, Jinan 250014, China*

Correspondence should be addressed to Gaoshan Wang; gaoshanwang@126.com

This paper investigates whether and how the research reports issued by securities companies affect stock returns from the perspective of investor sentiment in China. By collecting research reports and investor comments from a popular Chinese investor community, i.e., East Money, we derive two indices that represent the information contained in research reports: one is the attention of research reports and the other is the average stock rating given by research reports; then we develop an investor sentiment indicator using the machine learning method. Based on behavioral finance theory, we hypothesize that research reports have a significant effect on stock returns and investor sentiment plays a mediating role in it. The empirical analysis results confirm the above hypotheses. Specifically, the average stock rating given by research reports can better predict future stock returns, and investor sentiment plays a partial mediating role in the relationship between stock rating and stock returns.

1. Introduction

News reports, including research reports, play an important informational role in the stock market, and the effect of news reports has attracted interest from many researchers in the financial area [1–3]. According to the efficient market hypothesis (EMH), stock prices will change as new information comes. Then what information is and how it is incorporated into stock prices become interesting questions [4]. In essence, information is something that can reduce uncertainty, and from a statistical perspective, information refers to any observable results that can change people's beliefs based on conditional probability rule or Bayes' theorem. In practice, the term of information is often used interchangeably with the term of news or message. Nowadays, various kinds of news reports or messages (e.g., financial reports, research reports, corporate annual reports) flood social media and the Internet. They become the main information source for retail investors, who usually have limited time, money,

attention, and ability [5–8]. As for research reports issued by securities companies, although some people argue that they merely assemble public information that is already released, other people believe that research reports convey useful information and offer pertinent signals because securities companies have advantages at information collecting, processing, and analyzing [9]. Even news reports contain useful information; however, due to cognitive limitations and emotional biases, different investors pay different degrees of attention and make different reactions to the same information [10–12]. Therefore, how news reports are reflected in stock price is still not very clear [13]. Though there is a growing strand of literature investigating the effect of news reports on stock markets, the research results and conclusions are not consistent yet [6].

There are many types of information in the financial market. Some are private, and others are public [14]. According to its sources, information can be firm-released information, professional investment institution-released

information, government-released information, etc. [15]. And based on the content of information, information can be macroeconomic information, company financial and accounting information, research reports, investor comments, etc. Information also can be divided by its spread channel, e.g., the mainstream media websites, TVs, print media, social media, and news feeds [16]. Among various kinds of information in financial markets, research reports issued by securities companies are a very important kind of information, which usually contains stock ratings, rating adjustments, company earnings and price-earnings ratio forecasts, investment recommendations, and risk warnings, etc. Today, most people access financial information via social media and the Internet and there is no exception for individual investors [17, 18]. In China, one of the most popular financial information portal websites is East Money (https://www.eastmoney.com), which collects various kinds of information for investors. A featured function of East Money is that it collects all research reports issued by various securities companies and displays them in a simple and clear way, through which ordinary investors can easily find recommendations about specific stocks. According to a research report issued by Huajin Securities Co., Ltd on August 6, 2021, there are 14.62 million active users on the East Money platform per month, ranked top among all financial websites in China. Therefore, this paper mainly focuses on research reports issued by securities companies and displayed on the East Money platform.

Prior researchers have investigated the effect of research reports on stock returns [19]. However, the effect of research reports on the stock market remains an open challenge as well as the mechanism of how research reports influence the stock market [20]. In addition, in Chinese stock forums, people often complain about research reports and regard recommendations in research reports as inverse indicators [21]. Based on behavioral finance theory, this paper will investigate the effect and mechanism of research reports on stock returns in China from the perspective of investor sentiment. Specifically, we assume that research reports issued by securities companies have a significant effect on stock returns and investor sentiment plays a mediating role in it. We use the machine learning method to get an online investor sentiment index, which has advantages over the traditional direct or indirect sentiment indicators. The contributions of the paper lie in four aspects. First, prior researches mainly focus on either the relationship between research reports and stock returns or the relationship between investor sentiment and stock returns; however, few studies have examined the relations among the three variables together. To the best of our knowledge, this study is the first to explore the mediating effect of investor sentiment in the relationship between research reports published by securities companies and stock returns. Second, we develop a Chinese stock comment corpus, which contains 40,000 items of manually sorted stock comments. With the corpus, we use the supervised machine learning method to train a stock comment auto-classifier with high accuracy. Third, we use firm-specific sentiment and research report indices rather than market indices so that this study can use panel data.

Fourth, we focus on the behavior of retail investors, who constitute the main user group of the East Money financial platform.

The remainder of the paper is organized as follows. Section 2 provides the relevant literature. Section 3 proposes the hypotheses. Section 4 details the data and methodology. Section 5 reports the regression analysis results, robustness check results, and discussions. Section 6 concludes the paper.

2. Literature Review

Considerable practical and theoretical research results can be channeled into understanding the effect of research reports published by securities companies, which include a large number of prior researches that have long investigated the role of news reports and a wealth of studies that focus on investor sentiment.

2.1. The Impact of News Reports on the Stock Market. Fama [22] proposed the efficient market hypothesis (EMH) which implies that stock prices at any time fully reflect all available information, including public news reports. In this sense, the prediction of stock prices based on news is unsustainable [23]. However, the relevant news reports cannot be available to everyone at the same time due to their different contents, types, and transmission channels; moreover, different investors face different cost structures and have different reactions to new information [7]. As a result, there exists a lag period between the time information is introduced and when the market would correct itself, which makes short-term price prediction feasible [24–26]. In fact, both EMH and modern behavioral finance believe that the volatility of the stock market comes from the release, dissemination, and utilization of information, although they hold different views on how information shapes stock movements [1]. With more and more evidence emerging that stock prices cannot incorporate new information instantly, it is now widely accepted that news reports, whether they contain new information or only capture stale information, have significant effects on stock returns and volatility, and can help predict stock movement [27–30]. Tetlock [31] found that media tone has an impact on stock returns. Glasserman and Mamaysky [32] found that an increase in "unusual" news with negative sentiment predicts an increase in stock market volatility. El Ouadghiri et al. [33] found that media attention has a significant effect on stock returns. Though there is consensus that news reports have predictive power for stock returns, there is still a lack of understanding of the specific impacts that different kinds of news have on the stock market [34]. As an important kind of news, research reports should also influence the stock market; this study will further investigate how research reports issued by securities companies affect stock returns by following prior researches [35, 36].

2.2. The Role of Investor Sentiment. A growing body of literature examines the effect of investor sentiment on stock returns [37–42]. In the last decade, with the development of

social media and thanks to the advances in NLP (natural language processing) techniques and machine learning, online sentiments of retail investors have drawn more and more attention of researchers and practitioners [43–46]. Bollen et al. [47] derived moods of individual investors from Tweets and found that some mood dimensions have prediction powers while others do not. They also found that investor sentiment in terms of positive vs. negative mood measured by the OpinionFinder (https://mpqa.cs.pitt.edu/opinionfinder/) tool does not influence DJIA (Dow Jones Industrial Average). Azar and Lo [48] argued that traditional sentiment indicators based on survey data and transaction data have two drawbacks and the rise of social media allows people to overcome the drawbacks and measure the sentiment of a large number of individuals in real time. They proved that user messages do have useful information about future asset prices although social media data is generated by individual users and not investment professionals. Huang et al. [49] found that negative online individual investor sentiment is negatively correlated with market returns while positive sentiment does not have a significant effect on market returns. Lv et al. [50] derived online investor sentiment indicators from investor comments and found that the SSE (Shanghai Stock Exchange) Composite Index returns are significantly affected by online investor sentiment. While many researchers focused on the effect of investor sentiment on stock returns, some other researchers explored the influencing factors of individual investor sentiment. Yang et al. [51] found that news media has significant effect on investor sentiment. Gan et al. [52] pointed out that news reports such as macroeconomic announcements and corporate announcements can change investor sentiment.

2.3. Summary. By summarizing the above studies, we find that news reports could affect both individual investor sentiment and stock returns and may be released intendedly. However, prior researchers mainly focused on the whole market news sentiment or attention; the effects on news reports about specific domains, especially the research reports issued by securities companies, are far from being fully investigated. In China, more than 90% of investors are retail investors, whose investment decisions mainly rely on public information [44, 53]. Chen et al. [53] pointed out that the Chinese market is dominated by retail investors and financial information is always manipulated. Securities companies probably want to influence individual investor sentiment and subsequently affect stock movements by releasing research reports. In China, whether this is de facto the case is not fully understood. What is more, in Chinese stock forums, many investors express their dislike of research reports. Therefore, this paper will address the issue by making an empirical analysis of the relations among research reports, individual investor sentiment, and stock returns. Different from most similar researches on this topic, which still adopt sentiment dictionary to develop investor sentiment index [44, 54, 55], we adopt the supervised machine learning method that can achieve higher accuracy [56].

3. Hypothesis Development

3.1. The Effect of Research Reports on Stock Returns and Investor Sentiment. As it is well known, news reports play an influential role as information providers in stock markets [57–59]. According to agenda-setting theory, mass communication theory, and attention hypothesis, news reports serve as the foundation for market opinions and trading behaviors and thus can influence the stock market [23, 60]. Prior researches mainly investigated the media effect from two aspects. One is media coverage, i.e., the news volume covering specific stocks [23, 61]. The other is media content, i.e., media tone or news sentiment [62–64]. Huang and Zhang [65] found that media coverage has a significant effect on stock returns; in particular, stocks that are poorly covered by industry-specific media earn significantly higher future returns than stocks that are heavily covered. Wu et al. [66] found that positive news texts have a positive effect on stock returns, whereas negative news texts have a negative effect. Khedr et al. [67] found news sentiment has a significant effect on changes in stock prices.

As a very important kind of news report, research reports published by securities companies, which mainly focus on financial and market-related attributes to the stock price with the purpose of predicting stock prices and giving buy-sell-hold recommendations, play a prominent role in capital markets [11, 68, 69]. Securities companies have richer information channels and more research capabilities than ordinary financial news reporters, so for retail investors, research reports released by securities companies are important reference objects of investment decisions [70]. Prior researches showed that both the recommendations in the research reports and the number of research reports covering certain stock will influence stock returns. Vukovic et al. [35] pointed out that positive stock recommendations are likely to cause significant positive returns whereas negative recommendations lead to negative returns. Bouteska and Mili [9] provided evidence that shows stock returns are significantly related to recommendations in research reports. Ishigami and Takeda [71] showed that markets react strongly to the ratings in research reports and the numbers of research reports. Song et al. [70] found that research report attention and ratings in research reports can improve the prediction accuracy of excess returns.

Besides the relationship between research reports and stock returns, how research reports influence stock returns has also drawn people's attention. Research reports, which usually provide stock recommendations and earnings forecasts, directly affect the expectations and sentiments of investors, who interpret the research reports in their own unique manner and then make investment decisions [11]. In stock markets, even if research reports disclose all information in a truthful and timely manner, investors are limited by their cognition and capability and thus have heterogeneous beliefs about a company's prospects. Meanwhile, some investors exhibit overconfidence while others exhibit underconfidence after they read research reports. Therefore, investor sentiment may play a mediating role in the relationship between research reports and stock returns. Kim

et al. [5] found that investor sentiment responds to announcements of changes in analysts' recommendations significantly. Therefore, from the above analysis, we propose the following hypotheses:

H1: stock returns are significantly affected by research reports published by securities companies

H2: online investor sentiments are significantly influenced by research reports published by securities companies

3.2. The Effect of Investor Sentiment on Stock Returns.

A growing number of studies have shown that investor sentiment has a significant effect on stock returns [72, 73]. The sentiment is closely related to emotions. According to psychological theory, emotions refer to a complex psychological state such as happiness, sadness, anger, fear, surprise, and disgust, while sentiment can be regarded as a mental attitude that is created through the existence of the emotion or a thought that has been influenced by emotion [55, 74]. Behavioral financial scientists believe that emotions influence investors' judgments and decision-making, and so does sentiment [75]. Although the reasons and mechanisms of the effect of investor sentiment on stock returns are not very clear, researchers agree that the behaviors of investors are influenced by their emotions, cognitive limitations, and biases so that their behaviors cannot be completely rational [75–78]. To investigate the effect of investor sentiment on stock returns, researchers have used various kinds of measures, which can be divided into three categories, i.e., survey-based indicators or direct measures (e.g., the American Association of Individual Investors, Investors Intelligence), market-data-based indicators or indirect measures (e.g., adjusted turnover rate, buy-sell imbalance, BW index), and text-based indicators (e.g., investor sentiment derived from social media) [20, 54, 79]. Although investor sentiment measures developed in different ways have diverse effects on stock returns, a large number of researchers agree that investor sentiment has a significant effect on stock returns [80]. Shi et al. [81] found that investor sentiment has a significant effect on stock return using Consumer Confidence Index as a proxy of investor sentiment. Gao and Liu [82] constructed investor sentiment using market data and found that investor sentiment has a significant effect on stock returns. Sun et al. [44] derived investor sentiment from online posts and found that online investor sentiment has a significant effect on stock returns. Selin and Tas [42] found that Twitter sentiment has a significant effect on stock returns. Moreover, Bu [72] found that the indirect measures' predictive power on future stock return is remarkably higher than that of the direct measures. Anand et al. [73] found that a text-based sentiment measure is better than direct and indirect sentiment variables. Xu et al. [83] found that, among sentiment indices derived from social media, Internet news, and traditional newspapers, social media sentiment has the best predictive power. Therefore, we propose the following hypothesis:

H3: online investor sentiment influences stock returns significantly

Hence, based on hypothesis 2 and hypothesis 3, we further propose the following hypothesis:

H4: online investor sentiment plays a mediating role in the relationship between research reports and stock returns

4. Data and Methodology

4.1. Data.

Our data comes from East Money (https://www.eastmoney.com) and China Stock Market and Accounting Research Database (CSMAR), ranging from August 1, 2014, to July 31, 2018. First, we develop a web crawler to collect research reports, including reading volume, reply number, report title, stock rating (or recommendation), securities company, and issued date, as shown in Figure 1, of the constituents of CSI300 (China Securities Index) index on East Money using Python programming language. Second, we develop another web crawler to collect investor comments of the constituents of CSI300, including reading volume, reply number, title, author, and posting date. Finally, we download transaction data of the constituents of CSI300, including market value, book-to-market ratio, and BETA.

4.2. Measures of Research Reports.

We measure the influence of research reports from two aspects, i.e., attention and stock rating. Securities companies' attention to a stock reflects the popularity of the stock in the stock market. The stock ratings given by securities companies indicate their judgments on the future trend of the stock.

According to the study of Song et al. [70], we adopt formula (1) to calculate the attention of research reports (Att) and formula (2) to measure the average rating of a stock (Rat).

$$\text{Att}_{i,t} = \frac{N_{i,t}}{N_t}, \tag{1}$$

$$\text{Rat}_{i,t} = \frac{\sum_j R_{i,j,t}}{T_t}, \tag{2}$$

where $\text{Att}_{i,t}$ denotes the research report attention to stock i on day t, $N_{i,t}$ denotes the number of research reports that cover the stock i on day t, N_t indicates the total number of research reports on day t, $\text{Rat}_{i,t}$ indicates the average rating of stock i on day t, $R_{i,j,t}$ denotes the rating of stock i given by securities company j on day t, and T_t indicates the total number of research reports that give ratings to stock i on day t.

Generally, there are five levels of stock ratings, i.e., Buy, Overweight, Hold, Underweight, and Sell [5, 9, 84]. In reality, different securities usually use different terms to indicate the same rating. And there are 60 different kinds of rating terms in the research reports we collect from East Money. After categorizing and merging these terms, we use a 5-point scale to measure stock ratings, as shown in Table 1.

read	Comment	title	Rating	institution name	Post time
8719	8	Financial report comments: performance is in	Buy in	Guosen Securities Co.,	2021-08-13
7069	18	Supply chain management capabilities are hig	Buy in	Essence Securities Co.,	2021-08-12
4993	11	Multi-services work together, the annual per	Buy in	West China Securities	2021-08-11
12339	twenty	Rapid growth in performance	Buy in	Orient Securities Co.,	2021-05-06
8953	18	Intelligent manufacturing lays the foundatio	Buy in	Guosheng Securities Co.,	2021-05-05
8168	11	Comment on the first quarterly report of 202	Buy in	Essence Securities Co.,	2021-05-05
6448	5	Q1 earnings exceed expectations, and focus o	Overweight	Shanxi Securities Co.,	2021-05-05
9650	36	21Q1 performance exceeded expectations, and	Buy in	Tianfeng Securities Co.,	2021-04-30
8113	9	High performance in the first quarter repor	Buy in	West China Securities	2021-04-29
13932	32	Steady performance, determine the new energy	Buy in	Orient Securities Co.,	2021-04-05
6133	9	The full-year performance is in line with ex	Buy in	Essence Securities Co.,	2021-04-02
5480	2	The turning point in the fourth quarter rebo	Buy in	Tianfeng Securities Co.,	2021-04-01
4371	3	Financial report comment: Cloud computing co	Buy in	Guosen Securities Co.,	2021-03-31
8422	18	Intelligent manufacturing + industrial Inter	Buy in	Guosheng Securities Co.,	2021-03-30

Figure 1: Research reports on East Money.

Table 1: The score of stock rating.

Rating	Score	Relative frequency (%)
Strong buy/buy	2	50.97
Outperform/overweight/accumulate/add/moderate buy	1	18.84
Hold/neutral/equal weight	0	29.84
Underweight/underperform/moderate sell/weak hold	−1	0.09
Sell	−2	0.26

The distribution of the five levels is also displayed in Table 1, showing that the majority of stock ratings is "Strong Buy/Buy" in our sample. This skewed distribution is consistent with that in prior researches [21, 84].

4.3. Online Investor Sentiment. There are usually two kinds of methods to identify the sentiment polarities of texts. The first is sentiment-dictionary-based and the second is machine-learning-based. Boukes et al. [85] found the performances of dictionary-based approaches are not good and most researchers point out that machine-learning-based methods generally outperform dictionary-based methods [86, 87]. Therefore, we adopt the supervised machine learning method to retrieve investor sentiment from online investor comments. The process contains 4 steps, as follows.

First, we randomly select 50,000 items of investor comments and divide them into positive, negative, and neutral groups manually according to their emotion polarities. To ensure the quality of manual labeling, we hire three financial postgraduate students, and at the beginning, they label the comments independently. After they finish their labeling work independently, they compare their results and keep the comments whose emotional labels given by the three students are the same. Finally, we get a corpus, including 19,785 neutral comments, 4681 positive comments, and 4693 negative comments, which can be used for supervised training.

Second, we use Scikit-Learn (https://scikit-learn.org/stable/index.html), which is the most useful and robust library for machine learning in Python, to train an auto-sentiment-classifier [88]. We use Jieba (https://pypi.org/project/jieba/) to implement Chinese words segmentation and the chi-square statistic of each word is calculated by using Natural Language Toolkit (NLTK, https://www.nltk.org/). There are six well-known algorithms in the scikit-learn library. To examine which algorithm is most suitable and what the best number of feature words is, we run the above six models by setting the feature number from 800 to 5000. Because the numbers of the positive, negative, and neutral comments in the annotated corpus are not the same, we also repeat training the six models 100 times for each feature number and each time the same quantity of positive, negative, and neutral comments are randomly selected from the corpus. We compare the means of the accuracy scores of different classification models using one-way ANOVA (analysis of variance), with the results showing that Bernoulli Naïve Bayes (BernoulliNB) and Multinomial Naïve Bayes (MultinomiaNB) outperform the other algorithms, shown in Table 2. ANOVA results also show that the feature number should be between 2,500 and 3,000.

Third, according to the above tests, we choose BernoulliNB and set the feature number as 2,750. After executing the training program, we get an auto-sentiment-classifier, whose performance is shown in Table 3. As can be seen, the BernoulliNB algorithm using the corpus we

developed can achieve a 77.18% accuracy rate, better than that of Lutz et al. [89].

Fourth, using the well-trained classifier, we perform sentiment analysis of all comments of the constituents of CSI300, about 7,816,980 items. According to the study of Tsukioka et al. [90], we calculate investor sentiment with formula (3), as follows:

$$IS_{i,t} = \ln \frac{1 + pos_{i,t}}{1 + neg_{i,t}}, \tag{3}$$

where $IS_{i,t}$ denotes the investor sentiment of stock i on day t, $pos_{i,t}$ indicates the number of positive comments of stock i on day t, and $neg_{i,t}$ denotes the number of negative comments of stock i on day t.

4.4. Regression Models. Based on prior researches [65, 91, 92] and the hypotheses in Section 3, we adopt the following basic regression models to investigate the effect of research reports on stock returns.

$$SR_{i,t} = \alpha + \beta_1 Att_{i,t-1} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{4}$$

$$SR_{i,t} = \alpha + \beta_1 Rat_{i,t-1} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{5}$$

$$SR_{i,t} = \alpha + \beta_1 Att_{i,t-1} + \beta_2 Rat_{i,t-1} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{6}$$

where $SR_{i,t}$ denotes the stock return of stock i on day t. We also control for the lag stock return, value of market value (MV), book-to-market ratio (BMR), and BETA according to prior researches [65, 83, 93].

In order to explore the role of online investor sentiment, we also adopt the following equations:

$$IS_{i,t} = \alpha + \beta_1 Att_{i,t-1} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{7}$$

$$IS_{i,t} = \alpha + \beta_1 Rat_{i,t-1} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{8}$$

$$IS_{i,t} = \alpha + \beta_1 Att_{i,t-1} + \beta_2 Rat_{i,t-1} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{9}$$

$$SR_{i,t} = \alpha + \beta_1 Att_{i,t-1} + \beta_2 IS_{i,t} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{10}$$

$$SR_{i,t} = \alpha + \beta_1 Rat_{i,t-1} + \beta_2 IS_{i,t} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{11}$$

$$SR_{i,t} = \alpha + \beta_1 Att_{i,t-1} + \beta_2 Rat_{i,t-1} + \beta_3 IS_{i,t} + \sum_k \beta_k Controls_{k,i,t-1} + \varepsilon_{i,t}. \tag{12}$$

5. Analysis Results

5.1. Descriptive Statistics of Variables. After combining the investor sentiment (IS), attention of research report (Att), and stock rating (Rat) data derived from investor comments and research reports on East Money with the trading data from CSMAR, and removing missing values, we get a panel data, consisting of 96,033 records of 126 stocks, ranging from August 1, 2014, to July 31, 2018. The descriptive statistics of the variables are shown in Table 4. Figure 2 shows the daily time series of investor sentiment, research report

attention, stock rating, and stock return of one stock (SH600309).

To investigate the relationships among research reports, investor sentiment, and stock returns, we adopt the ordinary least squares (OLS) method, the same as that in Fang et al. [94], Blau [95], and Sun et al. [96], because it is used with the similar aim of analyzing the effect of investor sentiment on stock returns. Since the residuals show evidence of heteroscedasticity in our data, we use robust standard errors to calculate t statistics according to suggestions of statisticians [94, 97]. Furthermore, in order to avoid the multicollinearity

Table 2: One-way ANOVA results (factor: algorithms).

Group 1	Group 2	Mean diff	p-adj	Lower	Upper	Reject
BernoulliNB	LinearSVC	−0.0058	0.001	−0.0064	−0.0052	True
BernoulliNB	LogisticRegression	−0.0079	0.001	−0.0085	−0.0073	True
BernoulliNB	MultinomiaNB	−0.0001	0.900	−0.0007	0.0005	False
BernoulliNB	NuSVC	−0.0239	0.001	−0.0245	−0.0234	True
BernoulliNB	SVC	−0.0262	0.001	−0.0268	−0.0256	True
LinearSVC	LogisticRegression	−0.0020	0.001	−0.0026	−0.0015	True
LinearSVC	MultinomiaNB	0.0057	0.001	0.0051	0.0063	True
LinearSVC	NuSVC	−0.0181	0.001	−0.0187	−0.0175	True
LinearSVC	SVC	−0.0204	0.001	−0.0209	−0.0198	True
LogisticRegression	MultinomiaNB	0.0077	0.001	0.0071	0.0083	True
LogisticRegression	NuSVC	−0.0161	0.001	−0.0167	−0.0155	True
LogisticRegression	SVC	−0.0183	0.001	−0.0189	−0.0177	True
MultinomiaNB	NuSVC	−0.0238	0.001	−0.0244	−0.0232	True
MultinomiaNB	SVC	−0.0260	0.001	−0.0266	−0.0254	True
NuSVC	SVC	−0.0022	0.001	−0.0028	−0.0016	True

Table 3: Performances of different algorithms (feature number = 2750).

Algorithm	Overall accuracy	Positive group		Negative group	
		Accuracy	F_1	Accuracy	F_1
BernoulliNB	**0.7718**	**0.8745**	**0.8161**	0.8538	0.7884
MultinomiaNB	0.7693	0.8713	0.8136	**0.8545**	**0.7890**
LogisticRegression	0.7569	0.8713	0.8056	0.8520	0.7678
SVC	0.7312	0.8567	0.7810	0.8389	0.7393
LinearSVC	0.7508	0.8692	0.8043	0.8467	0.7632
NuSVC	0.7319	0.8567	0.7832	0.8389	0.7399

Each bold value is the maximum value of its column.

Table 4: Descriptive statistics.

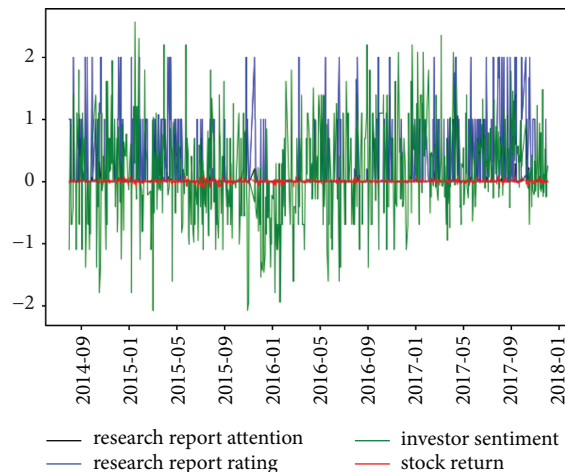| Variables	Mean	Std	Min	Max
Att	0.0031868	0.0188282	0	1
Rat	0.1412699	0.4147632	−2	2
IS	−0.1095563	0.6618756	−3.218876	3.912023
SR	0.0007777	0.0267761	−0.101426	0.101961
MV	$1.16e + 08$	$1.79e + 08$	1554300	$2.09e + 09$
BMR	0.5319294	0.3548162	0.000223	2.212037
BETA	1.018393	0.5117594	−2.371917	5.490196

Figure 2: Research report, sentiment, and stock return (SH600309).

problem, we use the variance inflation factor (VIF) to test where there is a multicollinearity problem in our OLS regression model. Our results show that the maximum VIF is 1.27, less than the suggested threshold, indicating that our model has no serious multicollinearity problem.

5.2. The Effect of Research Reports on Stock Returns.

In this section, we run the regression equations (4)–(6) separately, the results shown in Table 5. After controlling for the lag return, market value, book-to-market ratio, and BETA, either the research report attention alone or the research report rating alone has a significant effect on stock returns, with the coefficient equal to 0.0834628 ($p \leq 0.001$), 0.022528 ($p \leq 0.001$), respectively. However, when the regression model includes both the research report attention and the research report rating, the effect of research report attention is no longer significant ($\beta = 0.0314562$, $p = 0.300 > 0.05$) and the effect of research report rating is still significant ($\beta = 0.019971$, $p \leq 0.001$). The analysis results show that research reports issued by securities companies do influence stock returns. Although we divide the influence power of research reports into attention and rating; however, it seems that the stock rating in research reports is more important.

5.3. The Effect of Research Reports on Investor Sentiment.

We further examine the effect of research reports on investor sentiment, and the regression results are shown in Table 6. As can be seen, either the research report attention ($\beta = 0.0731084$, $p \leq 0.001$) alone or the stock rating ($\beta = 0.0211631$, $p \leq 0.001$) in research reports alone can influence investor sentiment significantly after controlling for the control variables. However, when the research report attention and the stock rating are both included in the regression model, only the stock rating ($\beta = 0.0193073$, $p \leq 0.001$) has a significant effect on investor sentiment, just like they do on stock returns.

5.4. The Effect of Investor Sentiment and Research Report on Stock Return.

In this section, we examine the synthetic influence of research reports and investor sentiment on stock returns, with results shown in Table 7. In model (10), both research report attention ($\beta = 0.0565279$, $p = 0.030 < 0.05$) and investor sentiment ($\beta = 0.3684253$, $p \leq 0.001$)

have significant effect on stock returns. In model (11), both stock rating ($\beta = 0.0147361$, $p \leq 0.001$) and investor sentiment ($\beta = 0.3681841$, $p \leq 0.001$) have significant effect on stock returns. However, in the model (12), the research report attention no longer has a significant effect on stock returns ($\beta = 0.0230509$, $p = 0.430 > 0.05$).

5.5. The Mediating Effect of Investor Sentiment.

In order to test the mediating effect of investor sentiment, we rewrite equations (5), (8), and (11) as follows:

$$SR_{i,t} = c + \lambda Rat_{i,t-1} + \sum_k \gamma_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{13}$$

$$IS_{i,t} = c + \alpha Rat_{i,t-1} + \sum_k \gamma_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{14}$$

$$SR_{i,t} = c + \lambda' Rat_{i,t-1} + \beta IS_{i,t} + \sum_k \gamma_k Controls_{k,i,t-1} + \varepsilon_{i,t}, \tag{15}$$

where λ is the total effect, λ' is the direct effect, and $\alpha\beta$ is the indirect effect. There exists $\lambda = \lambda' + \alpha\beta$. If $\alpha\beta$ does not equal zero significantly, then mediating effect exists. Furthermore, if λ' is not equal to zero significantly, there exists partial mediating effect.

According to prior researches, we test the mediating effect of investor sentiment by the Sobel test, the Z value for which is as follows:

$$Z = \frac{\alpha\beta}{\sqrt{\alpha^2 \sigma_\beta^2 + \beta^2 \sigma_\alpha^2}}, \tag{16}$$

where σ_α and σ_β are standard errors of estimations of α and β.

In model (13), the effect of research report rating on stock returns is significant ($\lambda = 0.022528$, $t = 5.66$). In model (14), the effect of research report rating on investor sentiment is also significant ($\alpha = 0.0211631$, $\sigma_\alpha = 0.0054672$, $t = 3.87$). In model (15), the effects of research report rating and investor sentiment on stock returns are both significant ($\lambda' = 0.0147361$, $t = 3.67$, $\beta = 0.3681841$, $\sigma_\beta = 0.01362$, $t = 27.03$). Therefore, the Z value for the Sobel test is

$$Z = \frac{\alpha\beta}{\sqrt{\alpha^2 \sigma_\beta^2 + \beta^2 \sigma_\alpha^2}} = \frac{0.0211631 * 0.3681841}{\sqrt{0.0211631^2 * 0.01362^2 + 0.3681841^2 * 0.0054672^2}} = 3.83. \tag{17}$$

At a two-tailed significance level of 0.05, the critical value is ±1.96; in this study, the Sobel Z value falls outside ±1.96, which means that $\alpha\beta$ is not equal to zero significantly. The statistics of the total, direct, and indirect effects are shown in Table 8. Both the direct and indirect effects are significant, representing that investor sentiment plays a partial mediating effect role in the relationship between research reports and stock returns.

5.6. Robustness Check.

In this section, we examine the robustness of the regression models by using another investor sentiment index. CSMAR also provides investor sentiment, which is also derived from online stock forums including East Money (https://guba.eastmoney.com/) and Sina Finance (https://finance.sina.com.cn/). After substituting the CSMAR's investor sentiment for our investor sentiment

TABLE 5: Regression results of stock return and research report.

	Model (4)	Model (5)	Model (6)
l.Att	0.0834628**		0.0314562
	(3.16)		(1.04)
l.Rat		0.0225280***	0.0199710***
		(5.66)	(4.50)
l.SR	0.0284335***	0.0280409***	0.0280219***
	(4.76)	(4.71)	(4.70)
l.MV	−0.0224525*	−0.0230832*	−0.0231090*
	(−2.30)	(−2.30)	(−2.29)
l.BMR	0.0121836***	0.0123575***	0.0124453***
	(4.85)	(4.92)	(4.97)
l.BETA	−0.0108032	−0.0106788	−0.0105945
	(−1.71)	(−1.69)	(−1.68)
Constant	0.4882944***	0.4766196***	0.4778434***
	(106.01)	(90.31)	(86.44)
R^2	0.0014	0.0016	0.0016
F-statistic	11.92***	14.42***	12.25***
N	71,031	71,031	71,031

Notes: t statistics are in parentheses. ***, **, and * denote significance at the 0.001, 0.01, and 0.05 level, respectively.

TABLE 6: Regression results of investor sentiment and research report.

	Model (7)	Model (8)	Model (9)
l.Att	0.0731084***		0.0228300
	(3.56)		(1.43)
l.Rat		0.0211631***	0.0193073***
		(3.87)	(3.56)
l.SR	0.0706226***	0.0702385***	0.0702248***
	(28.39)	(28.32)	(28.34)
l.MV	0.0847967***	0.0841808***	0.0841621***
	(3.89)	(3.90)	(3.90)
l.BMR	0.0241770	0.0243662	0.0244300
	(1.57)	(1.58)	(1.58)
l.BETA	−0.0578704**	−0.0577297**	−0.0576686**
	(−2.48)	(−2.48)	(−2.48)
Constant	0.4134446***	0.4024528***	0.4033410***
	(37.47)	(34.37)	(34.64)
R^2	0.0242	0.0245	0.0245
F-statistic	180.25***	176.96***	151.08***
N	71,031	71,031	71,031

Notes: t statistics are in parentheses. ***, **, and * denote significance at the 0.001, 0.01, and 0.05 level, respectively.

index, stock rating in research reports still has a significant effect on investor sentiment and stock return, as shown in Table 9. However, not only does research report attention has no significant effect on stock return when both the attention and rating are included in the regression model, but also it does not influence investor sentiment any longer.

We further examine the mediating effect of investor sentiment using the CSMAR's investor sentiment index, with results shown in Table 10. As can be seen, the mediation effect is still significant.

$$Z = \frac{\alpha\beta}{\sqrt{\alpha^2 \sigma_\beta^2 + \beta^2 \sigma_\alpha^2}} = \frac{0.0017288 * 0.0793405}{\sqrt{0.0017288^2 * 0.161291^2 + 0.0793405^2 * 0.0007003^2}} = 2.21. \tag{18}$$

TABLE 7: Regression results of investor sentiment, research report, and stock returns.

	Model (10)	Model (11)	Model (12)
l.Att	0.0565279*		0.0230509
	(2.20)		(0.79)
l.Rat		0.0147361***	0.0128627**
		(3.67)	(2.88)
IS	0.3684253***	0.3681841***	0.3681670***
	(27.02)	(27.03)	(27.02)
l.SR	0.0024143	0.0021802	0.0021675
	(0.42)	(0.38)	(0.38)
l.MV	−0.0536938***	−0.0540772***	−0.0540947***
	(−3.41)	(−3.40)	(−3.40)
l.BMR	0.0032762	0.0033862	0.0034510
	(0.55)	(0.57)	(0.58)
l.BETA	0.0105177	0.0105764	0.0106372
	(1.30)	(1.30)	(1.31)
Constant	0.3359709***	0.3284429***	0.3293466***
	(42.68)	(39.50)	(38.64)
R^2	0.0724	0.0725	0.0725
F-statistic	131.46***	128.29***	112.65***
N	71,031	71,031	71,031

Notes: t statistics are in parentheses. ***, **, and * denote significance at the 0.001, 0.01, and 0.05 level, respectively.

TABLE 8: Results of mediating effect analysis.

	Coefficient	Std. error	z value
Total effect	0.0225280	0.0039825	5.66
Direct effect	0.0147361	0.0040185	3.67
Indirect effect	0.0077919	0.0020335	3.83

TABLE 9: Regression results (using CSMAR sentiment index).

	Case (1)	Case (2)	Case (3)
l.Rat	0.0248163***	0.0017288*	0.0246792***
	(7.52)	(2.47)	(7.47)
IS			0.0793405***
			(4.92)
l.SR	0.0234505***	0.0013279*	0.0233452***
	(4.64)	(2.50)	(4.62)
l.MV	−0.0110012*	0.0021495	−0.0111717*
	(−2.37)	(1.74)	(−2.39)
l.BMR	0.0219775***	0.000058	0.0219729***
	(4.84)	(0.03)	(4.83)
l.BETA	−0.0197732**	−0.0012477	−0.0196742**
	(−2.78)	(−0.72)	(−2.76)
Constant	0.4805686***	0.4730054***	0.4430401***
	(93.65)	(430.12)	(45.94)
R^2	0.0013	0.0002	0.0015
F-statistic	21.35***	2.89*	20.87***
N	118,674	118,674	118,674

Notes: t statistics are in parentheses. ***, **, and * denote significance at the 0.001, 0.01, and 0.05 level, respectively.

We also conduct several robustness checks following previous researches to examine the effects of outliers or influential observations [98, 99]. And the results show that our findings are robust; i.e., the effects of research reports and investor sentiment on stock returns are still significant after controlling for the potential presence of outliers.

TABLE 10: Results of mediating effect analysis (using CSMAR sentiment index).

	Coefficient	Std. error	z value
Total effect	0.0248163	0.0033022	7.52
Direct effect	0.0246792	0.0033047	7.47
Indirect effect	0.0001372	$6.21665e - 05$	2.21

5.7. Discussion. We investigate the effect of research reports released by securities companies from two aspects of attention and rating. The empirical analysis results show that the influence of research report attention is not stable, whereas the rating index is a reliable and stable indicator. Moreover, the effects of rating in research reports on stock returns and investor sentiment are both positive, indicating that, with the optimism of analysts rising, the investor sentiment and stock returns will increase, and vice versa. In addition, we also find that investor sentiment plays a partial mediating role in the relationship between stock rating and stock returns. These findings imply that although some individual investors are disgusted with research reports issued by securities companies, there does exist a positive correlation between analysts' rating and stock returns in the Chinese stock market. In other words, in China, whether stock recommendations have relevance to conflicts of interest or not, in general, the analysts' recommendations have investment value. Therefore, this study has implications for investors who get information from social media or stock forums such as East Money. They can track the average rating and investor sentiment of the stock they are concerned about and adjust their investment strategies accordingly.

6. Conclusions

As an important public information source, research reports' effect has not been investigated fully in China. Although many researchers propose that recommendations in research reports issued by securities companies have a positive relation to stock markets, in Chinese stock forums, there is still a voice saying that the stock price will fall as soon as the research reports appear. Therefore, this study investigates the effect of research reports using data from Chinese stock forums, which is an innovative research perspective different from prior studies thanks to the advances of social media, big data, and natural language processing technologies. Based on behavioral finance theory, we hypothesize that research reports issued by securities companies have a significant effect on stock returns and investor sentiment plays a mediating role in it. To test the hypotheses, we first collect research reports from East Money and get two indicators, i.e., the attention of research reports and the average stock rating of research reports. Then, we collect online stock comments and derive investor sentiment by supervised machine learning method based on Scikit-Learn library. During the process, we develop a stock comment corpus, with which we find a satisfactory feature number and find that the Bernoulli Naïve Bayes and Multinomial Naïve Bayes are better than other algorithms for financial text classification. At last, we examine the relationship between research reports and stock returns and test the mediating role of investor sentiment. The empirical analysis results support our hypotheses, showing that stock ratings in research reports have a significant positive effect on stock returns and investor sentiment plays a partially mediating role, which is of great significance to participants of the Chinese stock markets.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Y. Chen, C. Y. Hu, W. J. Zhang, and Q. Li, "CEO exposure, media influence, and stock returns," *Journal of Global Information Management*, vol. 29, p. 43, 2021.

[2] H. S. Qiao and Y. Y. Su, "Media coverage and decomposition of stock market volatility: based on the generalized dynamic factor model," *Emerging Markets Finance and Trade*, vol. 56, no. 3, pp. 1–13, 2020.

[3] K. Kim, D. Ryu, and H. Yang, "Information uncertainty, investor sentiment, and analyst reports," *International Review of Financial Analysis*, vol. 77, Article ID 101835, 2021.

[4] C. Gu, D. Chen, and R. Stan, "Investor sentiment and the market reaction to macroeconomic news," *Journal of Futures Markets*, vol. 41, no. 9, pp. 1412–1426, 2021.

[5] K. Kim, D. Ryu, and H. Yang, "Investor sentiment, stock returns, and analyst recommendation changes: the KOSPI stock market," *Investment Analysts Journal*, vol. 48, no. 2, pp. 89–101, 2019.

[6] R. F. Engle, M. K. Hansen, A. K. Karagozoglu, and A. Lunde, "News and idiosyncratic volatility: the public information processing hypothesis," *Journal of Financial Econometrics*, vol. 19, no. 1, pp. 1–38, 2021.

[7] D. E. Allen, M. McAleer, and A. K. Singh, "Daily market news sentiment and stock prices," *Applied Economics*, vol. 51, no. 30, pp. 3212–3235, 2019.

[8] H. Jiang, S. Z. Li, and H. Wang, "Pervasive underreaction: evidence from high-frequency data," *Journal of Financial Economics*, vol. 141, no. 2, pp. 573–599, 2021.

[9] A. Bouteska and M. Mili, "The role of investor sentiment and valuation uncertainty in the changes around analyst recommendations: evidence from US firms," *The Journal of Behavioral Finance*, pp. 1–24, 2021.

[10] G. Birz, S. Dutta, and H. Yu, "Economic forecasts, anchoring bias, and stock returns," *Financial Management*, 2021.

[11] D. Lien, P. H. Hung, and L. W. Chen, "Analyst recommendations, herding intensity, and trading performance," *Asia-Pacific Journal of Financial Studies*, vol. 49, no. 5, pp. 772–802, 2020.

[12] P. C. Tetlock, "Information transmission in finance," *Annual Review of Financial Economics*, vol. 6, no. 1, pp. 365–384, 2014.

[13] Y. Duan, L. Liu, and Z. Wang, "COVID-19 sentiment and the Chinese stock market: evidence from the official news media and Sina weibo," *Research in International Business and Finance*, vol. 58, Article ID 101432, 2021.

[14] W. Y. Chen, L. G. Zhang, P. Y. Jiang, F. L. Meng, and Q. Y. Sun, "Can digital transformation improve the information environment of the capital market? evidence from the analysts' prediction behaviour," *Accounting and Finance*, 2021.

[15] Y. S. Dong and D. Q. Young, "Foreign macroeconomic and industry-related information transfers around earnings announcements: evidence from US-listed non-US firms," *Journal of Accounting and Economics*, vol. 71, no. 2-3, 2021.

[16] J. Peress and D. Schmidt, "Glued to the TV: distracted noise traders and stock market liquidity," *The Journal of Finance*, vol. 75, no. 2, pp. 1083–1133, 2020.

[17] A. Al-Nasseri, F. M. Ali, and A. Tucker, "Investor sentiment and the dispersion of stock returns: evidence based on the social network of investors," *International Review of Financial Analysis*, vol. 78, 2021.

[18] J. A. Cookson and M. Niessner, "Why don't we agree? evidence from a social network of investors," *The Journal of Finance*, vol. 75, no. 1, pp. 173–228, 2020.

[19] L. Guo, F. W. Li, and K. C. John Wei, "Security analysts and capital market anomalies," *Journal of Financial Economics*, vol. 137, no. 1, pp. 204–230, 2020.

[20] K. Kim, D. Ryu, and J. Yu, "Do sentiment trades explain investor overconfidence around analyst recommendation revisions?" *Research in International Business and Finance*, vol. 56, 2021.

[21] D. Kong, C. Lin, S. Liu, and W. Tan, "Whose money is smart? Individual and institutional investors' trades based on analyst recommendations," *Journal of Empirical Finance*, vol. 62, pp. 234–251, 2021.

[22] E. F. Fama, "Efficient capital markets: a review of theory and empirical work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[23] N. Strauss, R. Vliegenthart, and P. Verhoeven, "Intraday news trading: the reciprocal relationships between the stock market and economic news," *Communication Research*, vol. 45, pp. 1054–1077, 2018.

[24] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: a case of Hong Kong," *Information Processing and Management*, vol. 57, no. 5, Article ID 102212, 2020.

[25] R. P. Schumaker, Y. Zhang, C.-N. Huang, and H. Chen, "Evaluating sentiment in financial news articles," *Decision Support Systems*, vol. 53, no. 3, pp. 458–464, 2012.

[26] R. Arévalo, J. García, F. Guijarro, and A. Peris, "A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting," *Expert Systems with Applications*, vol. 81, pp. 177–192, 2017.

[27] X. C. Wan, J. Yang, S. Marinov, J. P. Calliess, S. Zohren, and X. W. Dong, "Sentiment correlation in financial news networks and associated market movements," *Scientific Reports*, vol. 11, p. 12, 2021.

[28] Y. Z. Li, S. R. Jiang, X. R. Li, and S. Y. Wang, "The role of news sentiment in oil futures returns and volatility forecasting: data-decomposition based deep learning approach," *Energy Economics*, vol. 95, 2021.

[29] P. K. Narayan, "Can stale oil price news predict stock returns?" *Energy Economics*, vol. 83, pp. 430–444, 2019.

[30] P. Jiao, A. Veiga, and A. Walther, "Social media, news media and the stock market," *Journal of Economic Behavior and Organization*, vol. 176, pp. 63–90, 2020.

[31] P. C. Tetlock, "Giving content to investor sentiment: the role of media in the stock market," *The Journal of Finance*, vol. 62, no. 3, pp. 1139–1168, 2007.

[32] P. Glasserman and H. Mamaysky, "Does unusual news forecast market stress?" *Journal of Financial and Quantitative Analysis*, vol. 54, no. 5, pp. 1937–1974, 2019.

[33] I. El Ouadghiri, K. Guesmi, J. Peillex, and A. Ziegler, "Public attention to environmental issues and stock market returns," *Ecological Economics*, vol. 180, 2021.

[34] N. Strauss, "Financial journalism in today's high-frequency news and information era," *Journalism*, vol. 20, pp. 274–291, 2019.

[35] D. B. Vukovic, V. Ugolnikov, and M. Maiti, "Sell-side analysts' recommendations a value or noise," *International Journal of Finance & Economics*, vol. 26, no. 2, pp. 3134–3151, 2021.

[36] Y. Yin, "Fund portfolio holdings and affiliated analyst's objectivity: do equity and employment relationship matter?" *The Journal of Behavioral Finance*, vol. 19, no. 1, pp. 16–29, 2018.

[37] M. Baker and J. Wurgler, "Investor sentiment and the cross-section of stock returns," *The Journal of Finance*, vol. 61, no. 4, pp. 1645–1680, 2006.

[38] S.-H. Kim and D. Kim, "Investor sentiment from internet message postings and the predictability of stock returns," *Journal of Economic Behavior & Organization*, vol. 107, pp. 708–729, 2014.

[39] L. Sun, M. Najand, and J. Shen, "Stock return predictability and investor sentiment: a high-frequency perspective," *Journal of Banking & Finance*, vol. 73, pp. 147–164, 2016.

[40] S. Behrendt and A. Schmidt, "The Twitter myth revisited: intraday investor sentiment, Twitter activity and individual-level stock return volatility," *Journal of Banking and Finance*, vol. 96, pp. 355–367, 2018.

[41] F. Audrino, F. Sigrist, and D. Ballinari, "The impact of sentiment and attention measures on stock market volatility," *International Journal of Forecasting*, vol. 36, no. 2, pp. 334–357, 2020.

[42] D. T. Selin and O. Tas, "Social media sentiment in international stock returns and trading activity," *The Journal of Behavioral Finance*, vol. 22, pp. 221–234, 2021.

[43] T. Renault, "Intraday online investor sentiment and return patterns in the U.S. stock market," *Journal of Banking and Finance*, vol. 84, pp. 25–40, 2017.

[44] Y. Sun, X. Zeng, S. Zhou, H. Zhao, P. Thomas, and H. Hu, "What investors say is what the market says: measuring China's real investor sentiment," *Personal and Ubiquitous Computing*, vol. 25, no. 3, pp. 587–599, 2021.

[45] R. Cervelló-Royo, F. Guijarro, and K. Michniuk, "Stock market trading rule based on pattern recognition and technical analysis: forecasting the DJIA index with intraday data," *Expert Systems with Applications*, vol. 42, no. 14, pp. 5963–5975, 2015.

[46] C. Gu and A. Kurov, "Informational role of social media: evidence from Twitter sentiment," *Journal of Banking and Finance*, vol. 121, 2020.

[47] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[48] P. D. Azar and A. W. Lo, "The wisdom of twitter crowds: predicting stock market reactions to FOMC meetings via twitter feeds," *Journal of Portfolio Management*, vol. 42, no. 5, pp. 123–134, 2016.

[49] C. X. Huang, S. G. Wen, X. Yang, J. D. Cao, and X. G. Yang, "Measurement of individual investor sentiment and its application: evidence from Chinese stock message board," *Emerging Markets Finance and Trade*, pp. 1–11, 2020.

[50] Y. Z. Lv, J. Z. Piao, B. N. Li, and M. J. Yang, "Does online investor sentiment impact stock returns? Evidence from the Chinese stock market," *Applied Economics Letters*, 2021.

[51] W. Yang, D. Lin, and Z. Yi, "Impacts of the mass media effect on investor sentiment," *Finance Research Letters*, vol. 22, pp. 1–4, 2017.

[52] B. Q. Gan, V. Alexeev, R. Bird, and D. Yeung, "Sensitivity to sentiment: news vs social media," *International Review of Financial Analysis*, vol. 67, p. 17, 2020.

[53] S. H. Chen, X. P. Cao, K. B. Lin, J. B. Huang, Y. B. Zhang, and H. W. Lin, "Financial transparency, media coverage, and momentum in China," *Emerging Markets Finance and Trade*, vol. 13, 2020.

[54] W. G. Zhang, X. Gong, C. Wang, and X. Ye, "Predicting stock market volatility based on textual sentiment: a nonlinear analysis," *Journal of Forecasting*, 2021.

[55] J. Chun, J. Ahn, Y. Kim, and S. Lee, "Using deep learning to develop a stock price prediction model based on individual investor emotions," *The Journal of Behavioral Finance*, vol. 22, no. 4, pp. 480–489, 2020.

[56] Y. Li, H. Bu, J. Li, and J Wu, "The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning," *International Journal of Forecasting*, vol. 36, no. 4, pp. 1541–1562, 2020.

[57] A. Bhardwaj and S. Imam, "The tone and readability of the media during the financial crisis: evidence from pre-IPO media coverage," *International Review of Financial Analysis*, vol. 63, pp. 40–48, 2019.

[58] T. Marty, B. Vanstone, and T. Hahn, "News media analytics in finance: a survey," *Accounting and Finance*, vol. 60, no. 2, pp. 1385–1434, 2020.

[59] M. Alomari, A. R. Al Rababa'a, G. El-Nader, A. Alkhataybeh, and M. Ur Rehman, "Examining the effects of news and media sentiments on volatility and correlation: evidence from the UK," *The Quarterly Review of Economics and Finance*, vol. 82, pp. 280–297, 2021.

[60] T. G. L. A. Van der Meer and R. Vliegenthart, "The consequences of being on the agenda: the effect of media and public attention on firms' stock market performance," *Communications*, vol. 43, no. 1, pp. 5–24, 2018.

[61] C. C. Li, H. C. Xu, and W. X. Zhou, "News coverage and portfolio returns: evidence from China," *Pacific-Basin Finance Journal*, vol. 60, 2020.

[62] P. K. Narayan and D. Bannigidadmath, "Does financial news predict stock returns? New evidence from islamic and non-islamic stocks," *Pacific-Basin Finance Journal*, vol. 42, pp. 24–45, 2017.

[63] S. L. Heston and N. R. Sinha, "News vs. Sentiment: predicting stock returns from news stories," *Financial Analysts Journal*, vol. 73, no. 3, pp. 67–83, 2017.

[64] G. Loffler, L. Norden, and A. Rieber, "Negative news and the stock market impact of tone in rating reports," *Journal of Banking and Finance*, vol. 133, no. 2, Article ID 106256, 2021.

[65] T. Huang and X. Y. Zhang, "Media Coverage of Industry and the Cross-Section of Stock Returns," *Accounting and Finance*, 2021.

[66] G. Guan-Ru Wu, T. Chieh-Tse Hou, and J.-L. Lin, "Can economic news predict Taiwan stock market returns?" *Asia Pacific Management Review*, vol. 24, no. 1, pp. 54–59, 2019.

[67] A. E. Khedr, S. E. Salama, and N. Yaseen, "Predicting stock market behavior using data mining technique and news sentiment analysis," *International Journal of Intelligent Systems and Applications*, vol. 9, no. 7, pp. 22–30, 2017.

[68] W. L. G. Hsieh and C. S. Lee, "Who reacts to what information in securities analyst reports? Direct evidence from the investor trade imbalance," *Pacific-Basin Finance Journal*, vol. 65, Article ID 101492, 2021.

[69] R. Barniv, W. Li, and T. Miller, "Independence of intrinsic valuations and stock recommendations - experimental evidence from equity research analysts and investors," *The Journal of Behavioral Finance*, 2021.

[70] H. L. Song, D. Y. Peng, and X. Huang, "Incorporating research reports and market sentiment for stock excess return prediction: a case of mainland China," *Scientific Programming*, vol. 2020, Article ID 8894757, 7 pages, 2020.

[71] S. Ishigami and F. Takeda, "Market reactions to stock rating and target price changes in analyst reports: evidence from Japan," *Journal of International Financial Markets, Institutions and Money*, vol. 52, pp. 134–151, 2018.

[72] Q. Bu, "Are all the sentiment measures the same?" *The Journal of Behavioral Finance*, 2021.

[73] A. Anand, S. Basu, J. Pathak, and A. Thampy, "The impact of sentiment on emerging stock markets," *International Review of Economics and Finance*, vol. 75, pp. 161–177, 2021.

[74] Y. Y. Xu and J. C. Zhao, "Can sentiments on macroeconomic news explain stock returns? Evidence form social network data," *International Journal of Finance and Economics*, 2020.

[75] H.-C. Shu and J.-H. Chang, "Investor sentiment and financial market volatility," *The Journal of Behavioral Finance*, vol. 16, no. 3, pp. 206–219, 2015.

[76] M. Qadan and D. Kliger, "The short trading day anomaly," *Journal of Empirical Finance*, vol. 38, pp. 62–80, 2016.

[77] A. Rapp, "Sentiment versus mood: a conceptual and empirical investigation," *Journal of Capital Markets Studies*, vol. 3, no. 1, pp. 6–17, 2019.

[78] H. K. Sul, A. R. Dennis, and L. I. Yuan, "Trading on twitter: using social media sentiment to predict stock returns," *Decision Sciences*, vol. 48, no. 3, pp. 454–488, 2017.

[79] J. Li, "The term structure effects of individual stock investor sentiment on excess returns," *International Journal of Finance & Economics*, vol. 26, no. 2, pp. 1695–1705, 2021.

[80] M. Dahmene, A. Boughrara, and S. Slim, "Nonlinearity in stock returns: do risk aversion, investor sentiment and, monetary policy shocks matter?" *International Review of Economics and Finance*, vol. 71, pp. 676–699, 2021.

[81] J. Shi, M. Ausloos, and T. T. Zhu, "If global or local investor sentiments are prone to developing an impact on stock returns, is there an industry effect?" *International Journal of Finance and Economics*, 2020.

[82] B. Gao and X. Liu, "Intraday sentiment and market returns," *International Review of Economics and Finance*, vol. 69, pp. 48–62, 2020.

[83] Y. A. Xu, J. Q. Wang, Z. L. Chen, and C. Liang, "Sentiment indices and stock returns: evidence from China," *International Journal of Finance and Economics*, 2021.

[84] C.-Y. Chan, H.-C. Lo, and Y.-R. Su, "Distribution of stock ratings and analyst recommendation revision," *The North*

American Journal of Economics and Finance, vol. 28, pp. 273–286, 2014.

[85] M. Boukes, B. Van de Velde, T. Araujo, and R. Vliegenthart, "What's the tone? easy doesn't do it: analyzing performance and agreement between off-the-shelf sentiment analysis tools," *Communication Methods and Measures*, vol. 14, no. 2, pp. 83–104, 2020.

[86] W. Van Atteveldt, M. A. C. G. Van der Velden, and M. Boukes, "The validity of sentiment analysis: comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms," *Communication Methods and Measures*, vol. 15, no. 2, pp. 121–140, 2021.

[87] S. T. Zhang, "Sentiment classification of news text data using intelligent model," *Frontiers in Psychology*, vol. 12, 2021.

[88] M. Schnaubelt, T. G. Fischer, and C. Krauss, "Separating the signal from the noise - financial machine learning for Twitter," *Journal of Economic Dynamics and Control*, vol. 114, 2020.

[89] B. Lutz, N. Prollochs, and D. Neumann, "Predicting sentence-level polarity labels of financial news using abnormal stock returns," *Expert Systems with Applications*, vol. 148, p. 11, 2020.

[90] Y. Tsukioka, J. Yanagi, and T. Takada, "Investor sentiment extracted from internet stock message boards and IPO puzzles," *International Review of Economics and Finance*, vol. 56, pp. 205–217, 2018.

[91] X. Gao, C. Gu, and K. Koedijk, "Institutional investor sentiment and aggregate stock returns," *European Financial Management*, vol. 27, no. 5, pp. 899–924, 2021.

[92] Z. Su, M. Lu, and L. Yin, "Chinese stock returns and the role of news-based uncertainty," *Emerging Markets Finance and Trade*, vol. 55, no. 13, pp. 2949–2969, 2019.

[93] B. Kim and S. Suh, "Overnight stock returns, intraday returns, and firm-specific investor sentiment," *The North American Journal of Economics and Finance*, vol. 55, 2021.

[94] X. Fang, Y. Jiang, and Z. Qian, "The effects of individual investors' attention on stock returns: evidence from the ChiNext market," *Emerging Markets Finance and Trade*, vol. 50, no. 3, pp. 158–168, 2014.

[95] B. M. Blau, "Price clustering and investor sentiment," *The Journal of Behavioral Finance*, vol. 20, no. 1, pp. 19–30, 2019.

[96] Y. Sun, X. Liu, G. Y. Chen, Y. H. Hao, and Z. P. Zhang, "How mood affects the stock market: empirical evidence from microblogs," *Information and Management*, vol. 57, p. 11, 2020.

[97] J. Hao and X. Xiong, "Retail investor attention and firms' idiosyncratic risk: evidence from China," *International Review of Financial Analysis*, vol. 74, 2021.

[98] G. Birz, "Stale economic news, media and the stock market," *Journal of Economic Psychology*, vol. 61, pp. 87–102, 2017.

[99] A. J. Leone, M. Minutti-Meza, and C. E. Wasley, "Influential observations and inference in accounting research," *The Accounting Review*, vol. 94, no. 6, pp. 337–364, 2019.