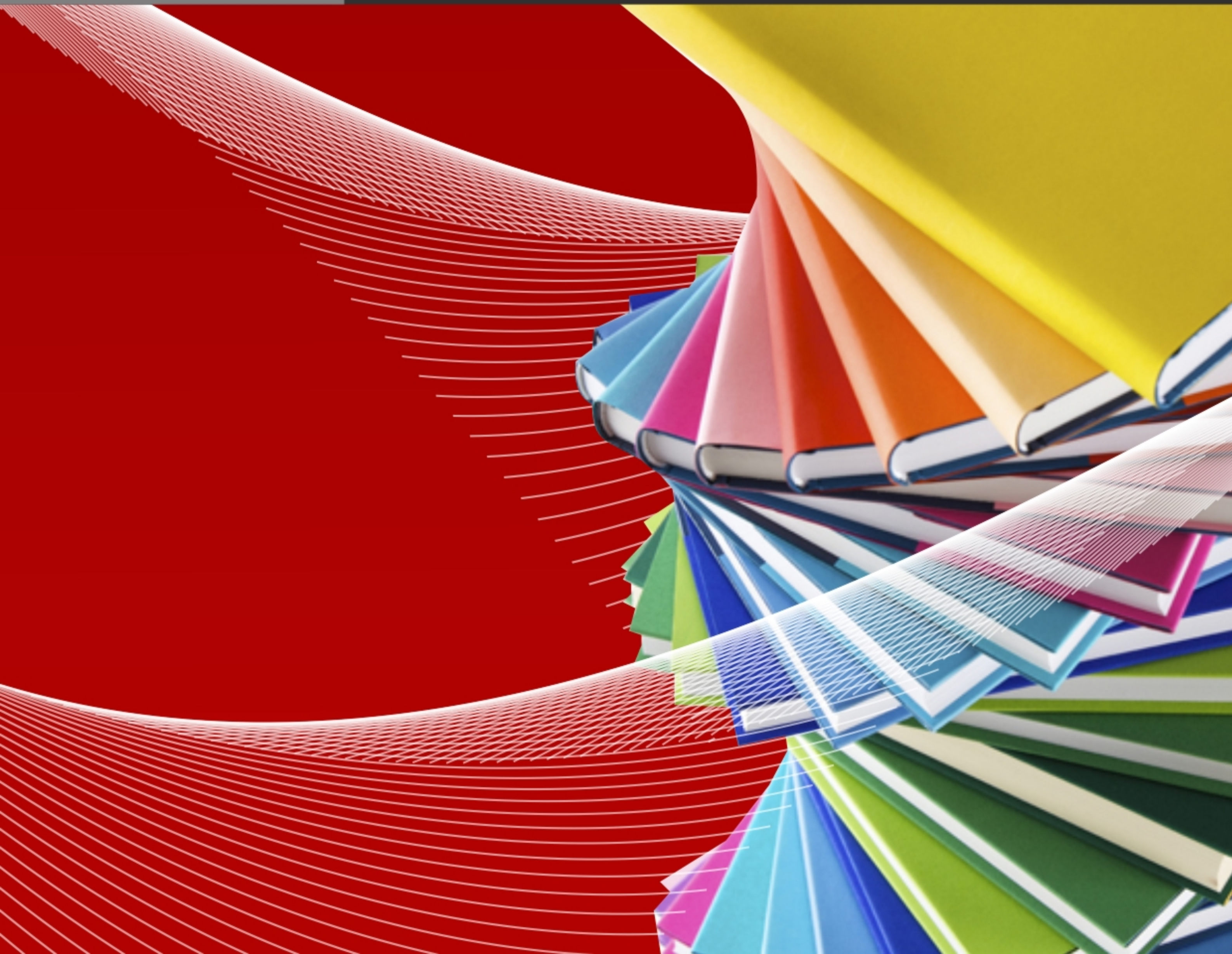
澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Outstanding Academic Papers by Students
學 生 優 秀 作 品



UNIVERSITY OF MACAU

FACULTY OF BUSSINESS ADMINISTRATION

Attention and Retail Investors' Behavior in China

Oscar, LEONG HOK CHON

MB546627

Thesis presented to the

Faculty of Business Administration

University of Macau

In partial fulfillment for granting the

Master of Science in Finance Degree

2017

UNIVERSITY OF MACAU

FACULTY OF BUSSINESS ADMINISTRATION

Academic Thesis

Attention and Retail Investors' Behavior in China

Oscar, LEONG HOK CHON

MB546627

Supervisor: Prof. Jinjuan REN

Abstract

Berkman, Koch, Tuttle, and Zhang (2012) find positive overnight returns and negative intraday returns in the U.S. market, and they attribute such pattern to Barber and Odean (2008)'s limited attention. However, I find negative overnight returns and positive intraday returns in China. While I measure investors' attention using 15-minutes returns at the market open with intraday high-frequency transaction data, I find that high attention leads to overpricing in the rest of the trading day and reversal on the next day. The evidence from China is essentially consistent with the attention-driven intraday return pattern.

I. Introduction

Compared with institutional investors, retail investors tend to be more irrational and uninformed. Barber and Odean (2008) argue that if retail investors want to buy, they would focus on stocks that attract their attention. They find that retail investors are net buyers following both positive and negative return surprises. Consistently, Berkman et al. (2012) measure attention by the square of return yesterday and retail net buying yesterday. They separate the sample into three portfolios every day based on each of the attention measure. Using the square of return yesterday as the attention measure, they find that the close-to-open overnight return of high-attention stocks outperforms low-attention stocks by 11 basis point. The outperformance would reverse by 10 basis point in the intraday period. They argue that the outperformance is caused by the attention-driven behavior of retail investors and followed by intraday reversal. They also find that the retail buying of high attention stocks is greater near the market open. Their results are robust to the control of institution ownership, transaction costs and investor sentiment.

Motived by Berkman et al. (2012), I study the attention-driven behavior of retail investors in China. I examine the overnight (close-to-open) and intraday (open-to-close) return patterns of stocks with high and low attention. Using the square of return yesterday as attention, my evidence shows that in China, high attention stocks tend to have lower overnight return and higher intraday return. This result seems inconsistent with Berkman et al. (2012).

I suggest an alternative explanation that in China, retail investors' attention does not focus on yesterday's historical returns, but on the within trading-day pattern. Accordingly, I utilize the high-frequency intraday data, divide the trading day into 15-minute at open and the rest of the trading day and find that extreme returns in the 15-minute window at market open, which is my new attention measure, attract retail investors' net-purchase later on. In the rest of the trading day, both the retail net-purchase and the return of high attention stocks are significantly greater than those stocks with low attention. The return of high attention stocks on average outperforms by about 38.5 basis point daily, which is about 96% annual. The huge outperformance will reverse on the following night period by about 13 basis point.

The model extended by Barber and Odean (2008) assumes that the return surprise (attention) is generated by institutional investors, since institutions can gather information in the first place. At that time, retail investors should not notice the information yet. That is, the retail buying should not be different from the retail selling at that time. Next, retail investors are attracted by the return surprise so they start purchasing the high attention stocks. This is the processes of attention-driven behavior presented by the model. The attention measure under no difference of retail buying and selling is 'pure', which means that the attention is generated by institutional investors. In this study, in order to best follow the processes, I control the retail net buying before I define stocks as high, medium or low attention. The results after controlling the retail net buying are consistent with the pervious results which do not violate the model. Moreover, the results are robust if I use another division point (2 hour) as the division point or adjust the attention by its median.

My study makes three contributions. Firstly, the Chinese market is the largest emerging market and is growing increasingly important in the world. Empirical evidence from this emerging market helps market participants better understand the financial economy. Secondly, behavioral biases, such as attention, arguably influence the Chinese market by a greater extent than in the U.S. market. This study is expected to add into the large literature in how behavioral biases influence asset pricing. Specifically, I give an explanation for understanding how the attention-driven phenomenon works in China. This is an important consideration for other research studying the Chinese market, especially for intraday analysis. It also provides references for other developing markets. Thirdly, this study has important implication for the investing community. The phenomenon in China implies a market-timing in contrast to that of US. In China, high-attention stocks are normally overpriced at the close of the day. Also, the strategy of buying stocks with low attention and selling stocks with high attention earns a bid-ask-spread-adjusted return of 38.5 basis point per day, translating to about 96% per annual.

II. Literature review

Barber and Odean (2008) extend the theoretical model presented by Kyle (1985) that simulates the attention-driven behavior of retail investors. They hypothesize that if retail investors want to buy, they would focus on stocks which attract their attention. They find that retail investors are net buyers following both positive and negative return surprises. It is consistent with Hirshleifer, Myers, Myers, and Teoh (2008) that earnings surprises also attract retail net buying.

Berkman et al. (2012) extend Barber and Odean (2008) and predict that the attention-driven behavior of retail investors would concentrate near the open of trading days. They divide the close-to-close daily return into two parts, overnight and intraday return, and examine how attention affects the return patterns. They find that the overnight returns of high-attention stocks would be higher and reverse in the intraday return. They also find that the impact on the open prices would be greater with high investor sentiment.

Both Barber and Odean (2008) and Berkman et al. (2012) study the behavior of retail investors in US. US, as a developed market, is dominated by institutional investors. Institutional investors would eliminate the irrational effect of retail investors by arbitrage. As a result, the phenomenon in US is highly rational.

But for China, the phenomenon implied in US may bias. For example, Berkman et al. (2012) result that the average overnight return of stocks with high measure of attention (square of yesterday return) is significantly positive (overpriced at opening). However, in preliminary test, I find that the corresponding return in China is significantly negative (overpriced at closing). This result seems inconsistent with the evidence in US, which is the main question I focus on.

I argue that in China, retail investors' attention focuses on the within trading-day pattern. Extreme returns in the early morning session of the trading day attract retail investors' attention and results in excess retail purchase. Such excess purchase will push up the price until the close of the trading day and reverse on the open of the next trading day. My preliminary evidence based on intraday transaction data supports this explanation.

The explanation is motivated by the manipulation disclosed by the China Securities Regulatory Commission. They investigate and treat quite a few short-term manipulators these years[1]. Short-term manipulators first execute some limit buy order with high prices, usually several minimum spread higher than the current price. Then they continuously input a great limit buy order with low prices (not to execute) and quickly cancel it. They create the appearance that the stock is supported with low price and executed with high price. Retail investors see the appearance as a buying signal and purchase the stock. After the price is pushed up and held on, they can sell their initial great purchase and make profit. Such strategy is available only when the retail investors focus on the within trading-day pattern and highly dominate the market. Otherwise the short-term manipulators using the strategy are not able to make profit.

Researchers use different attention measures to examine the attention-driven behavior. Lou (2014) show that increasing in advertising spending help to attract investor attention and lead to a rise in retail buying and abnormal stock returns. Da, Engelberg, and Gao (2011) conclude that Google Search Volume Index is a direct measure of retail investor attention. Similarly, Ying, Kong, and Luo (2015) use Baidu Search Volume Index as attention measure to study the retail investor behavior in China and find that investor attention has positive effect on the weekly stock returns. The effect is reversed incompletely in the following week. Titman, Wei, and Zhao (2016) find that in China, retail investors are net buyers after stock split announcements while institutional investors are net sellers. The announcements lead to positive return drift for 3 months. The positive return drift is not fully reversed in 12 months also. They conclude that retail investors are attracted by unusual split announcements, which is also a kind of attention. In my best understanding, measuring the attention within a trading day to study the attention-driven behavior of the same day is not yet concluded by researchers.

I study Chinese market in this paper for several reasons. Firstly, Chinese market as the largest emerging market is increasingly important in the world. By the end of 2016, the total market capitalization has exceeded 7.3 trillion US$ (4.1 trillion for Shanghai Stock Exchange and 3.2 trillion for Shenzhen Stock Exchange). China has already exceeded Japan and become the second largest market in the world. Secondly,

[1] One example in 2009 about Chen GuoSheng:
http://www.csrc.gov.cn/pub/newsite/jcj/aqfb/201107/t20110729_198300.html

part of Chinese stock market is open to foreign investors through Hong Kong2. Besides, the State Council of China recently messages that foreign invested enterprises are allowed to finance by listing in Chinese stock market[3]. Thus, it is important to examine whether the phenomenon in US exists in China. Finally, Chinese market is dominated by retail investors. The investor structure of Chinese market is different from US. The result in China provides a more suitable reference for other developing markets.

III. Research questions and methodology

3.1 Testable hypotheses

Compared with developed markets such as US, Chinese market has lower efficiency as Chinese market is dominated by retail investors. The Securities Association of China publishes a report about the investor structure in 2016[4]. The report shows that over 99.5% of investor accounts in China belongs to retail investors. Retail investors have around three times of influence comparing to institutional investors over the period of 2006-2014, in terms of shareholding. For example, in 2014, retail holding and institutional holding account for 37% and 10% of the circulation market value respectively. The arbitrage ability of institutional investors is quite low in China.

As a result, there should be attention-driven behavior of retail investors in China. That is, the implication presented by Berkman et al. (2012) should be obvious in China:

Stocks with high return surprises would attract retail investors' attention. Retail investors, which can only focus on stocks that attract their attention, release their buying power. As a result, the retail net buying should be higher.

<u>Hypothesis 1: other things being equal, stocks with higher attention have greater retail net buying.</u>

[2] Foreigners can set up an accounts in Hong Kong and trade some A-shares through Shanghai-Hong Kong Stock Connect (2014) and Shenzhen-Hong Kong Stock Connect (2016)
[3] http://www.gov.cn/zhengce/content/2017-01/17/content_5160624.htm
[4] http://www.sac.net.cn/yjcbw/zgzqzz/2016/2016_06/201607/P020160729391425134268.pdf

Since the Chinese market is highly dominated by retail investors, the high attention stocks that have relative high retail net buying should have higher return contemporaneously.

Hypothesis 2: other things being equal, stocks with higher attention have higher returns.

Berkman et al. (2012) find a high level of reversal following the temporary overpricing revealed by the attention-driven behavior. Titman et al. (2016) and Ying et al. (2015) also document that in China, the temporary overpricing incompletely reverse in the next period. I suggest that the attention-driven short-term high returns reverse following the temporary overpricing.

Hypothesis 3: other things being equal, stocks with higher attention have greater returns reversal.

3.2 Methodology of preliminary results

Similar to Berkman et al. (2012), I calculate the returns using quotations from CSMAR High-frequency Database over the period 2010–2015. The data provide the last trading price, quotes, and current volume for each A-shares every 5-6 seconds. However, the identity information is not provided. As a result, to access the retail trading, if the average trade value each interval is less than RMB 10,000, I assume the trading is made by retail investors (Lee (1992), Ng and Wu (2007), Wongchoti, Wu, and Young (2009) and Barber, Odean, and Zhu (2009)). The data also provide the buy-sell indicator whether the trading is initiated by the buyer or seller. If the trading is classified as a retail trading and initiated by the buyer (seller), I assume the trading is retail buying (selling). The data only provide the number of trade for Shenzhen Stock Exchange so the variables related to retail trading are only for Shenzhen market.

In order to avoid the bid-ask bounce, the prices are defined as the mid quotes of the highest bid and the lowest ask (Berkman et al., 2012). The opening prices are the first effective[5] mid quotes after 9:30[6]. The closing prices are the last effective mid

[5] "Effective" means that: 1. there is at least one pair of orders are matched and executed; 2. the lowest ask is smaller than 1.5 times the highest bid.

[6] In China, the stock market is open to trade at 9:30 and close at 15:00, with a 1.5 hours lunch break. There is a call auction 15 minutes before the market open. After the market is open, the unexecuted orders in call auction are still available to execute.

quotes before 15:00. After adjusting the daily opening and closing prices for stock splits and dividends, I measure the returns as following:

$$\mathrm{CTO}_t = \log(\mathrm{open}_t / \mathrm{close}_{t-1}) \qquad (1)$$

$$\mathrm{OTC}_t = \log(\mathrm{close}_t / \mathrm{open}_t) \qquad (2)$$

$$\mathrm{CTC}_t = \mathrm{CTO}_t + \mathrm{OTC}_t \qquad (3)$$

After defining the retail buying and selling, I compute 3 measures of retail buying near the open (Berkman et al., 2012):

$$\text{Retail Buy\_1}^{st} = 1 \text{ if the first trade is purchase by Retail investors or otherwise 0, minus 0.5.} \quad (4)$$

$$\text{Retail NetBuy\_15min} = \frac{\text{Retail Net Buy Volume in first 15 minutes}}{\text{Retail Volume in first 15 minutes}} \qquad (5)$$

$$\text{Retail NetBuy\_Diff\_1h} = \frac{\text{1st hour Retail Net Buy Volume} - \text{average Retail Net Buy Volume in the rest of the day}}{\text{Share outstanding}} \qquad (6)$$

To access the attention of each stock, Berkman et al. (2012) use yesterday square return (*SQ_return*) and yesterday retail net buying (*Retail_NetBuy*) to proxy attention. They conclude that both of the attention measures help to explain the attention-driven behavior of retail investors. Therefore, I use the same method as Berkman et al. (2012) to proxy the attention:

$$\text{SQ\_return}_t = \mathrm{CTC}_t^{\,2} \qquad (7)$$

$$\text{Retail\_NetBuy}_t = \frac{\text{Retail Buy Volume}_t - \text{Retail Sell Volume}_t}{\text{Total Retail Volume}_t} \qquad (8)$$

The sorting strategy is that, firstly, for each trading day, I divide all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (*SQ_return* or *Retail_NetBuy*) into three portfolios, namely, Low, Medium and High. I combine the three size-grouped Low portfolios into one larger Low portfolio (similar for Medium and High). As a result, the three portfolios are size-adjusted. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference of each variables between High and Low portfolio.

3.3 The stock exchanges in China

There are two stock exchanges in China, Shanghai Stock Exchange (SH market) and Shenzhen Stock Exchange (SZ market). To access the retail trading through the high-frequency data, if the average trade value each interval is less than RMB 10,000, I assume the trading is made by retail investors. The data only provide the number of trade for SZ market so the variables related to retail trading are only for SZ market. The evidence in SZ market about the attention-driven behavior of retail investors may not be valid for SH market. Thus, it is important to study the difference between the two markets.

Table I shows the size-adjusted return patterns under different attentions. The three columns in Panel A of Table I show CTO, OTC, CTC on the whole market. For all the attention portfolios, CTO is significantly negative, OTC is significantly positive, and CTC is not significant. Also, the difference of CTO between High and Low portfolio is significantly negative, and the difference of OTC between High and Low portfolio is significantly positive. The columns in Panel B & C show similar results as Panel A, except the difference of OTC in Panel C is not significant. Moreover, Panel D shows the return-difference between the two markets. Most of the return-differences are not significant except CTO_{diff} in Low and Medium portfolio. Most importantly, for the stocks with high attention, all the three measures of the return-difference are not significant. Overall, I conclude that the result for the attention-driven behavior of retail investors in China is not biased by markets.

3.4 Preliminary results

I do a preliminary test about the hypotheses. In Table II, for both of the panels, the differences of *Retail NetBuy_15min* and *Retail NetBuy_Diff_1h* are significantly positive. However, the differences of *Retail NB_1st* are significantly negative. That is, Hypothesis 1 is not supported. High-attention stocks attract retail buying only near the open, but not at the open. Meanwhile, as the differences of *CTO* are significantly negative, Hypothesis 2 is also not supported. Similarly, given that the differences of *OTC* are significantly positive, Hypothesis 3 is not supported either. Thus, a deeper analysis is needed.

3.5 Alternative attention measures

Table II reports a result which is quite different from previous literatures.

Barber and Odean (2008) as well as Berkman et al. (2012) test their hypothesis by defining the attention-driven event on one day and then study the reaction of retail investors in the next day (in total two days). The reason they study on daily-basis lies in that they want to avoid endogeneity problems- whether return surprises attract buyers or buyers bring surprises. However, the model proposed by Barber and Odean (2008) do not limit the attention-driven event on daily-basis.

How about the processes have already finished within a day? One possible explanation is that in the Chinese market, when there is an attention appearing on a stock within a trading day, retail investors purchase the stock immediately and overreact on the rest of the trading day. As I mentioned above, institutional investors may not have enough power to do arbitrage in such stocks. Thus, at the end of the trading day, the attention-driven stocks would achieve a high close price. After that, the enthusiasm of retail investors cools down during the night period. As a result, the open of the attention-driven stocks next day will revise downward, because of arbitraging by institutional investors and regretting by retail investors.

Also, the explanation does not violate the model presented by Barber and Odean (2008). The model is constructed under a rational understanding. When there is news on a stock, no matter good or bad, institutional investors will collect the news first and then trade in the market. The price of the stock would thus face a rise or decline. This price movement attracts retail investors' attention. The retail investors who are looking for buying will release their buy power and push up the price.

In other words, I suggest that in China, retail investors' attention focuses on the within trading-day pattern. Extreme returns in the early morning session of the trading day attract retail investors' attention and result in excess retail purchase. Also, as the Chinese market is highly dominated by retail investors, the expected function of institutional investors is low in China. Such excess purchase will push up the price until the close of the trading day and reverse on the open of the next trading day. My result in Table I based on intraday transaction data supports this explanation.

Barber and Odean (2008) divide the happening of two days into two part, before and after, by the close of the first day. While I want to reduce the happening into one day, I have to select an intraday division point. Thus, the studying period, from yesterday close to today close, is divided into two part, before and after, by the intraday division point. As a result, I can measure the attention before the intraday division point and study the following impact. This methodology is similar to Barber and Odean (2008) and Berkman et al. (2012):

	First period: Attention identification	Second period: Attention-driven impact	Third period: Reversal
Barber and Odean (2008)	Day t	Day t+1	/
Berkman et al. (2012)	Day t	Overnight period following day t	Day t+1
This paper	The foregoing part of day t	The rest of day t	Overnight period following day t

In practice, to construct the intraday attention, I do the following. Firstly, I select one time point as a division point for every trading days. Secondly, I calculate the return before the division point (*Return_Before*). Thus the square of *Return_Before* (*SQ_Return_Before*) becomes my proxy of intraday attention.

$$\text{Return\_Before}_t = \log(\text{price at division point}_t / \text{close}_{t-1}) \quad (9)$$

$$\text{SQ\_Return\_Before}_t = \text{Return\_Before}_t^{\,2} \quad (10)$$

After the construction of the intraday attention, I examine the rest of the trading day. I focus on the retail net buying after the division point (*NetBuy After*) and the return after the division point (*Return_After*). For reference purpose, I also compute the retail net buying before the division point (*NetBuy Before*), the overnight return next day (*CTO next*) and the intraday return next day (*OTC next*).

$$\text{NetBuy After} = \frac{\text{Retail Net Buy Volume After the division point}}{\text{Total Retail Volume After the division point}} \quad (11)$$

$$\text{Return\_After}_t = \log(\text{close}_t / \text{price at division point}_t) \quad (12)$$

$$\text{NetBuy Before} = \frac{\text{Retail Net Buy Volume Before the division point}}{\text{Total Retail Volume Before the division point}} \quad (13)$$

$$\text{CTO next} = \text{CTO}_{t+1} = \log(\text{open}_{t+1} / \text{close}_t) \quad (14)$$

$$\text{OTCnext} = \text{OTC}_{t+1} = \log(\text{close}_{t+1} / \text{open}_{t+1}) \qquad (15)$$

Here I suggest using 15 minutes after the market open as the division point. The model presented by Barber and Odean (2008) suggests that the attention is inferred by the trading of institution. Generally, institution obtains the first-hand information. If the information is released on the non-trading period[7], the first 15 minutes becomes the first opportunity to trade for the new information. As the stock market is closed at 15:00 in China, new information is more likely to be released on the non-trading period. The attention should be more likely to appear during the first 15 minutes. Also, if the result of using 15-minute division point really account for some implication, selecting a further division point should reflect a stronger result. It is because 15 minutes may not enough for institution to trade for the information and as a result infer an attention and retail investors are more likely to know that there is an attention. That is, the result of using further division points should enhance the implication.

The sorting strategy is same as *3.2*, except that the attention proxy is the square of the return before the 15-minute division point (*SQ_Return_Before*).

IV. Empirical result

4.1 The effect of the intraday attention on the rest of the day

If retail investors are attracted by the intraday attention and purchase the stocks with high attention, the retail net buying of the stocks with high attention during the rest of the trading day should be greater than the stocks with low attention. As retail investors dominate the Chinese market, such greater net purchase may push up the prices, resulting a greater return on the rest of the trading day. After one night of cool down period, the prices should adjust downward.

To test the three hypotheses, I treat each day as a separate event. I initially divide all stocks into three groups based on size. Within each size groups, I sort the stocks by the intraday attention into three portfolios, namely, Low, Medium and High. I combine the three size-grouped Low portfolios into one larger Low portfolio (similar for Medium and High). Then, within each size-adjusted portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series

[7] Except the 1.5 hours lunch break.

mean of each variables and the difference of each variables between High and Low portfolio. The cut points of the three group separation are the 30^{th} percentile and the 70^{th} percentile. All the significances are based on the Newey and West (1986) adjusted standard errors of the time-series means.

Table III reports the time-series means of the variables each portfolios. The column of *NetBuy After* brings the result of Hypothesis 1. Although *NetBuy After* is significantly negative within each portfolio, the difference of *NetBuy After* is 0.016 and significantly positive with 15.99 of t value. This evidence supports Hypothesis 1. The column of *Return_After* brings the result of Hypothesis 2. *Return_After*, which is the return of the rest of the trading day, is significantly positive within each portfolio. The difference of *Return_After* is 0.385 and significantly positive. This evidence supports Hypothesis 2. Also, the difference of 0.385 means that the return difference between stocks with high and low attention on average is 38.5 basis point daily, which can be transformed into about 96% yearly (almost a double). The column of *CTO next* brings the result of Hypothesis 3. Higher attention result in higher prices/returns. Such relation revise on the cool down period, which is the overnight period. Table III shows that the difference of *CTO next* is -0.130 and significantly negative, which is around 33% of the difference of *Return_After*. That is, around 33% of the influence of the attention-driven behavior are revised during the following overnight period. This evidence supports Hypothesis 3. Moreover, the difference of *OTC next* is significantly positive, which means that the impact of high attention also affect the returns on the next trading day. Overall, the results support all the hypotheses. Using 15 minute after the market open as an intraday division point help imply the attention-driven behavior in China.

Table III show that only 33% of the influence of the attention-driven returns are reversed during the following overnight period. However, Berkman et al. (2012) find that the attention-driven impact in US is almost completely reversed. One of the reason is that in China, the arbitrage ability of institutional investors is low. Retail ownership is around three times of institutional ownership in the Chinese market, as reported by The Securities Association of China in 2016[8]. The overpricing may not be able to be adjusted completely. Another reason refers to the study of Information

[8] http://www.sac.net.cn/yjcbw/zgzqzz/2016/2016_06/201607/P020160729391425134268.pdf

Uncertainty (IU). Jiang, Lee, and Zhang (2005) use firm age, return volatility, trading volume and duration to measure IU and find that stocks with low IU tend to have higher future return. My measure of intraday attention, which is based on the model extended by Barber and Odean (2008), might also proxy for the trading of institutional investors on new information. The revealed new information lowers the level of IU so that the following return is higher. In this sense, the difference of *Return_After* should be higher also. However, this part of the return increment should not be revised in the following period. As a result, the percentage of reversal is not high. This is also the limitation of my intraday attention measure.

Meanwhile, Table III also shows that the difference of *NetBuy Before* is significantly positive. This evidence implies one concern that the measure of the intraday attention is already the outcome of the attention-driven behavior, since I measure the intraday attention by the square of returns. A high intraday attention measure may actually come from yesterday's attentions. The result is week and inappropriate for the hypotheses because the causality is not clearly defined, just like a Mathematical Induction proof without the basis step of n=0. Given this, I also report the return during the intraday attention identification period (*Return_Before*) in Table III. The difference of *Return_Before* is significantly negative. That is, even though the retail net buying of the high attention stocks is greater, the returns on the same period is lower. The retail buying power may be recovered by the institution. This evidence contradict with the concern in the sense that the intraday attention is not the outcome of the attention-driven behavior. In order to best addressing the concern, in the next part, I control the retail net buying initially.

4.2 Net-buying-controlled attention-driven behavior

In the first place, the return surprises should be generated by the trading of institutional investors because institution gather the information firstly. At that time, retail investors should not have any reaction. This is the starting point of the attention-driven behavior. As a result, controlling the retail net buying initially can best address the concern mentioned above. That is, if the intraday attention can be defined without the influence of the retail net buying, such intraday attention is "pure", which means that the result under the intraday attention do not rise the concern. The smaller the difference of the retail net buying between the High and Low portfolio, the more pure the measure of the attention.

In specific, I equally divide all the stocks into five NetBuy quintiles every day, based on the absolute retail net buying before the division point. Then within each NetBuy quintiles, I do the same sorting as the above.

Table IV reports the net-buying-controlled result. Overall, the finding in Table IV is similar to Table III. For all NetBuy quintiles, the differences of *NetBuy After* are significantly positive, the differences of *Return_After* are significantly positive and the differences of *CTO next* are significantly negative. The evidences support the hypotheses. Most importantly, within the Lowest NetBuy tertiles, the difference of *NetBuy Before* is almost zero and insignificant. That is, within the Lowest NetBuy quintile, the intraday attention measure is pure that it do not affected by the retail net buying. Based on that, the results still imply the validness of the hypotheses, which means that the concern does not affect the implication.

Moreover, NetBuy quintiles, which are based on the absolute retail net buying before the division point, represent the degree of the variation of the retail investors' trading. The greater the retail net buying and selling, the harder institution investors executing arbitrage. As a result, NetBuy quintiles also represent the degree of difficulty of institution investors executing arbitrage. In Table IV, the differences of *NetBuy After* do not deviate a lot across different NetBuy quintiles. It means that the attention-driven behavior of retail investors is similar within different NetBuy quintiles. On the other hand, the difference of *Return_After* is greater within higher NetBuy quintiles. That is, even though the impact of the attention-driven behavior is similar, the return after the division point is greater in higher NetBuy quintile. Similarly, the reversal is smaller in higher NetBuy quintile, given that the difference of *CTO next* is negative and greater in higher NetBuy quintile. Also, within the Lowest NetBuy quintile, the reversal (the difference of *CTO next*) is around 54% of the difference of *Return_After*, which is the greatest of all the NetBuy quintiles and greater than the percentage in Table III. The percentage for the Highest NetBuy quintile is around 10% only, which is smaller than the percentage in Table III. Overall, the implication is consistent with the understanding of the participation of institution investors. The 54% reversal, which is not high, also refers to the limitation of my intraday attention stated in above.

4.3 Alternative intraday division point: 2h

Using 15 minutes after the market open as an intraday division point imply another concern, whether 15 minutes is appropriate. The attention appears on the first 15 minutes but retail investors may not notice what is happening. Rather than looking for new stocks, during the trading period, retail investors may only focus on the stocks they owned. Their attention on the stocks they owned relaxes during break times. In China, there is a 1.5 hours lunch break from 11:30 to 13:00. The lunch break provides retail investors with a chance to look at other stocks. As a result, retail investors are more likely to focus and trade high attention stocks after the lunch break. The lunch break starts 2 hours after the market open. In this part, I use 2 hours (2h) after the market open as the intraday division point to retest the hypotheses.

Table V & VI report the results using 2h after the market open as the intraday division point. Overall, the results are similar to and consistent with the results in Table III & IV.

In Table V, the difference of *NetBuy After* is 0.020 and significantly positive, which supports Hypothesis 1. It is slightly larger than the number (0.016) in Table III. Meanwhile, the difference of *Return_After* is 1.167 and significantly positive, which supports Hypothesis 2 and is around 3 times of the number (0.385) in Table III. It suggests that over time retail investors are more likely to focus and trade high attention stocks, leading a great return in the following period. Also, *NetBuy After* represents the proportion of the retail net buying to total retail trading and does not reflect the volume of the retail trading. The slightly greater *NetBuy After* is not inconsistent with the great *Return_After*. The difference of *CTO next* is -0.127 and significantly negative, which supports Hypothesis 3.

Table VI reports the net buying controlled results. All the NetBuy quintiles have similar results- the differences of *NetBuy After* is significantly positive, the differences of *Return_After* is significantly positive and the differences of *CTO next* is significantly negative. Within the Lowest NetBuy quintiles, although the differences of *NetBuy Before* is significantly negative, the magnitude is only 0.002 or 0.2%. The effect of *NetBuy Before* on the intraday attention is highly controlled. Also, the higher NetBuy quintiles, the greater the difference of *Return_After*. It is similar with Table IV and implies that the degree of difficulty of institution investors executing arbitrage is higher when the variation of the retail net buying increase. The reversal is about 22% of the difference of *Return_After*, which is lower than the number (33%) in Table III. It also suggests that using 2h as the intraday division point implies higher retail investors' participation.

4.4 Median-adjusted intraday attention

Both Table III & V report that the difference of *Return_Before* is significantly negative. The trend of *Return_Before* from Low portfolio to High portfolio is opposite to the trend of *Return_After*. Since the sum of *Return_Before* and *Return_After* is the intraday daily return, the implication about *Return_After* may be related to *Return_Before* also.

Given that *Return_Before* of all the three portfolios in Table III & V are significantly negative, it is possible that in China, the returns in the early section of the trading day is generally negative. This may affect the intraday attention since I measure the intraday attention by the square of the return. If most of stocks have negative returns early the day, the high attention stock portfolio will contain a high proportion of loser stocks. However, in this case the winner stocks may be more likely to attract the retail investors. Using a simple measure of intraday attention may bias the results and result in the trend of *Return_Before*. As a result, in this part, I adjust the intraday attention by its median. The median-adjusted intraday attention is calculated by:

$$\text{Median-adjusted intraday attention}_t = (\text{Return\_Before}_t - \text{median of Return\_Before}_t)^2 \quad (16)$$

Using the median-adjusted intraday attention ensures the High portfolio containing half winner stocks and half loser stocks[9] (similar to the Low and Medium portfolio). Therefore the difference of *Return_Before* should be small. The effect of the *Return_Before* on the implication should be best controlled.

Table VII is consistent with the pervious finding that the hypotheses are supported. In addition, the difference of *Return_Before* is insignificant. That means even *Return_Before* has no difference, the attention-driven behavior are observed.

[9] Winner stocks and loser stocks are defined comparing to the market. That is, the winner (loser) stocks are the stocks with highest (lowest) return.

The difference of *CTO next* is -0.155 and around 53% of the difference of *Return_After*, which means a 53% of reversal.

4.5 Robustness of the attention-driven behavior using Fama-Macbeth regression

In pervious parts, I test the hypotheses by dividing the sample into portfolios with different controls. Here I use Fama-Macbeth regression to retest the intraday phenomenon. If both of the approach imply same implication, the hypotheses are strongly supported since the results are less likely to be produced by construction.

Here are the regressions to be tested:

$$\text{Dependent variables} = \beta_0 + \beta_1 \text{Attention} + \text{controls} \quad (17)$$

Each hypothesis is tested using the corresponding dependent variables. The attention variable is *SQ_Return_Before*. All of the variables are defined same as before:

For Hypothesis 1: $$\text{NetBuy Before} = \beta_0 + \beta_1 \text{Attention} + \text{controls} \quad (18)$$

For Hypothesis 2: $$\text{Return\_After} = \beta_0 + \beta_1 \text{Attention} + \text{controls} \quad (19)$$

For Hypothesis 3: $$\text{CTO next} = \beta_0 + \beta_1 \text{Attention} + \text{controls} \quad (20)$$

The control variables include *MVlog*, *BMlog*, *MOM(-1,0)* and *MOM(-5,-1)*. *MVlog* is the log of the average capitalization of the previous 22 trading day. *BMlog* is the log of the book-to-market ratio. *MOM(i, j)* is the average past close-to-close return from day *i* to day *j-1*. The accounting data is from CSMAR databases. The return measure for day t is CTC_t, which is defined in Equation (3). The four control variables are aim to control the size effect, the book-to-market ratio effect and the momentum effect.

Table VIII, IX & X report the estimation result of Equation (18), (19) & (20) respectively. Each table contains three models that include different control variables. The last models include all the control variables. *SQ_return_Before* is calculated using the time 15 minutes after the market open as an intraday division point. Table VIII shows that all of the coefficient of *SQ_return_Before* are significantly positive. Hypothesis 1 is supported. The coefficients of *SQ_return_Before* are almost same among all the models, meaning that the size effect, the book-to-market ratio effect and the momentum effect do not affect the relation between the attention and the retail net buying. Table IX shows that all of the coefficient of *SQ_return_Before* are significantly positive. Hypothesis 2 is supported. Similar to Table VIII, including the control variables do not affect the relation a lot. Table X shows that all of the coefficient of *SQ_return_Before* are significantly negative is supporting Hypothesis 3. Once again, including the control variables do not affect the relation a lot. Overall, the three tables which use the time 15 minutes after the market open as an intraday division point support the hypotheses. This implication is not affected by the size effect, the book-to-market ratio effect and the momentum effect.

SQ_return_Before measuring the attention on the square of returns may deviate across different day, e.g. the highest return on day t is 3% but 10% on day t+1. Both the highest returns should be considered as same level of attention. As a result, I transpose *SQ_Return_Before* to *Score_Attention* which is a daily ranking score adjusted to range between 0 and 1 and reestimate Equation (18), (19) & (20):

$$\text{Score\_Attention} = (\text{Rank of SQ\_Return\_Before} - 1) / (\text{Number of stocks} - 1) \qquad (21)$$

Table XI reports the estimation of Equation (18), (19) & (20) using *Score_Attention* as the attention variable. Model 1, 2 & 3 are for Equation (18), (19) & (20) respectively. All the models include all the control variables. Similarly, the hypotheses are supported, given that the coefficient of *Score_Attention* is significantly positive in model 1, significantly positive in model 2 and significantly negative in model 3. Also, using ranking score gives some meaning on the value of the coefficients. The coefficient of *Score_Attention* in model 2 is 0.2559, meaning that the return of the stock with the highest attention on the rest of the trading day is 25 basis points over the stock with the lowest attention. The coefficient of *Score_Attention* in model 3 is -0.1503, which is about 59% reversal. The reversal is not high and also suggests the limitation of my intraday attention stated in above.

Moreover, using 2h after the market open as the intraday division point is also appropriate. Table XII & XIII report the estimation of Equation (18), (19) & (20) using 2h after the market open as the intraday division point. The attention variable in Table XII & XIII is *SQ_Return_Before* and *Score_Attention* respectively. Both of the table have similar results with the pervious tables and support the hypotheses. The reversal is about 11% only, which is because using 2h as the intraday division point imply higher retail investors' participation, comparing to using 15min as the intraday division point.

V. Conclusion

This paper finds that in China, retail investors' attention focuses on the within trading-day pattern. The result from China using the methodology of Berkman et al. (2012) is inconsistent with the phenomenon of the attention-driven behavior. While I measure investors' attention using 15-minutes returns at the market open with intraday high-frequency transaction data, I find that high attention leads to retail net buying and overpricing in the rest of the trading day and reversal on the next day. The evidence is consistent when I measure the attention using 2 hours returns or adjust the attention by its median. The evidence from China is essentially consistent with the attention-driven intraday return pattern. Moreover, this paper reveals the possibility that the absolute retail net buying, which represents the degree of the variation of the retail investors' trading, may be related to the degree of difficulty of institution investors executing arbitrage.

This paper reveals the different trading background in China, which is the largest emerging market. The phenomenon of the attention-driven behavior in US has a different form in China. It is an important consideration for other research studying the Chinese market, especially for intraday analysis. The phenomenon in China implies a market-timing in contrast to that of US. In China, high-attention stocks are normally overpriced at the close of the day.

References

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. doi:10.1093/rfs/hhm079

Barber, B. M., Odean, T., & Zhu, N. (2009). Do Retail Trades Move Markets? *The Review of Financial Studies*, *22*(1), 151-186. doi:10.1093/rfs/hhn035

Berkman, H., Koch, P. D., Tuttle, L., & Zhang, Y. J. (2012). Paying Attention: Overnight Returns and the Hidden Cost of Buying at the Open. *Journal of Financial and Quantitative Analysis*, *47*(04), 715-741.

Da, Z. H. I., Engelberg, J., & Gao, P. (2011). In Search of Attention. *The Journal of Finance, 66*(5), 1461-1499. doi:10.1111/j.1540-6261.2011.01679.x

Hirshleifer, D. A., Myers, J. N., Myers, L. A., & Teoh, S. H. (2008). Do Individual Investors Cause Post-Earnings Announcement Drift? Direct Evidence from Personal Trades. *The Accounting Review*, *83*(6), 1521-1550. doi:10.2308/accr.2008.83.6.1521

Jiang, G., Lee, C. M. C., & Zhang, Y. (2005). Information Uncertainty and Expected Returns. *Review of Accounting Studies*, *10*(2), 185-221. doi:10.1007/s11142-005-1528-2

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, *53*(6), 1315-1335. doi:10.2307/1913210

Lee, C. M. C. (1992). Earnings news and small traders. *Journal of Accounting and Economics*, *15*(2), 265-302. doi:http://dx.doi.org/10.1016/0165-4101(92)90021-S

Lou, D. (2014). Attracting Investor Attention through Advertising. *The Review of Financial Studies*, *27*(6), 1797-1829. doi:10.1093/rfs/hhu019

Newey, W. K., & West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix: National Bureau of Economic Research Cambridge, Mass., USA.

Ng, L., & Wu, F. (2007). The trading behavior of institutions and individuals in Chinese equity markets. *Journal of Banking & Finance*, *31*(9), 2695-2710. doi:http://doi.org/10.1016/j.jbankfin.2006.10.029

Titman, S., Wei, C., & Zhao, B. (2016). Are Chinese Investors Really Different? Evidence from Stock Splits in the US And China.

Wongchoti, U., Wu, F., & Young, M. (2009). Buy and sell dynamics following high market returns: Evidence from China. *International Review of Financial Analysis*, *18*(1–2), 12-20. doi:http://doi.org/10.1016/j.irfa.2009.01.003

Ying, Q., Kong, D., & Luo, D. (2015). Investor Attention, Institutional Ownership, and Stock Return: Empirical Evidence from China. *Emerging Markets Finance and Trade, 51*(3), 672-685. doi:10.1080/1540496X.2015.1046339

Appendix

Table I

Difference of the exchanges in China

There are two stock exchanges in China, Shanghai Stock Exchange (SH market) and Shenzhen Stock Exchange (SZ market). Similar to Berkman et al. (2012), I calculate the returns using the data from CSMAR High-frequency Database over the period 2010–2015. CTO, OTC, CTC are defined as same as Berkman et al. (2012), respectively (All the open prices and close prices are mid quotes):

$$CTO_t = \log(open_t / close_{t-1}) \qquad OTC_t = \log(close_t / open_t) \qquad CTC_t = CTO_t + OTC_t$$

The four panels are regarding the whole sample, SZ market, SH market and the difference between SZ and SH market, respectively.

The panels are constructed by first dividing all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (SQ_return or Retail_NetBuy) into three portfolios, namely, Low, Medium and High. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference between High and Low portfolio in Panel A, B & C. I also compute the difference of each variables in each group between SZ and SH market and report the time-series mean of the difference in panel D. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significances and the t-values are based on the Newey and West (1986) adjusted standard errors of the time-series means.

	Panel A: Whole market			Panel B: SZ market		
Attention	CTO	OTC	CTC	CTO	OTC	CTC
low	-0.173***	0.211***	0.035	-0.168***	0.206***	0.038
Medium	-0.159***	0.206***	0.047	-0.15***	0.204***	0.048
High	-0.246***	0.244***	-0.004	-0.242***	0.245***	0.001
Diff (H-L)	-0.073***	0.033**	-0.039**	-0.082***	0.039***	-0.043**
t (Diff)	-7.31	2.26	-2.17	-7.94	2.73	-2.43
	Panel C: SH market			Panel D: Difference between SZ & SH		
Attention	CTO	OTC	CTC	CTO	OTC	CTC
low	-0.184***	0.219***	0.03	0.021***	-0.013	0.012
Medium	-0.171***	0.21***	0.04	0.018***	-0.005	0.011
High	-0.257***	0.238***	-0.02	0.006	0.007	0.015
Diff (H-L)	-0.067***	0.019	-0.046**	-0.015***	0.020*	0.002
t (Diff)	-7.2	1.20	-2.42	-2.78	1.89	0.21

Table II

Attention-driven Retail Net Buying at the open, and the effect on overnight and intraday

Following Berkman et al. (2012), I use yesterday square return (**SQ_return**) and yesterday retail net buying (**Retail_NetBuy**) to proxy attention.

$$SQ\_return_t = CTC_t^{\,2} \qquad Retail\_NetBuy_t = \frac{\text{Retail Buy Volume}_t - \text{Retail Sell Volume}_t}{\text{Total Retail Volume}_t}$$

All the variables are defined as same as Berkman et al. (2012), respectively:

$$\text{Retail Buy\_1}^{st} = 1 \text{ if the first trade is purchase by Retail investors or otherwise } 0, \text{ minus } 0.5.$$

$$\text{Retail NetBuy\_15min} = \frac{\text{Retail Net Buy Volume in first 15 minutes}}{\text{Retail Volume in first 15 minutes}}$$

$$\text{Retail NetBuy\_Diff\_1h} = \frac{\text{Retail Net Buy Volume in 1st hour} - \text{average Retail Net Buy Volume in the rest of the day}}{\text{Share outstanding}}$$

$$CTO_t = \log(open_t / close_{t-1}) \qquad OTC_t = \log(close_t / open_t) \qquad CTC_t = CTO_t + OTC_t$$

The panels are constructed by first dividing all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (**SQ_return** or **Retail_NetBuy**) into three portfolios, namely, Low, Medium and High. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference between High and Low portfolio. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significance are based on the Newey and West (1986) adjusted standard errors of the time-series means.

Panel A: Use yesterday ***SQ_Return*** as attention proxy

Attention	Retail Buy_1st	Retail NetBuy_15min	Retail NetBuy_Diff_1h	CTO	OTC	CTC
low	-0.071***	-0.009**	-0.005**	-0.168***	0.206***	0.038
Medium	-0.069***	-0.003	-0.001	-0.15***	0.204***	0.048
High	-0.131***	0.004	0.005	-0.242***	0.245***	0.001
Diff (H-L)	-0.060***	0.013***	0.010***	-0.082***	0.039***	-0.043**
t (Diff)	-28	11.3	5.6	-7.94	2.73	-2.43

Panel B: Use yesterday ***Retail_NetBuy*** as attention proxy

Attention	Retail Buy_1st	Retail NetBuy_15min	Retail NetBuy_Diff_1h	CTO	OTC	CTC
low	-0.066***	-0.025***	-0.011***	-0.151***	0.189***	0.038
Medium	-0.093***	-0.003	-0.001	-0.19***	0.223***	0.033
High	-0.105***	0.019***	0.009***	-0.226***	0.253***	0.027
Diff (H-L)	-0.039***	0.044***	0.020***	-0.075***	0.064***	-0.011
t (Diff)	-12.12	32.2	11.09	-12.18	4.95	-0.94

Table III

The effect of the intraday attention on the rest of the day: 15 minutes as the intraday division point

Using the time **15 minutes** after the market open as an intraday division point, I calculate the return before the division point (Return_Before). The proxy of the intraday attention is the square of Return_Before (**SQ_Return_Before**). In this table, I mainly focus on: 1. the retail net buying after the division point (NetBuy After); 2. The return after the division point (Return_After); 3. The overnight return next day (CTO next). For reference purpose, I also report the retail net buying before the division point (NetBuy Before) and the intraday return next day (OTC next). In specific:

$$\text{Return\_Before}_t = \log(\text{price at division point}_t / \text{close}_{t-1}) \qquad \text{NetBuy Before} = \frac{\text{Retail Net Buy Volume Before the division point}}{\text{Total Retail Volume Before the division point}}$$

$$\text{Return\_After}_t = \log(\text{close}_t / \text{price at division point}_t) \qquad \text{NetBuy After} = \frac{\text{Retail Net Buy Volume After the division point}}{\text{Total Retail Volume After the division point}}$$

$$\text{SQ\_Return\_Before}_t = \text{Return\_Before}_t^{\,2} \qquad \text{CTO next} = \text{CTO}_{t+1} = \log(\text{open}_{t+1} / \text{close}_t) \qquad \text{OTC next} = \text{OTC}_{t+1} = \log(\text{close}_{t+1} / \text{open}_{t+1})$$

The table is constructed by first dividing all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (SQ_Return_Before) into three portfolios, namely, Low, Medium and High. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference between High and Low portfolio. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significances are based on the Newey and West (1986) adjusted standard errors of the time-series means.

		9:30-9:45		9:45-15:00		Next trading day	
Attention	No. Firms	NetBuy Before	Return_Before	NetBuy After	Return_After	CTO next	OTC next
Low	305	-0.042***	-0.054***	-0.039***	0.126***	-0.137***	0.182***
Medium	408	-0.043***	-0.248***	-0.04***	0.209***	-0.14***	0.191***
High	305	-0.01***	-0.433***	-0.023***	0.511***	-0.267***	0.241***
Diff (H-L)		0.033***	-0.379***	0.016***	0.385***	-0.130***	0.059***
t (Diff)		13.36	-4.73	15.99	4.95	-11.41	4.52

Table IV

Controlling the retail net buying, the effect of the intraday attention on the rest of the day: 15 minutes as the intraday division point

Using the time **15 minutes** after the market open as an intraday division point, I can calculate the return before the division point (Return_Before). The proxy of intraday attention is the square of Return_Before (**SQ_Return_Before**). The variables in this table is defined same as Table III.

I initially divide all the stocks into five NetBuy tertiles, based on the absolute value of the retail net buying before the division point (The absolute of **NetBuy Before**). Within each NetBuy tertiles, I do the same sorting as the above: Firstly, dividing all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (SQ_Return_Before) into three portfolios, namely, Low, Medium and High. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference between High and Low portfolio. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significance are based on the Newey and West (1986) adjusted standard errors of the time-series means.

			9:30-9:45		9:45-15:00		Next trading day	
NetBuy quintiles	Attention	No. Firms	NetBuy Before	Return_ Before	NetBuy After	Return_ After	CTO next	OTC next
1 (Low)	Low	61	-0.001***	-0.035***	-0.036***	0.144***	-0.148***	0.179***
	Medium	81	-0.002***	-0.193***	-0.035***	0.22***	-0.15***	0.181***
	High	61	-0.001***	-0.266***	-0.02***	0.383***	-0.299***	0.251***
	Diff (H-L)		0	-0.271***	0.017***	0.285***	-0.151***	0.066***
2	Low	61	-0.01***	-0.041***	-0.037***	0.131***	-0.15***	0.173***
	Medium	82	-0.011***	-0.159***	-0.036***	0.158***	-0.155***	0.198***
	High	61	-0.004***	-0.318***	-0.021***	0.424***	-0.293***	0.241***
	Diff (H-L)		0.006***	-0.277***	0.016***	0.293***	-0.143***	0.068***
3	Low	61	-0.029***	-0.053***	-0.038***	0.119**	-0.136***	0.173***
	Medium	82	-0.029***	-0.204***	-0.039***	0.151***	-0.14***	0.192***
	High	61	-0.009***	-0.382***	-0.024***	0.454***	-0.282***	0.263***
	Diff (H-L)		0.02***	-0.329***	0.014***	0.335***	-0.146***	0.09***
4	Low	61	-0.058***	-0.068***	-0.041***	0.119**	-0.128***	0.178***
	Medium	82	-0.053***	-0.236***	-0.042***	0.154***	-0.141***	0.196***
	High	61	-0.016***	-0.434***	-0.025***	0.461***	-0.259***	0.24***
	Diff (H-L)		0.041***	-0.366***	0.016***	0.342***	-0.131***	0.063***
5 (High)	Low	61	-0.12***	-0.08***	-0.044***	0.117**	-0.123***	0.198***
	Medium	81	-0.103***	-0.28***	-0.047***	0.185**	-0.124***	0.2***
	High	61	-0.034***	-0.764***	-0.028***	0.821***	-0.199***	0.221***
	Diff (H-L)		0.086***	-0.684***	0.016***	0.704***	-0.076***	0.023

Table V

The effect of the intraday attention on the rest of the day: 2h as the intraday division point

Using the time **2h** after the market open as an intraday division point, I calculate the return before the division point (Return_Before). The variables in this table is defined same as Table III.

The table is constructed by first dividing all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (SQ_Return_Before) into three portfolios, namely, Low, Medium and High. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference between High and Low portfolio. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significance are based on the Newey and West (1986) adjusted standard errors of the time-series means.

		9:30-11:30		13:00-15:00		Next trading day	
Attention	No. Firms	NetBuy Before	Return_Before	NetBuy After	Return_After	CTO next	OTC next
Low	305	-0.055***	-0.066***	-0.033***	0.133***	-0.139***	0.158***
Medium	408	-0.053***	-0.358***	-0.036***	0.198**	-0.138***	0.195***
High	305	-0.017***	-1.054***	-0.014***	1.3***	-0.266***	0.26***
Diff (H-L)		0.037***	-0.988***	0.020***	1.167***	-0.127***	0.102***
t (Diff)		17.6	-8.5	7.45	10.28	-9.96	7.84

Table VI

Controlling the retail net buying, the effect of the intraday attention on the rest of the day: 2h as the intraday division point

Using the time **2h** after the market open as an intraday division point, I can calculate the return before the division point (Return_Before). The proxy of intraday attention is the square of Return_Before (**SQ_Return_Before**). The variables in this table is defined same as Table III.

I initially divide all the stocks into five NetBuy tertiles, based on the absolute value of the retail net buying before the division point (The absolute of **NetBuy Before**). Within each NetBuy tertiles, I do the same sorting as the above: Firstly, dividing all stocks into three groups according to past size (average market capitalization of previous 20 trading days). Secondly, within each size group, I sort the stocks by the attention (SQ_Return_Before) into three portfolios, namely, Low, Medium and High. Then, within each portfolios, I calculate the cross-sectional averages of each variables every day. Lastly, I report the time-series mean of each variables and the difference between High and Low portfolio. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significance are based on the Newey and West (1986) adjusted standard errors of the time-series means.

			9:30-11:30		13:00-15:00		Next trading day	
NetBuy quintiles	Attention	No. Firms	NetBuy Before	Return_ Before	NetBuy After	Return_ After	CTO next	OTC next
1 (Low)	Low	61	-0.003***	0.025	-0.022***	0.14***	-0.167***	0.145***
	Medium	81	-0.003***	0.017	-0.021***	0.13**	-0.169***	0.189***
	High	61	-0.001***	-0.144	-0.005*	0.841***	-0.324***	0.301***
	Diff (H-L)		0.002***	-0.168	0.017***	0.701***	-0.156***	0.156***
2	Low	61	-0.021***	-0.001	-0.025***	0.126***	-0.153***	0.161***
	Medium	82	-0.017***	-0.132**	-0.026***	0.135**	-0.155***	0.2***
	High	61	-0.005***	-0.4***	-0.01***	0.881***	-0.312***	0.303***
	Diff (H-L)		0.016***	-0.399***	0.015***	0.754***	-0.16***	0.142***
3	Low	61	-0.048***	-0.068***	-0.039***	0.118***	-0.133***	0.153***
	Medium	82	-0.039***	-0.315***	-0.033***	0.151*	-0.141***	0.203***
	High	61	-0.013***	-0.759***	-0.014**	1.009***	-0.287***	0.289***
	Diff (H-L)		0.036***	-0.69***	0.024**	0.892***	-0.155***	0.136***
4	Low	61	-0.082***	-0.147***	-0.036***	0.134***	-0.119***	0.16***
	Medium	82	-0.068***	-0.461***	-0.041***	0.141*	-0.127***	0.193***
	High	61	-0.029***	-1.195***	-0.02***	1.193***	-0.243***	0.232***
	Diff (H-L)		0.053***	-1.048***	0.016***	1.059***	-0.124***	0.072***
5 (High)	Low	61	-0.139***	-0.232***	-0.05***	0.142**	-0.104***	0.168***
	Medium	81	-0.119***	-0.681***	-0.055***	0.237**	-0.121***	0.191***
	High	61	-0.048***	-2.765***	-0.017***	2.606***	-0.169***	0.189***
	Diff (H-L)		0.091***	-2.574***	0.033***	2.51***	-0.065***	0.015

Table VII

The effect median-adjusted intraday attention on the rest of the trading day: 15 minutes as the intraday division point

This table uses the time **15 minutes** after the market open as an intraday division point. Unlike Table III, here I adjust the attention by its median. The median-adjusted intraday attention is calculated by:

$$\text{Median-adjusted intraday attention}_t = (\text{Return\_Before}_t - \text{median of Return\_Before}_t)^2$$

The table is constructed same as Table III with the same variables, except I use the median-adjusted intraday attention to define the Low, Medium and High portfolio. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. The significance are based on the Newey and West (1986) adjusted standard errors of the time-series means.

		9:30-9:45		9:45-15:00		Next trading day	
Median-adjusted attention	No. Firms	NetBuy Before	Return_ Before	NetBuy After	Return_ After	CTO next	OTC next
Low	305	-0.052***	-0.228***	-0.043***	0.172**	-0.117***	0.2***
Medium	408	-0.04***	-0.27***	-0.039***	0.212***	-0.141***	0.187***
High	305	-0.004	-0.23***	-0.02***	0.462***	-0.278***	0.229***
Diff (H-L)		0.048***	-0.002	0.023***	0.291***	-0.155***	0.029**
t (Diff)		18.46	-0.07	20.99	8.21	-25.77	2.17

Table VIII

The effect of the intraday attention on retail net buying: Fama-Macbeth regression

In this table, I test the following equation:

$$NetBuy\_After = \beta_0 + \beta_1 SQ\_Return\_Before + controls$$

Same as Table III, I use the time **15min** after the market open as an intraday division point to calculate *NetBuy_After* and *SQ_Return_Before*, which are the retail net buying after the division point and the proxy of the intraday attention respectively. According to **Hypothesis 1**, the coefficient of *SQ_return_Before* (β_1) is expected to be positive.

The control variables include *MVlog*, *BMlog*, *MOM(-1,0)* and *MOM(-5, -1)*. *MVlog* is the log of the average capitalization of the previous 22 trading day. *BMlog* is the log of the book-to-market ratio. *MOM(i, j)* is the average past close-to-close return from day *i* to day *j-1*.

I run the cross-sectional regression each day and report the time-series mean of the estimate of the coefficients. Each models include several control variables. Model 3 includes all the control variables. All the variables are winsorized at 0.5% and 99.5% each day. The t-statistics are calculated based on Newey and West (1986) adjusted standard errors and report in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Dependent variable: NetBuy_After			
	1	2	3
SQ_Return_Before	0.0030***	0.0030***	0.0031***
	(27.76)	(27.36)	(28.22)
MVlog		0.0092***	0.0093***
		(9.71)	(9.72)
BMlog		-0.0051***	-0.0049***
		(-10.7)	(-11.1)
MOM(-1, 0)			0.0010***
			(4.3)
MOM(-5, -1)			-0.0006***
			(-6.79)

Table IX

The effect of the intraday attention on the rest of the intraday return: Fama-Macbeth regression

In this table, I test the following equation:

$$Return\_After = \beta_0 + \beta_1 SQ\_Return\_Before + controls$$

Same as Table III, I use the time **15min** after the market open as an intraday division point to calculate *Return_After* and *SQ_Return_Before*, which are the return after the division point and the proxy of the intraday attention respectively. According to **Hypothesis 2**, the coefficient of *SQ_return_Before* (β_1) is expected to be positive. The control variables are defined same as in Table VIII.

I run the cross-sectional regression each day and report the time-series mean of the estimate of the coefficients. Each models include several control variables. Model 3 includes all the control variables. All the variables are winsorized at 0.5% and 99.5% each day. The t-statistics are calculated based on Newey and West (1986) adjusted standard errors and report in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Dependent variable: Return_After			
	1	2	3
SQ_Return_Before	0.0130***	0.0119***	0.0120***
	(10.36)	(8.89)	(9.26)
MVlog		-0.0688***	-0.0700***
		(-8.99)	(-9.21)
BMlog		-0.0021	-0.0027
		(-0.31)	(-0.42)
MOM(-1, 0)			0.0149***
			(5.81)
MOM(-5, -1)			-0.0078***
			(-6.73)

Table X

The effect of the intraday attention on the overnight return next day: Fama-Macbeth regression

In this table, I test the following equation:

$$\text{CTO next} = \beta_0 + \beta_1 \text{SQ\_Return\_Before} + \text{controls}$$

Same as Table III, I use the time **15min** after the market open as an intraday division point to calculate *CTO next* and *SQ_Return_Before*, which are the overnight return next day and the proxy of the intraday attention respectively. According to **Hypothesis 3**, the coefficient of *SQ_return_Before* (β_1) is expected to be negative. The control variables are defined same as in Table VIII.

I run the cross-sectional regression each day and report the time-series mean of the estimate of the coefficients. Each models include several control variables. Model 3 includes all the control variables. All the variables are winsorized at 0.5% and 99.5% each day. The t-statistics are calculated based on Newey and West (1986) adjusted standard errors and report in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Dependent variable: CTO next			
	1	2	3
SQ_Return_Before	-0.0169***	-0.0166***	-0.0140***
	(-20.51)	(-20.05)	(-17.07)
MVlog		0.0425***	0.0386***
		(14.97)	(14.42)
BMlog		0.0179***	0.0129***
		(6.74)	(4.99)
MOM(-1, 0)			-0.0235***
			(-25.07)
MOM(-5, -1)			-0.0125***
			(-30.93)

Table XI

The effect of the intraday attention using the ranking of the attention: Fama-Macbeth regression

In this table, I test the following equation:

$$\text{Dependent Variable} = \beta_0 + \beta_1 \text{Score\_Attention} + \text{controls}$$

I use the time **15min** after the market open as an intraday division point to calculate the dependent variables and *SQ_Return_Before*. The dependent variables in different models are *NetBuy_After*, *Return_After* and *CTO next*, which represent the retail net buying after the division point, the return after the division point and the overnight return next day, respectively. *SQ_Return_Before* is the proxy of the intraday attention. I transpose *SQ_Return_Before* to *Score_Attention* which is a daily ranking score adjusted to range between 0 and 1:

$$\text{Score\_Attention} = (\text{Rank of SQ\_Return\_Before} - 1) / (\text{Number of stocks} - 1)$$

According to Hypotheses, the coefficients of *Score_Attention* (β_1) are expected to be positive. The control variables are defined same as in Table VIII.

I run the cross-sectional regression each day and report the time-series mean of the estimate of the coefficients. Each models have different dependent variables and include all the control variables. Except *Score_Attention*, all the variables are winsorized at 0.5% and 99.5% each day. The t-statistics are calculated based on Newey and West (1986) adjusted standard errors and report in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	1	2	3
Dependent variable:	NetBuy_After	Return_After	CTO next
Score_Attention	0.0225***	0.2559**	-0.1503***
	(18.62)	(2.3)	(-9.78)
MVlog	0.0090***	-0.0781***	0.0402***
	(9.5)	(-9.33)	(15.45)
BMlog	-0.0052***	-0.0182**	0.0125***
	(-11.72)	(-2.21)	(4.88)
MOM(-1, 0)	0.0014***	0.0192***	-0.0232***
	(5.96)	(7.33)	(-24.74)
MOM(-5, -1)	-0.0005***	-0.0068***	-0.0126***
	(-5.68)	(-5.76)	(-30.47)

Table XII

The effect of the intraday attention under a different intraday division point: Fama-Macbeth regression

In this table, I test the following equation:

$$\text{Dependent Variable} = \beta_0 + \beta_1 SQ\_Return\_Before + controls$$

I use the time **2h** after the market open as an intraday division point to calculate the dependent variables and *SQ_Return_Before*. The dependent variables in different models are *NetBuy_After*, *Return_After* and *CTO next*, which represent the retail net buying after the division point, the return after the division point and the overnight return next day, respectively. *SQ_Return_Before* is the proxy of the intraday attention. According to Hypotheses, the coefficients of *SQ_return_Before* (β_1) are expected to be positive. The control variables are defined same as in Table VIII.

I run the cross-sectional regression each day and report the time-series mean of the estimate of the coefficients. Each models have different dependent variables and include all the control variables. All the variables are winsorized at 0.5% and 99.5% each day. The t-statistics are calculated based on Newey and West (1986) adjusted standard errors and report in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	1	2	3
Dependent variable:	NetBuy_After	Return_After	CTO next
Attention	0.0016***	0.0196***	-0.0063***
	(19.49)	(23.34)	(-16.53)
MVlog	0.0059***	-0.0447***	0.0401***
	(5.66)	(-9.69)	(16.09)
BMlog	-0.0048***	0.0093**	0.0158***
	(-10.03)	(2.39)	(6.15)
MOM(-1, 0)	0.0012***	-0.0078***	-0.0257***
	(5.42)	(-4.41)	(-27.61)
MOM(-5, -1)	-0.0006***	-0.0059***	-0.0130***
	(-5.8)	(-8.28)	(-31.93)

Table XIII

The effect of the intraday attention using the ranking of the attention under a different intraday division point: Fama-Macbeth regression

In this table, I test the following equation:

$$\text{Dependent Variable} = \beta_0 + \beta_1 \text{Score\_Attention} + \text{controls}$$

I use the time **2h** after the market open as an intraday division point to calculate the dependent variables and *SQ_Return_Before*. The dependent variables in different models are *NetBuy_After*, *Return_After* and *CTO next*, which represent the retail net buying after the division point, the return after the division point and the overnight return next day, respectively. *SQ_Return_Before* is the proxy of the intraday attention. I transpose *SQ_Return_Before* to *Score_Attention* which is a daily ranking score adjusted to range between 0 and 1:

$$\text{Score\_Attention} = (\text{Rank of SQ\_Return\_Before} - 1) / (\text{Number of stocks} - 1)$$

According to Hypotheses, the coefficients of *Score_Attention* (β_1) are expected to be positive. The control variables are defined same as in Table VIII.

I run the cross-sectional regression each day and report the time-series mean of the estimate of the coefficients. Each models have different dependent variables and include all the control variables. Except *Score_Attention*, all the variables are winsorized at 0.5% and 99.5% each day. The t-statistics are calculated based on Newey and West (1986) adjusted standard errors and report in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	1	2	3
Dependent variable:	NetBuy_After	Return_After	CTO next
Score_Attention	0.0264***	1.4198***	-0.1498***
	(12.67)	(8.11)	(-8.85)
MVlog	0.0056***	-0.1114***	0.0398***
	(5.4)	(-14.23)	(15.77)
BMlog	-0.0051***	-0.0641***	0.0129***
	(-10.84)	(-6.7)	(5.14)
MOM(-1, 0)	0.0014***	0.0384***	-0.0237***
	(6.76)	(9.91)	(-25.5)
MOM(-5, -1)	-0.0005***	-0.0010	-0.0127***
	(-5.1)	(-0.67)	(-30.88)