Finance Research Letters 29 (2019) 411–417

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

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

Gao Ya[a], Xiong Xiong[a,b], Feng Xu[a,b,*], Li Youwei[c], Samuel A. Vigne[d]

[a] *College of Management and Economics, Tianjin University, Tianjin 300072, China*
[b] *China Centre for Social Computing and Analytics, Tianjin 300072, China*
[c] *Hull University Business School, University of Hull, HU6 7RX UK*
[d] *Queen's Management School, Queen's University Belfast, BT9 5EE UK*

ARTICLE INFO

Keywords:
Investor attention
Heterogeneous trading behaviour
Chinese stock market
Eastmoney
Guba

JEL classification:
G10
G12
G41

ABSTRACT

In this paper, we propose a new direct proxy for investors' attention in the Chinese stock market: daily abnormal reading quantity of each stock's posts on the *Eastmoney guba* website. Using A-shares samples of the Shanghai Stock Exchange, we find that our proposed proxy (i) is significantly correlated to existing attention proxies; (ii) leads to contemporarily high returns and long-time reversal; (iii) is related to heterogeneous trading behaviour of different investors. In summary, we add value to the field of investor attention approximation with a new and efficient measure that can be useful for guiding and modelling investor's trading

1. Introduction

As Kahneman (1973) points out: attention is a scarce resource and investors – especially individual investors, have limited attention. Fellow academics have proposed various proxies for investors' attention, such as: (i) abnormal trading volume (see Gervais et al., 2001, Barber and Odean, 2008, and Da et al., 2011), (ii) extreme return (see Seasholes and Wu, 2007, Barber and Odean, 2008, Da et al., 2011, Hsu and Chen, 2017, and Hood and Lesseig, 2017), (iii) news or advertisement (see Hirshleifer and Teoh, 2003, Tetlock, 2011, Lou, 2014, Yuan, 2015, Yang et al., 2017, and Tsukioka et al., 2018), and finally, (iv) aggregated search frequency (see Da et al., 2011, Ben-Rephael et al., 2017, and Yung and Nafar, 2017).

Based on the fact that social media has become a popular venue for individuals to share their analysis and opinion on financial securities (Chen et al., 2014), we build a new proxy based on the popular Chinese social media website for investors: *Guba* from the *Eastmoney* website.[1]

The motivation of finding a new proxy stems directly from previous re- search and from the new technologies available today. Indeed, as mentioned in Da et al. (2011), former proxies like volume, return and news are indirect proxies. They all share the assumption that if the volume and the return of a stock faces extreme movements, or that a stock is mentioned in news or

☆ The authors would like to thank Xiangtong Meng for data support. Funding by the Natural Science Foundation of China (71532009, 71790594, 71871157) and the Ministry of Education Fund on Humanities and Social Science (14YJC790029) is gratefully acknowledged. The author Gao. Y also want to thanks for the support from the China Scholarship Council, this paper was finished when she was an exchanged Ph. D. student in Queen's University Belfast.

* Corresponding author at: College of Management and Economics, Tianjin University, Tianjin 300072, China.
E-mail addresses: gaoyatju@tju.edu.cn (Y. Gao), xxpeter@tju.edu.cn (X. Xiong), fengxu@tju.edu.cn (X. Feng), youwei.li@hull.ac.uk (Y. Li), s.vigne@qub.ac.uk (S.A. Vigne).

[1] http://guba.eastmoney.com/

https://doi.org/10.1016/j.frl.2018.11.009
Received 3 September 2018; Received in revised form 9 November 2018; Accepted 15 November 2018
Available online 20 November 2018
1544-6123/

advertisement, then investors will pay attention to that stock. However, volume and return can be driven by factors unrelated to investors' attention, and news or advertisements can only be efficient when investors really read them; pushing for the need for a more direct measurement of investors' attention.

Da et al. (2011) propose a direct proxy by relying on the aggregated search frequency in Google; a similar index exists using the Baidu search engine, but this proxy is inconvenient and inefficient for Chinese investors. Indeed, investors might rely on Baidu to find news on fundamental changes in the price of a stock (Kou et al. (2018)), but if they want to find out how many people are interested in the same stock, they have to rely on a more specific website.[2] This proxy can therefore serve as a post hoc analysis indicator, but has little influence on investors' current attention. Furthermore, many investors have more interest in a website that provides comprehensive information about stocks and provides a place to discuss their views; the activity frequency on that website is therefore more intuitive in measuring investors' attention. While there are many analogous websites in China, the *guba* themed community from *Eastmoney* is the most popular and influential one.

Our results indicate that an abnormal value of daily readings of information on a specific stock posted on *guba* is strongly correlated to existing attention proxies and has a significant influence on the return of the stock; indicating the reliability of our attention measure. Our results also offer sup- port to Barber and Odean (2008) who find that individual investors will be net buyers if their attention is attracted.

2. Data and variables

We rely on two main samples: the first spans from the 26th of May 2011 to the 21st of February 2017, relying on A-shares which have at least 600 trading days, while the second spans from the 1st of August 2013 to the 31th of July 2014. There are 961 stocks in the first sample, therefore being a good representation for the whole market, and a random selection of 200 stocks in the second sample, representing 35.71% of A-shares' daily volume. Transaction data has been obtained from the Thomson Reuters Tick History database and the CSMAR database, and includes stock price, trading volume, market capitalisation, book to market ratio, turnover, and spread. News and post data were obtained from the *Eastmoney* website and include time, content, number of reads, and number of comments.

We obtain detailed account data for 200 stocks from the Shanghai Stock Exchange, including twelve specific categories of investors and their daily transactions. We differentiate them into five categories of individual investors (*Ind*), six categories of professional institutional investors (*Pro*), and ordinary institutional investors (*OI*) based on exchange's classification.

Individual investors are divided by capital size: from *Ind*1 to *Ind*5, holding 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.

Professional institutional investors are divided by the nature of their institution: from *Pro*1 to *Pro*6, standing for Investment Funds, Qualified Foreign Institutional Investor, Insurance Fund, Self-Brokerage, Asset Management Agency, and Social Insurance Funds, respectively.

Descriptive statistics are provided in Table 1.

The trading ratio is the percentage of trading from different investors to total transactions. Between 2008 and 2016, more than 80% of trading is from individual investors. The number of observations (*N*), daily buy volume (*Buy*), and daily sell volume (*sell*) are also displayed. On average, trading from individual investors is larger and trading frequency of individual investors is higher.

Transaction frequency and trading volume are decreasing with individual investor's capital size; compared to institutional investors, individual investors (especially those who hold small amounts of capital) have less information and are easily influenced by attention-grabbing events. Furthermore, individual investors are the main users of the *guba* platform, underlining again that our measure is a good proxy in describing their attention to stock market news.

We define our attention proxy as the abnormal amount of daily reads of a stock post *GB*1.[3]

$$Ab\ GB1_t = GB1_t - median(GB1_{t-1}...GB1_{t-10}) \tag{1}$$

*GB*1 is therefore the logarithmic value of daily reading quantity of each stock post on the *guba* website.

We also rely on abnormal trading volume (*Ab volume*), the absolute value of abnormal return (*Abs ab ret*), and the abnormal reading quantity of an individual stock's news coverage (*Ab News*1) as existing attention proxies.

The variables are defined as follows:

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

where we rely on the logarithmic value of trading volume in a way similar to Barber and Odean (2008) to standardise values.[4]:

$$Abs\ ab\ ret = abs(ret - rm) \tag{3}$$

where *ret* is the return of a stock calculated in regard to the closing price, and *rm* is the average daily value of *ret*.

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

[3] A similar standardised method is used in Da et al. (2011); we also use other lag lengths in Eqs. 1 and 4, results are similar.

[4] The reason we choose 244 is to reflect the average annual trading days in the Chinese stock market, which is comparable to Barber and Odean (2008) using average annual trading days in the U.S. stock market in their paper

Table 1
Description statistics of different investors.

Trading ratio (%)					Investors	*N*	*Buy*(million)	*Sell*(million)
	Ind	*OI*	*Pro*	*Others*	*Ind*1	48,004	7.189	6.967
2016	85.62	1.41	12.21	0.75	*Ind*2	48,004	4.629	4.481
2015	86.91	2.06	10.47	0.56	*Ind*3	48,004	2.546	2.448
2014	85.19	2.98	11.60	0.22	*Ind*4	46,929	0.745	0.723
2013	82.24	2.46	15.30		*Ind*5	44,563	1.296	1.279
2012	80.78	2.10	17.12		*Pro*1	40,939	1.266	1.681
2011	83.52	2.09	14.39		*Pro*2	37,912	1.059	1.227
2010	84.59	2.43	12.98		*Pro*3	16,150	0.542	0.430
2009	85.36	3.82	10.82		*Pro*4	44,304	0.982	1.031
2008	83.21	3.96	12.83		*Pro*5	37,324	0.349	0.312
					*Pro*6	8885	0.206	0.217
					OI	47,631	0.532	0.795

Table 2
Correlation matrix between posts and news.

	*GB*1	*GB*2	*GB*3	*News*1	*News*2	*News*3
*GB*1	1.000					
	(1.00)					
*GB*2	0.644	1.000				
	(1.00)	(1.00)				
*GB*3	0.836	0.756	1.000			
	(1.00)	(1.00)	(1.00)			
*News*1	0.132	0.097	0.130	1.000		
	(0.89)	(0.76)	(0.87)	(1.00)		
*News*2	0.137	0.121	0.149	0.912	1.000	
	(0.89)	(0.85)	(0.91)	(1.00)	(1.00)	
*News*3	0.125	0.099	0.129	0.959	0.890	1.000
	(0.86)	(0.77)	(0.87)	(1.00)	(1.00)	(1.00)

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

where *News*1 is the logarithmic value of the number of reads of each stock's news. In light that *News*1 is not daily, t is the number of days for which there is news.

We also consider the quantity of daily replies to posts on a particular stock (*GB*2) and the amount of news (*News*2), as well as the quantity of daily posts on a particular stock (*GB*3) and the amount of news (*News*3).

A correlation matrix of these variables is provided in Table 2.[5]

The correlation between *GB*1 and *GB*3 is relatively high and all correlations are significant at the 95% confidence level. Correlation between *News*1 and *News*3 is also very high. In this paper, we use variables *GB*1 and *News*1 as the main variables, results obtained through the other variables are very similar and available on demand.

The logarithmic value of market capitalisation ($Log(cap + 1)$), the book to market ratio (*B/M*), turnover (*Turnover*), and spread (*Spread*) are used as control variables.[6]

3. Empirical studies

In order to serve as a good proxy for investors' attention, the variable needs to fulfil certain requirements:

- Frequency: Investors need to be able to observe the measure on a daily basis.
- Generalisability: The measure needs to be correlated with other attention proxies.
- Correlation: The measure needs to adequately represent changes in the financial market.[7]

[5] The values in parentheses are the ratio of correlations significant at the 95% confidence level.
[6] The control variables are used to control for stock heterogeneity.
[7] We propose these three requirements based on trading experience and previous studies. Firstly, a good investor attention proxy should be computable at a high-frequency, which can be available for investors every day, one of advantage of our proxy. The second requirement is in line with Da et al. (2011), who also study similar relationships. A good news proxy should be related to existing proxies, which proves the reasonability of this new proxy. The third requirement is derived from results in Barber and Odean (2008), who find attention pushes investors' buy transactions and high return temporarily, and in the long term the stock price will go back to its' fundamental value.

Table 3
Correlation between the four attention proxies.

	*Ab GB*1	*Ab volume*	*Abs ab ret*	*Ab News*1
*Ab GB*1	1.000			
	(1.00)			
Ab volume	0.694	1.000		
	(1.00)	(1.00)		
Abs ab ret	0.249	0.308	1.000	
	(0.98)	(0.99)	(1.00)	
*Ab News*1	0.119	0.082	0.040	1.000
	(0.88)	(0.66)	(0.35)	(1.00)

Table 4
Vector auto-regression model of attention proxies.

	Lagged one day					
	*Ab GB*1	*Ab volume*	*Abs ab ret*	*Ab News*1	*cons*	R^2
*Ab GB*1	**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 News*1	**0.143**	**0.157**	0.002	**0.137**	**−3.527**	**0.044**
	(0.019)	**(0.027)**	(0.011)	**(0.008)**	**(0.413)**	**(0.003)**

There is an average amount of 17.3 daily posts per stock on the *guba* website, hence meeting the first requirement highlighted above.

We display the correlation matrix between *Ab GB*1 and the other proxies in Table 3[8]

Table 3 reveals that the correlations between these proxies are high, and mostly significant. The only exception is between *Ab News*1 and the other proxies, which is also reasonable due to the infrequency release of public news.

We follow Da et al. (2011) and rely upon a vector auto-regression model (VAR) to compare our four attention proxies; results are displayed in Table 4.[9]

A constant and a time trend is included in the VAR. The four proxies serve as independent variables, while their lagged values are the dependent variables. A VAR is estimated for each stock and the average is obtained. We find that *Ab GB*1 leads the other proxies, and therefore captures investors' attention more timely than extreme return, volume, or news. We explain this by arguing that Chinese individual investors tend to trade only after paying attention to a stock, while price pressure and extreme volume can persist over several days. We can also reasonably assume that before the release of public news, there has been some indication about it in exactly such websites as *guba; Ab GB*1 therefore leads *Ab News*1. More generally, after the release of public news, there will be some discussion about it, explaining why *Ab News*1 has a positive and significant influence on *Ab GB*1. The positive and significant relationship between *Ab volume* and *Ab GB*1 underlines the idea that investors continue to pay attention to a stock after extreme volume. Finally, the negative and significant relationship between *Abs ab ret* and *Ab GB*1 is likely due to mean-reversion of *Ab GB*1 after extreme returns when *Ab GB*1 spikes. This is particularly appropriate for the Chinese market, as Barber and Odean (2008) argue that attention shocks lead to net buying by retail traders. Considering the large proportion of uninformed traders in China, it is understandable to observe abnormal high return and return reversal after a high value of the attention proxy.

Relying upon Fama and MacBeth (1973) cross-sectional regressions, we study the influence of four attention proxies on contemporary and future returns to meet the third requirement. Results are displayed in Table 5.[10]

The dependent variable is specified as the abnormal return (in basis points) on the event day and the following four days, with the four attention proxies being the independent variables. Abnormal return is defined as difference between a stock's return and average return computed daily by the 961 stocks in the sample. The independent variables are standardized in order to be comparable. After controlling for the other proxies, the influence of *Ab GB*1 is found be significant. *Ab GB*1 has a positive influence on current return on the event day, while this relationship is reversed the following day - for the three days after that, the relationship becomes

[8] Statistical significance is provided in parentheses.

[9] Values in parentheses are standard errors computed following Newey and West (1987). Coefficients and standard errors are in bold if significant at the 95% confidence interval.

[10] Values in parentheses are standard errors computed following Newey and West (1987). Coefficients and standard errors are in bold if significant at the 95% confidence interval.

Table 5
Influence of four attention proxies on stock return.

	$return_t$	$return_{t+1}$	$return_{t+2}$	$return_{t+3}$	$return_{t+4}$
*Ab GB*1	**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 News*1	**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

insignificant. We can therefore conclude that the influence of *Ab GB*1 is instantaneous and will disappear with the release of new posts. The influence of *Ab volume* is the most significant and has long-term reversal during the following days, while the influence of *Abs ab ret* is small, it is the most durable. Finally, the influence of *Ab News*1 is only significant on the event day.

*Ab GB*1 is therefore a good proxy for investors' attention in the Chinese stock market: it has daily frequency, is highly correlated to existing attention proxies, and has significant influence on stock return.

In a last step, we study the influence of *Ab GB*1 on order imbalance *im* for different investors by relying upon a time and firm fixed panel data model:

$$im_{i,t} = \beta_0 + \beta_1 {*}Ab\ GB1 + \beta_2 {*}Log(cap + 1)_{i,t} + \beta_3 {*}B/M_{i,t} + \beta_4 {*}Spread_{i,t} + \beta_5 {*}Turnover_{i,t} + \mu_{i,t} \tag{5}$$

where *im* is calculated as follows:

$$im_{i,t} = \frac{buy\ volume_{i,t} - sell\ volume_{i,t}}{buy\ volume_{i,t} + sell\ volume_{i,t}} \tag{6}$$

while buy and sell volumes are considered for each type of investor in stock *i* on day *t*. The other variables are control variables; we report results in Table 6.[11]

The influences of *Ab GB*1 on different investors' order imbalance vary significantly. Individual investors with less capital are net buyers under an increase of *Ab GB*1, while those with more capital are insignificantly influenced, and institutional investors are net sellers under and increase of *Ab GB*1. Our results are indeed consistent with Barber and Odean (2008), who use abnormal trading volume, extreme returns and news as attention proxies and find that order imbalances of individual investors and institutional investors are oppositely influenced. Based on results in Table 6, we can also find that individual investors prefer stocks with low capital, high B/M ratio, high spread and low turnover; while institutional investors have no discernible preferences. This is further evidence of individual investors being uninformed traders acting on noise, which are therefore likely to be influenced by attention-grabbing events.

4. Robustness test

A presupposition of the Barber and Odean (2008) model is that individual investors don't have access to short sale; they can only choose to sell the few stocks they own, but can choose from several thousands of stocks to buy – leading to an order imbalance.

Barber and Odean (2008) eliminate the influence of short sale by studying stocks that individual investors already hold; we propose a more direct research in the light of the lift of the short sale constrains in China after 2010. Based on short sale lists, we divide our sample into two groups and study whether *Ab GB*1 influences them differently. We rely upon the short sale list of the 31st of January 2013, and identify 97 stocks in our sample for which short selling is allowed, and 102 for which short selling is illegal. We apply Eq. (5) on both sub-samples and display results in Table 7.[12]

In general, individual investors prefer to buy rather than sell high attention stocks, and interestingly, that phenomenon is more pronounced for stocks that can be sold short. In order to truly identify whether or not there is a true difference between stocks that allow short selling and those that do not, we rely upon the methodology of Paternoster et al. (1998) .[13]

$$Z = \frac{b_1 - b_2}{\sqrt{SEb_1^2 + SEb_2^2}} \tag{7}$$

where b_1 and b_2 are the coefficients obtained from both samples, and SEb_1^2 and SEb_2^2 the corresponding standard errors.

[11] Values in parentheses are standard errors computed following Newey and West (1987). Coefficients and standard errors are in bold if significant at the 95% confidence interval.

[12] Values in parentheses are standard errors computed following Newey and West (1987). Coefficients and standard errors are in bold if significant at the 95% confidence interval.

[13] This methodology is also used by Acharya and Xu (2017) and Xiong et al. (2017).

Table 6
The relationship between *Ab GB*1 and order imbalance.

	*Ind*1	*Ind*2	*Ind*3	*Ind*4	*Ind*5	*Pro*1	*Pro*2	*Pro*3	*Pro*4	*Pro*5	*Pro*6	*OI*
*Ab GB*1	**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	47,760	47,760	47,760	46,685	44,320	40,695	37,674	15,949	44,060	37,080	8817	47,387
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

Table 7
The influence of *Ab GB*1 on individual investors' order imbalance.

	*Ab GB*1					*Ab GB*1 *Short sell*				
	*Ind*1	*Ind*2	*Ind*3	*Ind*4	*Ind*5	*Ind*1	*Ind*2	*Ind*3	*Ind*4	*Ind*5
*Ab GB*1	**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	24,193	24,193	24,193	23,213	21,152	23,567	23,567	23,567	23,472	23,168
R-squared	0.053	0.025	0.017	0.015	0.020	0.106	0.082	0.031	0.015	0.027

Results indicate that the influences of *Ab GB*1 on order imbalance are significantly larger for stocks that can be sold short: *Diff* in Table 7 is calculated as the difference between the right hand and the left hand coefficient. The first *Diff* value is 0.945, which is calculated by the influence of *Ab GB*1 on stocks that can be sold short (2.110) minus the influence of *Ab GB*1 on stocks that can't be sold short (1.165). The value in the parentheses is the significant value of *Diff* (0.945), which is obtained by Eq. (7). In the light that the value of the *Z* statistic is larger than the critical value at the 1% critical level (4.258), we can draw the conclusion that *Ab GB*1 have different influences for stocks sold short and not sold short.[14] This result should be reasonable: in Chinese stock market, individual investors only account for 24% of all short sellers in the period,[15] and the ratio of short sales by individual investors to total short sales is only of about 4%. Short selling doesn't play a great role in investors' trading, and it should be noted that stocks allowing for short selling are usually larger in size and have greater liquidity, increasing the probability to be mentioned on *guba*.

In conclusion, *Ab GB*1 remains a good proxy for individual investors' attention, even after lifting short selling constraints.

5. Conclusion

In this paper we propose a new proxy for investors' attention: daily ab- normal reading quantity of each stock's posts on the *Eastmoney guba* website. we derive multiple results: *Ab GB*1 has a strong correlation with other proxies, namely Ab News1, *Abs ab ret*, and *Ab volume*; furthermore, *Ab GB*1 is efficient in predicting short term stock returns as well as future reversal; and finally, *Ab GB*1 is

[14] Except for individual investors who hold largest assets. Considering their capital size, they might have similar trading behaviour as institutional investors.

[15] Threshold of short selling is 500,000 CNY, which is relatively high for individual investors.

highly correlated with heterogeneous trading behaviour of different types of investors. The proxy is therefore a useful expansion of available investors' attention proxies in the Chinese stock market.

Our study will be of interest to Chinese policy makers especially, as individual investors are the main participants in the Chinese stock market, while still imperfect processors of publicly available information due to their limited analytical abilities.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.frl.2018.11.009.

Reference

Acharya, V.V., Xu, Z., 2017. Financial dependence and innovation: the case of public versus private firms. J. Financ. Econ. 124 (2), 223–243.
Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21 (2), 785–818.
Ben-Rephael, A., Da, Z., Israelsen, R.D., 2017. It depends on where you search: institutional investor attention and underreaction to news. Rev. Financ. Stud. 30 (9), 3009–3047.
Chen, H., De, P., Hu, Y.J., Hwang, B.-H., 2014. Wisdom of crowds: the value of stock opinions transmitted through social media. Rev. Financ. Stud. 27 (5), 1367–1403.
Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. J. Financ. 66 (5), 1461–1499.
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. J. Polit. Econ. 81 (3), 607–636.
Gervais, S., Kaniel, R., Mingelgrin, D.H., 2001. The high volume return premium. J. Financ. 56 (3), 877–919.
Hirshleifer, D., Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. J. Account. Econ. 36 (1- 3), 337–386.
Hood, M., Lesseig, V., 2017. Investor inattention around stock market holiday. Financ. Res. Lett. 23, 217–222.
Hsu, C.-C., Chen, M.-L., 2017. The timing of low-volatility strategy. Financ. Res. Lett. 23, 114–120.
Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Inc, Englewood Cliffs, New Jersey.
Kou, Y., Ye, Q., Zhao, F., Wang, X., 2018. Effects of investor attention on commodity futures markets. Financ. Res. Lett. 25, 190–195.
Lou, D., 2014. Attracting investor attention through advertising. Rev. Financ. Stud. 27 (6), 1797–1829.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
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.
Seasholes, M.S., Wu, G., 2007. Predictable behavior, profits, and attention. J. Empir. Financ. 14 (5), 590–610.
Tetlock, P.C., 2011. All the news that's fit to reprint: do investors react to stale information? Rev. Financ. Stud. 24 (5), 1481–1512.
Tsukioka, Y., Yanagi, J., Takada, T., 2018. Investor sentiment extracted from internet stock message boards and IPO puzzles. Int. Rev. Econ. Financ. 56, 205–217.
Xiong, X., Gao, Y., Feng, X., 2017. Successive short-selling ban lifts and gradual price efficiency: evidence from China. Account. Financ. 57 (5), 1557–1604.
Yang, W., Lin, D., Yi, Z., 2017. Impacts of the mass media effect on investor sentiment. Financ. Res. Lett. 22, 1–4.
Yuan, Y., 2015. Market-wide attention, trading, and stock returns. J. Financ. Econ. 116 (3), 548–564.
Yung, K., Nafar, N., 2017. Investor attention and the expected returns of reits. Int. Rev. Econ. Financ. 48, 423–439.