# A new attention proxy and order imbalance: Evidence from China

Gao, Ya; Xiong, Xiong; Feng, Xu; Li, Youwei; Samuel A. Vigne

## Abstract

We find a new and direct proxy for investors' attention in the Chinese stock market: daily abnormal reading quantity of each stock's posts in "guba" of Eastmoney website ($Ab\,GB1$). Using samples of Shanghai Stock Exchange A-shares, we find $Ab\,GB1$ (1) is significantly correlated to existing attention proxies; (2) leads to contemporarily high return and long-time reversal; (3) is related to heterogeneous trading behaviours of different investors. In summary, we find $Ab\,GB1$ is an efficient attention proxy and it can be useful for government to guide investors' trading.

Key words: investors' attention; heterogeneous trading behaviour; Chinese stock market

JEL: G10, G12, G41

## 1. Introduction

Attention is a scarce resource (Kahneman [1973]) and investors especially individual investors have limited attention. Things that attract their attention may have some influence on their trading behaviour.

There have been various proxies for investors' attention, such as abnormal trading volume (Gervais, Kaniel, and Mingelgrin [2001], Barber and Odean [2008], Da, Engelberg and Gao [2011]), extreme return (Seasholes and Wu [2007], Barber and Odean [2008], Da, Engelberg and Gao [2011]), news or advertisement (Hirshleifer and Teoh [2003], Tetlock [2011], Lou [2014], Yuan [2015]) and aggregated search frequency (Da, Engelberg and Gao [2011], Ben-Rephael, Da and Israelsen [2017]). Based on the fact that social media has

become a popular venue for individuals to share their analysis of financial securities (Chen et al. [2014]), we find a new proxy ($Ab\ GB1$) using a popular investors' social media: "guba" of Eastmoney website[1].

The reason why we want to find a new attention proxy is as follows: Firstly, as mentioned in Da, Engelberg and Gao (2011), former proxies like volume, return and news are indirect proxies. They have an assumption that if the volume and return are extreme or a stock is mentioned in news or advertisement, then investors will pay attention to that stock. However, volume and return can be driven by factors unrelated to investors' attention. News or advertisements can only be efficient when investors actually read them. We need a more direct measurement of investors' attention. Secondly, even though Da, Engelberg and Gao (2011) give a direct proxy by using aggregated search frequency in Google, and we also have a similar index by using Baidu search engine, this proxy is still inconvenient and inefficient for Chinese investors. Investors may search stocks in the engine to know news and fundamental changes, but if they want to know how many people are also interested in the same stock, they have to turn to a specific website[2] and search again. This proxy can be a great post hoc analysis indicator, but has less influence on investors' current attention. What's more, investors may have more interest in the website that gives comprehensive information about stocks and provides a place to discuss their interested information. Activity frequency in that website may be more intuitive to measure investors' attention. There are several analogous websites in China, but Eastmoney is the most popular and influential one. "guba" themed community in Eastmoney provides a place for investors to give posts about their interested stocks, read others' posts and give comments. Besides, "guba" gives sub-communities named by stocks' name, so investors can easily find relevant information about stocks they are interested.

---

[1] http://guba.eastmoney.com
[2] http://index.baidu.com/?from=pinzhuan

We find $Ab\ GB1$ is strongly correlated to existing attention proxies and has significant influence on stock return, which indicates our proxy can be reasonable. According to Baber and Odean (2008), individual investors will be net buyers if their attention is attracted and we also find that result.

The remainder of this paper is as follows. Section 2 describes data and variables. Section 3 gives empirical results of our attention proxy. Section 4 performs robustness test and Section 5 concludes.

## 2. Data and variables

In this paper, we have two main samples. The first one is from 2011.05.26 to 2017.02.21, using A-shares which have at least 600 trading days.[3] We have 961 stocks in this sample and it can be a good representative for the whole market. The second sample is randomly selected 200 stocks during 2013.08.01 to 2014.07.31 and they occupies 35.71% of A-shares' daily volume. Transaction data is from the Thomson Reuters Tick History database and CSMAR database, including stock price, trading volume, market capitalization, book to market ratio, turnover and spread. News and posts data include time, their content, number of people read them and number of comments. We grab them from "Eastmoney" website by Matlab program using stock code as key words.

We get detailed account data of 200 stocks from the Shanghai Stock Exchange, including specific twelve categories investors and their daily transactions. We have five individual investors ($Ind$), six professional institutional investors ($Pro$) and ordinary institutional investors ($OI$). Individual investors are divided by capital size: from $Ind1$ to $Ind5$, they hold less than 0.1 million, between 0.1 and 1 million, between 1 and 3 million, between 3 and 10 million, and more than 10 million CNY, respectively. And from $Pro1$ to $Pro6$, they are Investment Fund, QFII, Insurance Fund, Self-Brokerage, Asset Management

---

[3] The maximum trading day is 1397 and average value is 1245, so we choose 600 as threshold.

Agency, and Social Insurance Fund, respectively. We give description statistics of these

investors in table 1.

Table 1 description statistics different investors

| | Trading ratio (%) | | | | Investors | $N$ | $Buy$(million) | $Sell$(million) |
|---|---|---|---|---|---|---|---|---|
| | $Ind$ | $OI$ | $Pro$ | $Others$[4] | $Ind1$ | 48004 | 7.189 | 6.967 |
| 2016 | 85.62 | 1.41 | 12.21 | 0.75 | $Ind2$ | 48004 | 4.629 | 4.481 |
| 2015 | 86.91 | 2.06 | 10.47 | 0.56 | $Ind3$ | 48004 | 2.546 | 2.448 |
| 2014 | 85.19 | 2.98 | 11.6 | 0.22 | $Ind4$ | 46929 | 0.745 | 0.723 |
| 2013 | 82.24 | 2.46 | 15.3 | | $Ind5$ | 44563 | 1.296 | 1.279 |
| 2012 | 80.78 | 2.1 | 17.12 | | $Pro1$ | 40939 | 1.266 | 1.681 |
| 2011 | 83.52 | 2.09 | 14.39 | | $Pro2$ | 37912 | 1.059 | 1.227 |
| 2010 | 84.59 | 2.43 | 12.98 | | $Pro3$ | 16150 | 0.542 | 0.43 |
| 2009 | 85.36 | 3.82 | 10.82 | | $Pro4$ | 44304 | 0.982 | 1.031 |
| 2008 | 83.21 | 3.96 | 12.83 | | $Pro5$ | 37324 | 0.349 | 0.312 |
| | | | | | $Pro6$ | 8885 | 0.206 | 0.217 |
| | | | | | $OI$ | 47631 | 0.532 | 0.795 |

Trading ratio is the percentage of trading from different investors to total transaction.

During 2008 to 2016, more than 80% trading is from individual investors and professional

institutional investors take the second place.

The right part of table 1 is the number of observations ($N$), daily buy volume ($Buy$) and

sell volume ($Sell$) of different investors. In average, trading from individual investors is larger

and trading frequency of individual investors is higher. Trading ratio in our second sample is

81.05%, 7.39% and 11.57%, which is similar to the integral market.

We also find transaction frequency and trading volume are decreasing with individual

investors' capital size, those who hold the least capital take the majority part. Compared to

institutional investors, individual investors (especially for those who hold small capital) have

less information and they are easily influenced by attention-grabbing events. What's more,

individual investors are main users of "guba", therefore, our proxy may have good effects in

describing their attention.

Our attention proxy $Ab\ GB1$ is defined as

$$Ab\ GB1_t = GB1_t - median(GB1_{t-1} \dots GB1_{t-10}) \quad (1)\ [5]$$

---

[4] Others is "Shanghai-Hong Kong Stock Connect". Data of trading ratio is from the statistical yearbook of the Shanghai Stock Exchange.
[5] Similar standardized method is used in Da, Engelberg and Gao [2011].

$GB1$ is logarithm value of daily reading quantity of each stock's posts in "guba".

We use abnormal trading volume ($Ab\ volume$), Absolute value of abnormal return ($Abs\ ab\ ret$) and abnormal reading quantity of each stock's news ($Ab\ News1$) as representatives of existing attention proxies. They are defined as follows:

$$Ab\ volume_t = volume_t - \frac{1}{244}\sum_{i=1}^{244} volume_{t-i} \qquad (2)$$

$volume$ is logarithm value of trading volume, we use the similar way in Baber and Odean (2008) to do standardization.

$$Abs\ ab\ ret = abs(ret - rm) \qquad (3)$$

$ret$ is stock return calculated by close price and $rm$ is average value of $ret$ in each day.

$$Ab\ News1_t = News1_t - median(News1_{t-1}\dots News1_{t-10}) \qquad (4)$$

$News1$ is logarithm value of reading quantity of each stock's news. Considering $News1$ is not daily data, $t$ is the number of days when there is news.

We also have daily reply quantity of each stock's posts ($GB2$) / news (News2), daily number of each stock's posts ($GB3$) / news (News3), and correlations of these variables are in table 2.

Table 2 correlations between posts and news

|  | $GB1$ | $GB2$ | $GB3$ | $News1$ | $News2$ | $News3$ |
|---|---|---|---|---|---|---|
| $GB1$ | 1.000 |  |  |  |  |  |
|  | (1.00)[6] |  |  |  |  |  |
| $GB2$ | 0.644 | 1.000 |  |  |  |  |
|  | (1.00) | (1.00) |  |  |  |  |
| $GB3$ | 0.836 | 0.756 | 1.000 |  |  |  |
|  | (1.00) | (1.00) | (1.00) |  |  |  |
| $News1$ | 0.132 | 0.097 | 0.130 | 1.000 |  |  |
|  | (0.89) | (0.76) | (0.87) | (1.00) |  |  |
| $News2$ | 0.137 | 0.121 | 0.149 | 0.912 | 1.000 |  |
|  | (0.89) | (0.85) | (0.91) | (1.00) | (1.00) |  |
| $News3$ | 0.125 | 0.099 | 0.129 | 0.959 | 0.890 | 1.000 |
|  | (0.86) | (0.77) | (0.87) | (1.00) | (1.00) | (1.00) |

For each stock, we calculate correlations of these variables, and coefficients in table 2 are average values. Correlations between $GB1$ to $GB3$ are relatively high and all correlations

---

[6] Values in parentheses of table 2 and table 3 are ratio of correlations significant at 95% confidence level to all correlations.

are significant at 95% confidence level. Situations of $News1$ to $News$ 3 are similar but with

higher correlations. In this paper, we use $GB1$ and $News1$ as main variables. Results

calculated by others are similar, for simplicity, they are not tabulated.

Logarithm value of market capitalization ($Log(cap + 1)$), book to market ratio ($B/M$),

turnover ($Turnover$) and spread ($Spread$) are control variables[7].

## 3. Empirical studies

To be a good proxy of investors' attention must meet several requirements:

(1)        Enough frequency, investors can obverse it every day

(2)        Have relatively high correlations with existing attention proxies

(3)        Along with high value of proxy, there may be temporarily higher prices in

current day and long-term reversal

Daily average value of all stocks' posts in "guba" is 17.30, so our proxy meets the first

requirement. We give correlations between $Ab\ GB1$ and existing proxies in table 3.

Table 3 correlations between four attention proxies

|  | $Ab\ GB1$ | $Ab\ volume$ | $Abs\ ab\ ret$ | $Ab\ News1$ |
|---|---|---|---|---|
| $Ab\ GB1$ | 1.000 (1.00) | | | |
| $Ab\ volume$ | 0.694 (1.00) | 1.000 (1.00) | | |
| $Abs\ ab\ ret$ | 0.249 (0.98) | 0.308 (0.99) | 1.000 (1.00) | |
| $Ab\ News1$ | 0.119 (0.88) | 0.082 (0.66) | 0.04 (0.35) | 1.000 (1.00) |

In general, correlations between our proxy and others are high. Average correlation

between $Ab\ GB1$ and $Ab\ volume$ is 0.694 and all of them are significant. Correlation

between $Ab\ GB1$ and $Abs\ ab\ ret$ is 0.249 and 98% is significant. Correlation between

$Ab\ GB1$ and $Ab\ News1$ is 0.119 and 88% is significant. Other relationships are reasonable:

correlation between $Ab\ volume$ and $Abs\ ab\ ret$ is 0.308 and 99% is significant; correlations

[7] Individual investors may prefer to small stocks, book to market ratio reflects potential development of stocks, turnover and spread are popular proxies of liquidity. We use them to control stock heterogeneity.

between $Ab\ News1$ and other two proxies are not very high, which is understandable

considering the infrequency release of public news.

According to Da, Engelberg and Gao (2011), we use vector auto-regressions model

($VARs$) to compare four attention proxies, and results are in table 4.

Table 4 vector auto-regression model of attention proxies[8]

| | Lagged one day | | | | | |
| | $Ab\ GB1$ | $Ab\ volume$ | $Abs\ ab\ ret$ | $Ab\ News1$ | $cons$ | $R^2$ |
|---|---|---|---|---|---|---|
| $Ab\ GB1$ | **0.378** | **0.363** | **-0.083** | 0.001 | **0.773** | **0.340** |
| | **(0.013)** | **(0.014)** | **(-0.003)** | (0.003) | **(0.195)** | **(0.013)** |
| $Ab\ volume$ | **0.088** | **0.757** | **-0.104** | **-0.006** | **3.133** | **0.572** |
| | **(0.005)** | **(0.008)** | **(-0.003)** | **(-0.002)** | **(0.129)** | **(0.010)** |
| $Abs\ ab\ ret$ | **0.172** | **0.587** | **-0.065** | **0.013** | **-10.002** | **0.141** |
| | **(0.015)** | **(0.019)** | **(-0.009)** | **(0.006)** | **(-0.316)** | **(0.005)** |
| $Ab\ News1$ | **0.143** | **0.157** | 0.002 | **0.137** | **-3.527** | **0.044** |
| | **(0.019)** | **(0.027)** | (0.011) | **(0.008)** | **(-0.413)** | **(0.003)** |

We include a constant and time trend in $VARs$. Independent variables are four

contemporary proxies, respectively, and dependent variables are lagged proxies. We do $VARs$

for each stock and average across stocks. We find $Ab\ GB1$ leads other attention proxies.

Coefficients on $Ab\ GB1$ are all positive and statistically significant, which suggests $Ab\ GB1$

captures investors' attention more timely than extreme return, volume or news. This is not

surprising: to the extent that individual investors trade only after paying attention to a stock

and price pressure / extreme volume can persist over day, $Ab\ GB1$ could lead $Ab\ volume$ and

$Abs\ ab\ ret$. It is also reasonable that before the release of public news, there has been some

indications about it, therefore, $Ab\ GB1$ can lead $Ab\ News1$. More general, after the release of

public news, there will be some discussions about it, so it is reasonable to find $Ab\ News1$ has

positive and significant influence on $Ab\ GB1$. $Ab\ volume$ is significantly and positively

related to current $Ab\ GB1$, consistent with the idea that investors continue to pay attention to

a stock after extreme volume. In the other direction, we find lagged $Abs\ ab\ ret$ is

significantly but negatively related to $Ab\ GB1$. This is likely due to mean-reversion of

---

[8] Values in parentheses from table 4 to table 7 are standard errors computed using Newey-West (1987). Coefficients and standard errors are in bold if coefficients are significant at 95% confidence.

$Ab\ GB1$ after extreme return when  $Ab\ GB1$ spikes.

The third characteristic is especially appropriate for the Chinese stock market than mature markets like the U.S stock market. According to Barber and Odean (2008) model, attention shocks lead to net buying by retail traders. Considering the large proportion of uninformed individual investors in China, it is possible to find abnormal high return and return reversal after a high value of attention proxy.

We use Fama-MacBeth (1973) cross-sectional regressions to study the influence of four attention proxies on contemporary and future returns and give results in table 5.

Table 5 the influence of four attention proxies on stock return

|  | $return_t$ | $return_{t+1}$ | $return_{t+2}$ | $return_{t+3}$ | $return_{t+4}$ |
|---|---|---|---|---|---|
| $Ab\ GB1$ | **0.273** | **-0.026** | 0.003 | 0.006 | 0.007 |
|  | **(0.010)** | **(0.008)** | (0.007) | (0.006) | (0.006) |
| $Ab\ volume$ | **0.574** | **-0.140** | **-0.084** | **-0.059** | **-0.062** |
|  | **(0.013)** | **(0.011)** | **(0.011)** | **(0.009)** | **(0.009)** |
| $Abs\ ab\ ret$ | **0.044** | 0.003 | **0.008** | **0.018** | **0.025** |
|  | **(0.004)** | (0.003) | **(0.003)** | **(0.003)** | **(0.003)** |
| $Ab\ News1$ | **0.367** | -0.013 | -0.221 | -0.184 | -0.112 |
|  | **(0.126)** | (0.110) | (0.128) | (0.127) | (0.099) |
| R-squared | 0.052 | 0.022 | 0.017 | 0.015 | 0.014 |

Dependent variable are abnormal return (in basis points) on current day and following four days. Abnormal return is the difference between stock return and average return of 961 stocks in that day. Independent variables are standardized so they can compare.

Influence of $Ab\ GB1$ is significant after control other proxies: it has positive influence on current return and reverses on the next day. In following three days, influence of $Ab\ GB1$ is insignificant. We can draw a conclusion that influence of $Ab\ GB1$ is instantaneous and will disappear with the release of new posts. Influence of $Ab\ volume$ is most significant, and has long-term reversal during following days; influence of $Abs\ ab\ ret$ is the smallest but lasts the longest time; influence of $Ab\ News1$ is significant only on current day.

Above all, we find $Ab\ GB1$ can be a good proxy for investors' attention in the Chinese stock market: it has daily frequency for investors to use every day; it is highly related to existing attention proxies and has significant influence on stock return.

Then we study the influence of $Ab\ GB1$ on order imbalance ($im$) of different investors

by using the time and firm fixed panel data regression (7). Results are in table 6.

$$im_{i,t} = \beta_0 + \beta_1 * Ab\ GB1 + \beta_2 * Log(cap + 1)_{i,t} + \beta_3 * B/M_{i,t} + \beta_4 * Spread_{i,t} + \beta_5 *$$

$$Turnover_{i,t} + \mu_{i,t} \tag{7}$$

$im$ is calculated as follows, where $buy\ volume_{i,t}$ and $sell\ volume_{i,t}$ are buy and sell

volume for each kind of investors on stock $i$, day $t$.

$$im_{i,t} = \frac{buy\ volume_{i,t} - sell\ volume_{i,t}}{buy\ volume_{i,t} + sell\ volume_{i,t}} \tag{8}$$

Other variables are control variables.

The influence of $Ab\ GB1$ on different investors' order imbalance are significantly

different. Individual investors who have less capital are net buyer with the increase of

$Ab\ GB1$, those who have more capital are insignificantly influenced and institutional

investors are net sellers with the increase of $Ab\ GB1$. Our results are consistent with Baber

and Odean (2008), who use abnormal trading volume, extreme return and news as attention

proxies, and find order imbalances of individual investors and institutional investors are

oppositely influenced. Based on results in table 6, we can also find individual investors prefer

stocks with low capital, high B/M, high spread and low turnover and institutional investors

have no specific preference. Which proves individual investors are uninformed and noise

traders, so they are possibly influenced by attention-grabbing events.

## 4. Robustness test

A presupposition of Barber and Odean (2008) model is that individual investors have no

access to short sale. They can only choose the few stocks that they own to sell, but they can

Table 6 the relationship between $Ab\ GB1$ and order imbalance

| | Ind1 | Ind2 | Ind3 | Ind4 | Ind5 | Pro1 | Pro2 | Pro3 | Pro4 | Pro5 | Pro6 | OI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ab GB1 | **1.559** | **1.714** | **0.991** | 0.420 | -0.242 | **-4.210** | **-3.148** | **-5.136** | **-1.844** | **-1.458** | **-8.670** | **-1.793** |
| | **(0.107)** | **(0.120)** | **(0.205)** | (0.445) | (0.511) | **(0.625)** | **(0.709)** | **(1.277)** | **(0.490)** | **(0.649)** | **(1.773)** | **(0.423)** |
| Log(cap + 1) | **-0.270** | **-0.368** | **-1.003** | -0.130 | -0.186 | 0.081 | 1.357 | 1.686 | 0.491 | -0.261 | 2.549 | 0.250 |
| | **(0.151)** | **(0.169)** | **(0.288)** | (0.621) | (0.700) | (0.804) | (0.999) | (1.819) | (0.657) | (0.859) | (2.022) | (0.592) |
| B/M | **0.757** | **0.883** | **0.829** | **1.530** | **1.488** | -1.031 | -0.089 | **-4.625** | 0.564 | -0.273 | -0.480 | 0.581 |
| | **(0.157)** | **(0.175)** | **(0.298)** | **(0.642)** | **(0.721)** | (0.844) | (0.963) | **(1.413)** | (0.688) | (0.868) | (1.897) | (0.614) |
| Spread | **0.181** | **0.189** | 0.092 | 0.089 | 0.299 | 0.003 | -0.163 | -0.844 | 0.142 | **-0.815** | -0.608 | 0.031 |
| | **(0.058)** | **(0.065)** | (0.111) | (0.239) | (0.267) | (0.320) | (0.351) | (0.467) | (0.258) | **(0.321)** | (0.534) | (0.228) |
| Turnover | **-0.333** | **-0.482** | **-0.459** | **-1.267** | **-1.926** | **3.804** | **-1.623** | 0.065 | -0.348 | 0.079 | **2.931** | -0.173 |
| | **(0.058)** | **(0.065)** | **(0.112)** | **(0.241)** | **(0.271)** | **(0.370)** | **(0.437)** | (0.816) | (0.279) | (0.355) | **(1.168)** | (0.230) |
| Constant | **-5.793** | 0.148 | **13.631** | -0.206 | 15.421 | **-34.982** | **47.022** | 9.812 | **38.831** | **34.761** | -24.232 | **14.773** |
| | **(1.833)** | (2.052) | **(3.493)** | (7.551) | (8.520) | **(10.141)** | **(12.542)** | (22.933) | **(8.142)** | **(11.144)** | (30.473) | **(7.189)** |
| Observations | 47760 | 47760 | 47760 | 46685 | 44320 | 40695 | 37674 | 15949 | 44060 | 37080 | 8817 | 47387 |
| R-squared | 0.064 | 0.034 | 0.014 | 0.009 | 0.016 | 0.079 | 0.074 | 0.045 | 0.140 | 0.147 | 0.162 | 0.028 |

Table7 the influence of $Ab\ GB1$ on individual investors' order imbalance

| | Ab GB1 | | | | | Ab GB1 Short sell | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Ind1 | Ind2 | Ind3 | Ind4 | Ind5 | Ind1 | Ind2 | Ind3 | Ind4 | Ind5 |
| Ab GB1 | **1.165** | **1.033** | -0.026 | -0.355 | -0.577 | **2.110** | **2.646** | **2.237** | **1.557** | 0.333 |
| | **(0.127)** | **(0.151)** | (0.301) | (0.655) | (0.775) | **(0.182)** | **(0.194)** | **(0.277)** | **(0.610)** | (0.681) |
| Diff | **0.945** | **1.613** | **2.263** | **1.912** | 0.910 | | | | | |
| | **(4.258)** | **(6.561)** | **(5.532)** | **(2.136)** | (0.882) | | | | | |
| Log(cap + 1) | **-0.507** | **-0.468** | **-0.828** | 0.039 | -0.089 | **-0.654** | **-0.821** | **-1.448** | -1.340 | -0.222 |
| | **(0.160)** | **(0.189)** | **(0.377)** | (0.802) | (0.921) | **(0.375)** | **(0.401)** | **(0.572)** | (1.291) | (1.435) |
| B/M | **2.477** | **1.889** | 0.275 | 1.793 | 0.508 | **0.774** | **0.953** | **0.919** | **1.396** | **1.938** |
| | **(0.358)** | **(0.424)** | (0.845) | (1.826) | (2.178) | **(0.195)** | **(0.208)** | **(0.297)** | **(0.653)** | **(0.724)** |
| Spread | **0.227** | **0.243** | 0.196 | 0.010 | 0.321 | 0.130 | **0.143** | 0.006 | 0.086 | 0.287 |
| | **(0.083)** | **(0.098)** | (0.196) | (0.421) | (0.482) | (0.081) | **(0.087)** | (0.124) | (0.272) | (0.301) |
| Turnover | **-0.129** | **-0.201** | **-0.241** | **-0.961** | **-1.706** | **-0.833** | **-1.165** | **-0.891** | **-1.975** | **-2.572** |
| | **(0.062)** | **(0.073)** | **(0.146)** | **(0.312)** | **(0.358)** | **(0.128)** | **(0.136)** | **(0.195)** | **(0.428)** | **(0.476)** |
| Constant | **-4.158** | 2.396 | **16.261** | 1.084 | **27.302** | -3.924 | 1.754 | **14.421** | 9.951 | 3.695 |
| | **(1.941)** | (2.299) | **(4.581)** | (9.844) | **(11.261)** | (4.237) | (4.533) | **(6.465)** | (14.531) | (16.162) |
| Observations | 24193 | 24193 | 24193 | 23213 | 21152 | 23567 | 23567 | 23567 | 23472 | 23168 |
| R-squared | 0.053 | 0.025 | 0.017 | 0.015 | 0.020 | 0.106 | 0.082 | 0.031 | 0.015 | 0.027 |

buy across thousands of stocks if their attention is attracted, therefore there might be a great order imbalance.

Barber and Odean (2008) eliminate the influence of short sale by studying stocks individual investors already hold. In this paper, we do a more direct research with the lift of short sale constrains in China after 2010. Based on short sale lists, we divide our sample into two groups and study whether there are some differences of $Ab\ GB1$'s influrnce on them.

Using the short sale list on 2013.01.31, we find 97 stocks in our sample can short sell and 102 can't. We use these sub-samples to repeat regression (7) and give results in table 7.

In general, individual investors prefer to buy rather than sell for stocks that attract their attention and that phenomenon is more significant for stocks can short sell. The number of stocks can and can't sell is similar, so we can do some comparison for their regressions.

The way we use to clarify whether there is a real gap is from Paternoster et al. (1998).[9] The general form is as follows.

$$Z = \frac{b_1 - b_2}{\sqrt{SEb_1^2 + SEb_2^2}} \qquad (9)$$

$b_1$ and $b_2$ are coefficients from two regressions. $SEb_1^2$ and $SEb_2^2$ are corresponding standard errors. Results are in column $Diff$ of table 7. Coefficients are differences between two regressions and values in parentheses are $Z$-values calculated using Paternoster et al. (1998). Influences of $Ab\ GB1$ on order imbalance are significantly larger for stock can be short, except for individual investors who hold the largest asset[10]. These results are deviated from the intuition that transactions on stocks can short sell may be less influenced by attention, but they are still reasonable. Individual investors only account for 24% of short sellers in this period [11] and ratio of short selling to total trading volume is about 4%. Short sale doesn't play a great role in investors' trading. What's more, stocks can short sell have

---

[9] Which is also used by Acharya and Xu (2017) and Xiong et al (2018).
[10] They might have similar trading behaviour as institutional investors considering their capital size.
[11] Threshold of short selling is 0.5 million, which is relatively high for individual investors.

larger size and better liquidity and they are easier to be mentioned in "guba".

Ab GB1 remains a good proxy for individual investors' attention when lifting short selling constraints and individual investors are net buyers of stocks that attract their attention.

## 5. Conclusion

In this paper, we use $Ab\ GB1$ as the proxy of investor's attention and find (1) it has a strong correlation with existing attention proxies such as $Ab\ News1$, $Abs\ ab\ ret$ and $Ab\ volume$; (2) it is efficient to predict contemporary stock return and future reversal and (3) it is highly related to heterogeneous trading behaviour of different investors, especially for individual investors. It can be a useful expansion of available investors' attention proxies in the Chinese stock market.

Our study also plays a great role in the government' guiding, as individual investors are the main composition of the Chinese stock market and they are imperfect processors of publicly available information because of their limited analytical abilities. It is helpful to give some instructions information in "guba" for investors to look through and have some discussions, which will make information spread faster.

## Acknowledgment

## Reference

[1]     Acharya, V., & Xu, Z. (2017). Financial dependence and innovation: The case of public versus private firms. Journal of Financial Economics, 124(2), 223-243.

[2]     Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies, 21(2), 785-818.

[3]     Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. The Review of Financial Studies, 30(9), 3009-3047.

[4]     Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. The Review of Financial Studies, 27(5), 1367-1403

[5]     Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. The Journal of Finance, 66(5), 1461-1499.

[6]     Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. Journal of political economy, 81(3), 607-636.

[7]     Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The high-volume return premium. The Journal of Finance, 56(3), 877-919

[8]     Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. Journal of accounting and economics, 36(1), 337-386.

[9]     Kahneman, D. (1973). Attention and effort (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

[10]    Lou, D. (2014). Attracting investor attention through advertising. The Review of Financial Studies, 27(6), 1797-1829.

[11]    Newey, W. K., & West, K. D. (1987). A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703-708.

[12]    Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. Criminology, 36(4), 859-866.

[13]    Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. Journal of Empirical Finance, 14(5), 590-610.

[14]    Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information?. The Review of Financial Studies, 24(5), 1481-1512.

[15]    Xiong, X., Gao, Y., & Feng, X. (2018). Successive short-selling ban lifts and gradual price efficiency: Evidence from China. Journal of Accounting and Finance, forthcoming.

[16]    Yuan, Y. (2015). Market-wide attention, trading, and stock returns. Journal of Financial Economics, 116(3), 548-564.