

Association for Information Systems AIS Electronic Library (AISeL)

WHICEB 2014 Proceedings

Wuhan International Conference on e-Business

Summer 6-1-2014

The Relationship between Online Attention and Share Prices

Yue Zhao

School of Management, Harbin Institute of Technology, China, zhaoyuehit@126.com

Qiang Ye

School of Management, Harbin Institute of Technology, China

Ziru Li

School of Management, Harbin Institute of Technology, China

Follow this and additional works at: <http://aisel.aisnet.org/whiceb2014>

Recommended Citation

Zhao, Yue; Ye, Qiang; and Li, Ziru, "The Relationship between Online Attention and Share Prices" (2014). *WHICEB 2014 Proceedings*. 2.

<http://aisel.aisnet.org/whiceb2014/2>

This material is brought to you by the Wuhan International Conference on e-Business at AIS Electronic Library (AISeL). It has been accepted for inclusion in WHICEB 2014 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

The Relationship between Online Attention and Share Prices

Yue Zhao[—], Qiang Ye, Ziru Li

¹School of Management, Harbin Institute of Technology, China

Abstract: Our objective is to explore the correlation between stock market and investors' attention in China's stock market through empirical analysis method. We use search volume (SV) to represent individual investors' attention. We choose 1380 a-share stocks in SSE and SZSE from 2008 to 2011. After summarizing time series data of different cross-sections, we get the panel data. According to the fixed effect analysis of panel data, we find that SV can react the changing of individual investors' attention. An increase in SV predicts higher volume and transaction amount in this week but turn down in the next week in all four years. And the price has the same fluctuation in 2010 and 2011. This is an abnormal phenomenon that more investors pay attention to one stock making its price get lower after a short increase. Through comparing with the analytical result of American, we predict that there will be a stock control and illegal cash in Chinese stock market. The different analytical results between Chinese and American stock market indicated that our stock supervision system is still not perfect.

Keywords: online attention, stock price, search volume, illegal cash

1. INTRODUCTION

Along with the development of the network and the progress of information technology, the value of network information data appears important gradually. Mass of data resources provide lots of available information and promote the development in many fields. The information online can also be used in financial area.

In the past decades, many educational experts have done the researches about stocks used some economic methods, achieved some valuable results. Nowadays online attention information starts to be used in this field. Therefore, in the following pages, there is an introduction to the relationship of online attention, investor attention and stock price. At first, investigators used some indirect proxies of investors' attention such as newspaper, news, company accounting statement. But since Herbert Simon the Nobel Laureate in Economic put forward the opinion that a wealth of information creates a poverty of attention. It means that the indirect proxies of attention can't response the attention of investors correctly. In 2008, Barber and Odean investigated that online search volume is the direct proxy of individual investors. And an increase in individual investors' attention results in temporary positive price pressure^[1].

Based on the characteristics of China stock market and the theory of finance market, we will study the basic operation mechanism, including the structure of China stock market, the reaction of the investors' psychology and the behavior in the stock market. We also describe the online attention of China's stock, which explains the conception of online attention and research of online attention. Then we will summarize methods and results of the former literatures. Meanwhile, according to China stock market characteristics, this paper chooses an appropriate method to illustrate the relationship between share prices and online attention. We will choose the prices and search volume data in four years and the panel data analysis method. According to the data analysis, the result displays that there is a connection between stock prices and online attention. Comparing with American stock market, we try to explain the difference of analysis result and make the appropriate interpretation.

[—] Corresponding author. Email: zhaoyuehit@126.com

2. LITERATURE REVIEW

Though there are many kinds of investigation about stock market, but it still remains an area which worth exploring. For instance, we can use the online search data to get more new information about the stock market. Many researchers have made their contribution to pave the way for building the relationship between internet and finance.

Barber and Odean (2006) analysed the stock attention and found that investors could not buy all stocks which catch their attention. They found that stock of much news follows extremely one-day returns for individual investors [2]. Different from indirect proxies of attention, Zhi Da chose search frequency in Google which is a direct measure. They collected information of Russell 3000 stocks as a sample. “A wealth of information creates a poverty of attention” means that indirect proxies could not reflect investors’ attention correctly. Their conclusion was that an increase of searching frequency predicts higher stock prices in the next two weeks and an eventual price reversal within the year.

Thomas Chemmanur and An Yan (2009) pointed out that a larger stock return associate with a great amount of advertising in the advertising year but a smaller stock return after this time [3]. Thomas Dimpfl and Stephan Jank (2011) stated that high searches follow high volatility, and high volatility follows high searches. And they also found that search requires are useful to predict volatility particularly [4]. Da, Engelberg and Gao (2011) suggested that retail investors are likely use the internet information to help them make a choice before they want to buy stocks [5]. Foucault, Sraer and Thesmar (2011) also found that retail investors’ trading activity leads to a higher level of volatility in individual stocks [6].

Besides this application, search attention data was also used in the field of prediction. Bank, Larch and Peter used search query data to measure retail investor attention in the financial field. Similarly, Da Engelberg and Gao (2010) used this data to predict investors’ earning. Then they used search queries related to household concerns to measure investor sentiment [5]. Ilaria Bordinol, Stefano Battiston’s study also proved that web search queries data can be used to predict stock market volume. The analysis showed that there is a positive correlation between search volume and stock trends [7]. Da Zhi found that higher search volume predicts higher stock prices in the next two weeks and it will reversal with the year [8-10].

Above all, there are correlations between stock data and attention data. And if you search for a stock on the Internet, you are paying attention to it undoubtedly.

3. THE ONLINE ATTENTION AND STOCK MARKET

As direct proxies of online attention approach is rarely used to analyze stock market in China, this paper stated a new way to comprehend some appearance behind economical and online data. This chapter will put forward to hypothesis model which is on the basic of existing theories. After that, we choose an appropriate way to receive the result of data analysis. We try to study the similarities and differences between China and USA. Finally, depending on the analysis and related theory, we try to interpret the economical appearance which was reflected through online attention.

3.1 Stock market of China and online attention findings.

China Stock can be divided into A-shares, B-shares, H-shares and etc. A-shares is short for RMB common stock within the territory of the issue by the company for domestic entities organizations or individuals (excluding Taiwan, Hong Kong, Macao investors) in RMB subscription and trading common stock; B-shares is short for RMB special stock with indicated in RMB, to foreign currency subscription and trading, in China (Shanghai Shenzhen) listed on the stock exchange transaction. So A-shares investment object is organizations and individuals for the mainland of China. Mainboard stock is one of A-shares with lower trading quantity

requirement. It is more suitable for individual investors.

Stock price also called quotations on the stock exchange standing for the trading price in the securities exchange. The influence factors of the stock price volatility are very complex. Compared with the stock market of Europe and the United States, the management system of stock market is still improving gradually. Policy loophole may provide an opportunity for some experienced stock investors to obtain larger interests.

Attention is the event or character by the degree of concern, usually achieved by opinion polls or network voting. Meanwhile, the stock online attention means the attention of stock on the Internet. The Internet users search information about stock, so the bigger concern may stand for more information queries. Researchers have used the search volume on the internet to represent for the online attention.

Barber and Odean said that attention is a scarce resource in 2006 [2]. Firstly, individual investors select some stocks which they are interested in. Compared with the vast amount of information for people, our time and experience is limited. So the changes of online attention can reflect investors' psychology and behavior. For stock investors, this action may influence whether they buy or sell the stock. But when investors want to sell stocks, it happens to be change. Most individual investors hold relatively few common stocks in their hands. They can only sell the stock that they have already own. In short, attention is a major factor determining the stocks individual investors buy, but when selling, does not apply with equal force to investors' attention. Due to the high share of Baidu search engine in mainland, Baidu index can display main trend of the keywords search. The higher search volume will make more investors interest in it. According to the price pressure hypothesis, more investors' attention produces more stock needs.

3.2 Research model.

Based on the result of article mentioned above and the price pressure hypothesis, we put forward to four assumptions:

- There is positive correlation between search volume of stocker ticker and that of company name.
- Search volume which can be measured by Baidu Index can present attention of investors.
- Stock prices will present growth trend after search volume goes up.
- Current online attention increases will lead to latter rise of the stock price, recovering gradually in long time

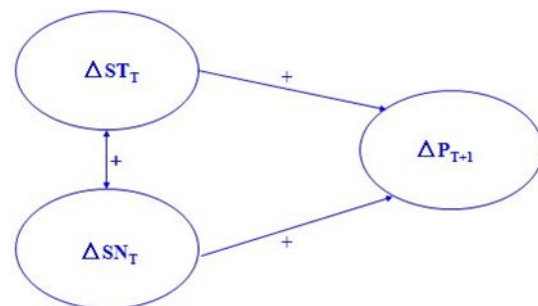


Figure 1. Hypothesis Model

Figure 1 shows us the hypothesis model of this paper. ΔST_T is the search volume variation of stock ticker in current time. ΔSN_T is the search volume variation of company name in current time. ΔP_{T+1} is the variation of stock price in the next time. The arrows signify action direction between variables and the '+' shows the positive relationship.

4. DATA COLLECTION AND ANALYSIS

This paper uses two parts of data including online attention data and stock related data. As RESSET supply many financial data for every year, we choose the stock price, trade volume and transaction sum to research the stock fluctuation. Because we use the weekly data of online attention, we choose the weekly stock data. On the other hand, in order to get the online attention data, we use a directly accessible method to interpret the fluctuation of it. Baidu index help us realize this purpose. When we check the index data through inputting stock ticker or company name as keyword, the web will return us a chart. In this chart, x axis shows the time and y

axis shows search volume for the keyword in every day. Finally, we get 1380 stocks with four variables from 2008 to 2011.

What shows in table 1 is the definition of variables. In the following passage, the variables will use for data analysis.

Table 1. Variable definition

Variable	Definition
st	search volume from Baidu index on stock ticker Sociology
d*Inst	the log of st during the week minus the log of previous * week
sn	search volume from Baidu index on company name
d*Insn	the log of sn during the week minus the log of previous * week
price	weekly closing price
vol	trading volume of every week
sum	transaction sum of every week

First, we separates the stock related data and search volume data for each year from 2008 to 2011. Then we use a short procedure to compute weekly data for every variable. And get 1380 stocks for 205 weeks points. Second, we try to choose a way to get the variation for every week. If using the D-value to represent the change between weeks, the current D-value will have strong relative with value of last week. We use time 1 to 205 to sign the whole time series. We use the stock tickers to remark 1380 cross-sections. Then we combine the seven variables (five operation variable and two identification variable) including time, id, d*Inst, d*Insn, price, vol and sum into one panel data table.

$$\Delta y_t = y_t - \frac{1}{n} * (y_{t-1} + \dots + y_{t-n}) \quad (1)$$

Table 2 shows the whole search volume data and the price number. Through data preprocess, we get 1380 stocks data finally. And table 3 states the composition of the sample. The valid data come from A-shares which come from SSE and SZSE. Forty-five percent is from SZSE and the other is SSE. There are 839 listed companies in SZSE excluding third board and 954 in SSE by the end of October 2012 [10, 11]. The number ratio in our sample is typical for all a-shares. So the sample can response overall situation accurately.

Table 2 Preprocessed Result

	price	st	sn
valid data	1380	1380	1380
total data	2207	2206	2222
percentage	62.5%	62.6%	62%

Table 3 Valid Data Source

stock type	number	ratio
A-share in SZSE	620	45.0%
A-share in SSE	760	55.0%
valid data	1380	100.0%

In order to understand the structure of variables, we do characterization analysis and summarize the result in table 4. We find the sn number is far more than st number. It proves that individual investors are used to search the stock information by company names. There are much information about the enterprise management situation and some financial events when you search the company names. This type information is interested and important for investors. It may be the reason why there is a great deal of search records of sn. So we use the sn as the direct proxy of online attention in the following analysis.

Table 4 Valid Data Characterization Analysis

	n	mean	max	min	sd
st	282900	371.91	11644	0	371.0077
sn	282900	405.1687	89441	0	1818.714
price	282900	13.03486	275.53	1.15	10.8336
vol	274061	46300000	2750000000	4400	74800000
sum	274061	539000000	3950000000	23320	951000000

After the elementary characterization analyze, we turn the data into panel data. Through the description analysis using the panel data from 2008 to 2011, we can understand the main characteristics of sample data. Avoiding repetition, we only show one result of 2008 in table 5. We can see from the analysis that the value of variables has significant fluctuations. However, it will make fixed effect model more suitable.

Table 5 Panel Data Characterization Analysis

Variable	Mean	Std. Dev.	Min	Max	Observations
price: overall	12.63956	13.02588	1.15	275.53	N=70533
between	--	10.39875	2.92	156.4437	n=1380
within	--	7.849601	58.02417	183.7345	t=51
sn: overall	234.4299	878.628	0	31616	N=70533
between	--	838.743	0	17550.06	n=1380
within	---	268.972	9540.688	14300.37	t=51
vol: overall	3.11e+07	5.82e+07	4400	2.45e+09	N=70533
between	--	4.55e+07	10578	6.86e+08	n=1380
within	--	3.58e+07	4.32e+08	1.80e+09	t=51
sum: overall	3.53e+08	7.99e+08	23320	3.95e+10	N=70533
between	--	6.37e+08	52378.93	1.32e+10	n=1380
within	--	4.75e+08	8.39e+09	2.67e+10	t=51

5. PANEL DATA ANALYSIS

5.1 Estimation model selection.

Stata data analysis software supply two analysis model for panel data analysis, fixed effect model (FE) and

random effect model (RE). Fixed effect model that includes individual effects variable is endogenous, and on the contrary, random effect model assumes that all contain individual random effects regression variable is exogenous; Random effect model intercept item in theory the same, produce difference reason is that intercept term random disturbance, individual effect as a random variation factors to consider. The characteristics of stock determine that it is more suitable for fixed effect model hypothesis.

Stata supply Hausman test to help reconfirm FE model. Table 7 shows the result of Hausman test. The $\chi^2(4)$ of sigmamore is 371.01 and $\chi^2(4)$ of sigmaless is 368.59. Both prob. value are 0.0000, the result refuses the original hypothesis. There is no systematic error of fixed effect. So we use FE for subsequent analysis.

Table 6 Hausman Inspection Result

	no addition option	sigmamore	sigmaless
chi2(4)	243.72	371.01	368.59
Prob.>chi2	0.0000	0.0000	0.0000

5.2 Fixed effects model principle

This section introduces the principle and analysis method of fixed effect model. The panel data has time and cross-section dimensions, so use $y_{i,t}$ to present explained variable. The subscript i and t denotes time and cross-section respectively. As shown on formula 2, $x_{i,t}$ is explaining variable with change of time and individual. The z_i is individual characteristic without change overtime. And it constitutes a perturbation terms, including unobservable individual characteristic without change overtime and perturbation terms with individual time change.

$$y_{i,t} = \alpha + x'_{i,t}\beta + z'_i\gamma + v_i + \varepsilon_{i,t} \quad (2)$$

Both sides on time dimension:

$$\bar{y}_i = \alpha + \bar{x}'_i\beta + z'_i\gamma + v_i + \bar{\varepsilon}_i \quad (3)$$

Use the OLS in formula 3 we will get the 'component estimator'. When individual characteristic is related to explain variable, Formula 2 is called fixed effect model.

Formula 2 minus formula 3:

$$(y_{i,t} - \bar{y}_i) = (x_{i,t} - \bar{x}_i)' \beta + (\varepsilon_{i,t} - \bar{\varepsilon}_i) \quad (4)$$

Use OLS estimate for formula 4, we can get FE estimator. The above is the principle of FE estimation. The disadvantage of FE is that it can't achieve coefficient estimate for no changing with time variables. So before FE estimation, we have to ensure that variables are changing with time. It is suitable the FE estimate conditions for our study variables.

5.3 Fixed effects analysis result

We do the panel data analysis in fixed effect model and get the results in table 7. The Prob. is 0.0000 in every panel data analysis. It means that the panel data analysis is significant. R-sq shows the interpretation extent of variable change. For instance, the R-sq between Δ pricet and Δ SNt is 0.0475 in 2008. It means that the search volume variation can account for 4.75% fluctuation of stock price at that time. Most R-sq is more than 0.01. This result verifies that price, vol and sum are related to online attention.

Table 7 Panel Data Analysis Result A

ΔSN_t		$\Delta price_t$	$\Delta price_{t+1}$	Δvol_t	Δvol_{t+1}	Δsum_t	Δsum_{t+1}
2008	coef	0.1029	0.0050	1.4081	-0.3838	1.5019	-0.4100
	R-sq	0.0475	0.0040	0.2804	0.0320	0.2845	0.0330
	Prob.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2009	coef	0.1121	0.0043	1.3494	-0.2588	1.4487	-0.2743
	R-sq	0.1683	0.0004	0.3614	0.0179	0.3723	0.0175
	Prob.	0.0000	0.0000	0.00000	0.0000	0.0000	0.0000
2010	coef	0.0597	-0.0400	1.6023	-1.4602	1.6503	-1.5271
	R-sq	0.1094	0.0658	0.6029	0.2356	0.6156	0.2444
	Prob.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2011	coef	0.0459	-0.0239	1.3585	-1.1517	1.3947	-1.1934
	R-sq	0.0689	0.0673	0.5295	0.2858	0.5383	0.2971
	Prob.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Then we use table 8 to make the result clear. Same as American stock market, the search volume of stock ticker has significant positive correlation with that of company name in China. Based on the existing theories, the stock price will goes up and recover in a long time while online attention goes up. In 2010 and 2011, stock price goes up in search volume ascending week, but it falls down in latter week. Meanwhile, from 2008 to 2011, the increasing of search volume makes trading volume and transaction sum increase in the first week and decline in next week. Theoretically, more individual investors spend their limited attention resource mean that more investors want to buy the stock. The more investors want to buy, the higher price it will be get. But on the contrary, high attention only makes price ascending in a very short time. After that, it pushes price down. However, the result doesn't meet our reasoning. The results don't confirm to stock market rule in some year. Perhaps illegal cash leads to the abnormal phenomenon in 2010 and 2011. The reason why the three targets descend in the next week is probably that someone controls the stock related information and it makes far more people pay attention to this stock. When the price gets higher, they will get the illegal cash.

Table 8 Panel Data Analysis Result B

ΔSN_t		$\Delta price_t$	$\Delta price_{t+1}$	Δvol_t	Δvol_{t+1}	Δsum_t	Δsum_{t+1}	
2008	2008	+	+	+	-	+	-	
	2009	+	+	+	-	+	-	
		2010	+	-	+	-	+	-
		2011	+	-	+	-	+	-

6. CONCLUSIONS

In this paper, we use a direct measure of investor attention, search volume in Baidu (ST and SN). In the sample of 1383 stocks in SZSE and SSE from 2008 to 2011, first, we collect search volumes separately using

company names and stock tickers as the keywords and study relationship between them. We find that the searching volume of company name is far more than that of stock code for one stock. After some detail analysis, we find that search using company name can get more information about the stock and company. Therefore we chose search volume of company name to represent individual investors' attention. Then we analyze the search volume and stock price, trading volume, transaction sum weekly. We test the price pressure hypothesis, stock price, trading volume, transaction sum all get higher when the investor attention goes higher in one week. In addition, the stock price also has positive correlation with investors' attention in last two weeks in 2008 and 2009. But there is a negative correlation between stock price and investor attention in the second week in 2010 and 2011. It is an abnormal phenomenon that investors pay attention to one stock, making more investors want to buy this stock, but the price goes higher and fell in short terms. There may be some people use illegal method to make investors pay attention to one stock. Then they sell the stock when investors push the stock price higher. The abnormal stock price fluctuation indicates that there are people who use this method to gain illegal profit.

ACKNOWLEDGEMENT

This research is partially supported by NKBRPC (2014CB340506) and the NSFC(71225003).

REFERENCES

- [1] Brad M. Barber, Terrance Odean, Ning Zhu. (2009). Do retail trades move markets?: *Review of Financial Studies*, 21: 151-186
- [2] Brad M. Barber, Terrance Odean. (2006). All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors: *Review of Financial Studies*, 21(2): 785-818
- [3] Thomas Chemmanur, An Yan. (2009). Advertising, Attention, and Stock Returns: Available at SSRN: <http://ssrn.com/abstract=1340605> or <http://dx.doi.org/10.2139/ssrn.1340605>
- [4] Thomas Dimp, Stephan Jank. (2011). Can internet search queries help to predict stock market volatility?: Available at SSRN :http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1941680
- [5] Zhi Da, Joseph Engelberg, Pengjie Gao. (2010). The Sum of All Fears: Investor Sentiment and Asset Prices: <http://public.kenan-flagler.unc.edu/faculty/engelbej/FEARS.pdf>
- [6] Foucault, Sraer, Thesmar. (2011). Individual Investors and Volatility: *The Journal Of Finance*, 66(4):1369-1406
- [7] Ilaria Bordino, Stefano Battiston, Guido Caldarelli. (2011). Web search queries can predict stock market volumes: <http://arxiv.org/abs/1110.4784>
- [8] Zhi Da, Joseph Engelberg, Pengjie Gao. (2011). In Search of Attention: *The Journal Of Finance*, 66 (5): 1461-1499
- [9] Corwin, Shane, Jeffrey H. Harris, and Marc L. Lipson. (2010) The development of secondary market liquidity for NYSE-listed IPOs: *Journal of Finance*, 59: 2339-2373
- [10] Da, Zhi, Joseph Engelberg, and Pengjie Gao. (2010). In search of fundamentals: Working paper, University of Notre Dame and University of North Carolina at Chapel Hill
- [11] (2012). <http://www.szse.cn/>
- [12] (2012). <http://www.sse.com.cn/sseportal/ps/zhs/home.html>