JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS

doi:10.1017/S0022109024000085

Retail Trading and Return Predictability in China

Charles M. Jones
Columbia Business School
cj88@gsb.columbia.edu

Donghui Shi
Fudan University Fanhai International School of Finance
dhshi@fudan.edu.cn

Xiaoyan Zhang
Tsinghua University PBC School of Finance
zhangxiaoyan@pbcsf.tsinghua.edu.cn (corresponding author)

Xinran Zhang
Central University of Finance and Economics School of Finance
zhangxinran@cufe.edu.cn

Abstract

Using comprehensive account-level data, we separate Chinese retail investors into 5 groups and document strong heterogeneity in trading dynamics and performances. Retail investors with smaller account sizes cannot predict future returns correctly, display daily momentum patterns, fail to process public news, and show overconfidence and gambling preferences, while retail investors with larger account balances predict future returns correctly, display contrarian patterns, and incorporate public news in trading. Using performance measures established in previous literature, we find that smaller retail investors suffer from poor stock selection abilities and trading costs, while large retail investors' stock selection abilities are offset by trading costs.

I. Introduction

Retail investors are important participants in financial markets, and many studies using data from the United States (U.S.) and Europe are devoted to understanding their trading motives, performance, and role in information transmission and price discovery.[1] China's equity market, which is the second largest

We thank an anonymous referee, Hendrik Bessembinder (the editor), Utpal Bhattacharya, Darwin Choi, Lauren H. Cohen, Ron Kaniel, Terrance Odean, Xintong Zhan, Hao Zhou, and seminar participants at Tsinghua PBC School of Finance, Renmin University, Shanghai Jiaotong University, Fudan University, Shanghai University of Finance and Economics, and conference audiences at the 2019 CIFFP, 2021 CFRC, 2021 CICF, and 2022 ABFER Annual Conferences for their helpful comments and suggestions. Xiaoyan Zhang acknowledges the financial support from the National Natural Science Foundation of China (Grant Nos. 72350710220 and 71790605). Xinran Zhang acknowledges the financial support from the National Natural Science Foundation of China (Grant No. 72303268). All remaining errors are our own.

[1]Related studies include Barber and Odean (2000), (2001), (2008), Barber, Odean, and Zhu (2009), Kaniel, Gideon, and Sheridan (2008), Kelley and Tetlock (2013), Boehmer, Jones, Zhang, and Zhang (2021), Barber, Huang, Odean, and Schwarz (2022), Eaton, Green, Roseman, and Wu (2022), and Welch (2022)).

in the world, provides an important setting for studying retail investors. According to the annual report of the Shanghai Stock Exchange in 2016, retail investors contribute about 85% of the daily trading volume on the exchange, whereas institutional investors contribute only around 15%. Behind the high trading volumes are tens of millions of retail investors in China, which accounts for the largest population of retail investors in the world. The dominance and prevalence of retail trading in China brings retail investors to the central stage of the capital market, making it crucial for researchers, regulators, and practitioners to understand their investment choices and how those choices affect information transmission and price discovery.

An easy reference for answering these questions might appear to be the earlier studies that use data from the U.S. or Europe. However, the development of the Chinese stock market has been substantially different from that of the U.S. or Europe. For instance, the U.S. market displays typical features of developed capital markets, such as the dominance of institutional investors in terms of trading and holding, mature information environment, and established legislature system. Meanwhile, the Chinese stock market displays typical features of emerging markets, with rising institutional investors accounting for only a small part of the daily trading volume, high valuations (or lower earnings yields), high volatilities, and high turnovers. Given these significant differences, it is not clear whether previous studies' conclusions regarding retail investors in other developed countries would readily apply to China, and a focused study of Chinese retail investors becomes essential.

Our study focuses on 2 core questions to understand retail investors in China: First, how do retail investors contribute to the price discovery process? In other words, can they predict or fail to predict future price movements? Second, if retail investors succeed or fail in predicting future price movements, what are the driving forces of their trading behaviors? Liquidity, information, or behavioral biases? Clearly, millions of Chinese retail investors in our sample are not homogenous individuals, and the answers to these 2 questions likely differ drastically for different groups of retail investors. Therefore, we rely on the rich cross section of our data to examine the heterogeneity of retail investors and how their trading interacts with stock returns, information flows, and behavioral features.

We are grateful that a major stock exchange in China provides us the access to proprietary account-level trading and holdings data from 2016 to 2019, accounting for over 53 million retail accounts. To comply with regulatory requirements, the exchange categorizes all Chinese retail accounts into 5 groups based on their account balances: less than 100,000 CNY (RT1), between 100,000 and 500,000 CNY (RT2), between 500,000 and 3,000,000 CNY (RT3), between 3,000,000 and 10,000,000 CNY (RT4), and greater than 10,000,000 CNY (RT5). The 5 groups account for 58.7%, 28.6%, 10.9%, 1.4%, and 0.4% of the total number of accounts, respectively. In other words, the majority of Chinese retail investors have account sizes below 500,000 CNY.[2]

To answer the first research question on how retail investors contribute to price discovery, we directly examine whether retail investors' buy and sell activities

[2]Additional gender and age information shows that most Chinese retail investors are young or middle-aged males.

predict future stock returns. If the market is perfectly efficient, stock prices would follow random walks and retail trading would not predict future returns. If the market is not perfectly efficient and some investors have value-relevant information for future stock prices, their order flows would positively predict future returns. On the other hand, if some investors have an information disadvantage, or fail to incorporate timely information into their trades, their order flows might negatively predict future returns. Using daily retail order imbalances (OIB) from each retail group, we examine their predictive power for future stock returns at different horizons. The smaller retail investors, RT1–RT4, predict next-day returns with negative and significant coefficients. That is, the prices of stocks they purchase experience negative returns the following day, whereas the stocks they sell experience positive returns. By contrast, the largest retail investors, RT5, positively and significantly predict next-day returns, indicating that they purchase and sell stocks in directions consistent with future price movements. When we examine longer horizons, the abovementioned predictive patterns persist for about 8–9 weeks. These patterns remain quite robust when we form long–short strategies on order flow information and for subsets of stocks with differences in size, value, liquidity, and share price levels.

The data allow us to examine the return predictive patterns among the characteristics that are not easily available in most other data sets. For instance, we directly observe the counterparties of each trade and find that the order flows from small retail investors always predict returns negatively and do not depend on counterparties. Interestingly, for large retail and institutional investors, their order flows typically predict future returns positively, but they become negative when they trade on the same side as small retail investors, suggesting that their prediction becomes erroneous when on the same side as smaller retail investors. We also examine how holding horizons are related to the predictive patterns of retail order flow. Small retail investors' negative return predictive power is stronger over the short horizon and slowly diminishes when holding horizons become longer, indicating that a longer holding horizon might lead to better returns. In contrast, large retail investors' positive return predictive power is stronger over the short horizon, which suggests that they might trade on time-sensitive information, and the predictive power becomes weaker for longer horizons.

For the second research question regarding trading motives of retail investors, previous literature provides multiple explanations, such as order flow persistence, liquidity provision, behavioral biases, and information (dis)advantages. These explanations also naturally connect with the predictive pattern of retail order flows for future returns. We adopt the 2-stage decomposition procedure in Boehmer et al. (2021) to examine whether these hypotheses explain the trading activities of different retail investor groups and how these trading activities contribute to the predictive patterns for future returns. Our results show that small retail investors display significant order flow persistence and that their trades have momentum patterns at daily horizons that demand immediate liquidity. Smaller retail investors also display strong behavioral biases, such as overconfidence and gambling preferences, and fail to predict and process informational earnings news. In explaining order flow's predictive power for future returns, order persistence, daily momentum trading patterns, behavioral biases, and information disadvantages all contribute to the negative predictive power of smaller retail investors. In contrast, the largest retail investors display contrarian trading patterns, trade against the behavioral biases of other retail groups, and are capable of predicting and processing earnings news. All of these features contribute to the positive predictive power of the largest retail investors.

To further understand the performance of retail trading flows, we follow Barber, Lee, Liu, and Odean (2009) to construct net buy and sell portfolios, track the net positions of each group of retail investors, and compute the performance of the net buy and sell portfolios. Consistent with previous findings that smaller retail investors negatively predict future returns, while large investors positively predict future returns, we find that the net buy and sell portfolios tracking smaller retail investors' order flows have an annual return of −5.61%, while the net buy and sell portfolios tracking large retail investors have an annual return of −0.29%. We also decompose the performance into 3 components: stock selection, market timing, and trading cost. We find that stock selection and trading costs are the 2 major drivers of total returns in the tracking portfolios.

This study is closely related to the retail investor literature. As mentioned earlier, previous studies on retail investors primarily use data from the U.S. and other markets and tend to treat retail investors as one group.[3] For instance, using data from a discount broker in the U.S., Barber and Odean (2000), (2001), (2008) document many behavioral biases. Kaniel et al. (2008), Barber et al. (2009), Kelley and Tetlock (2013), and Boehmer et al. (2021) use different data sets from the U.S. and find that retail trading can positively predict the cross section of future returns. Outside the U.S., Grinblatt and Keloharju (2000), Linnainmaa (2010), and Grinblatt, Keloharju, and Linnainmaa (2012) focus on Finland data; Bach, Calvet, and Sodini (2020) focus on Swedish data; and Dorn, Huberman, and Sengmueller (2008) study data from Germany. All of these studies provide important results regarding retail investors' trading activities. Interestingly though, the results are not always consistent, and evidence from China will provide more insights into this literature.

This study is also related to the recent studies of the rapidly growing Chinese stock market. Liu, Stambaugh, and Yuan (2019) and Liu, Zhou, and Zhu (2024) establish asset pricing factors for stock returns. For Chinese retail investors, An, Lou, and Shi (2022) study the wealth redistribution role of financial bubbles and crashes over July 2014 and Dec. 2015, documenting a net transfer of 250 billion CNY from the poor to ultra-wealthy retail investors over this period. Liu, Peng, Xiong, and Xiong (2021) and Liao, Peng, and Zhu (2022) both focus on the behavioral features of Chinese retail investors and document overconfidence, gambling preferences, and extrapolative expectations in these investors. Li, Geng, Subrahmanyam, and Yu (2017), Chen, Gao, He, Jiang, and Xiong (2019), Hu, Liu, and Xu (2021), Jiang, Liu, Peng, and Wang (2022), and Titman, Wei, and Zhao (2022) all focus on an earlier Chinese sample period and examine behavioral biases and reactions to corporate events. These Chinese studies primarily rely on low-frequency data, or data from a small number of brokerages covering a small part of the market, or investigate issues other than return predictability. As a result, there

[3]To facilitate reading, Appendix Table I in the Supplementary Material provides a summary table of previous papers on retail investors using data from different markets, and Appendix Table II in the Supplementary Material provides comparisons between the findings of this study with the results from previous literature.

remains no direct study on the heterogeneity of retail investors' trading behavior, their return predictive power, and how they process information using high-frequency trading data in one major stock market dominated by retail investors.

Compared with previous studies, this study makes 3 important contributions. First, our study, with its large market coverage of the market for a recent sample period, is one of the most thorough and comprehensive studies of Chinese retail investors and provides many important implications for regulators, practitioners, and academic researchers. Second, we separate retail investors into groups based on account size and provide unique and direct evidence of investor heterogeneity in terms of return predictability, counterparties, holding horizons, and performance. Third, we examine different hypotheses regarding return prediction patterns for different retail investor groups and provide clear evidence on the sources of the negative or positive predictive power of different retail investors.

II. Data

A. Data on Retail Investors

Our data come from one major stock exchange in China, which contains the trading and holding history of all stocks listed on the exchange from all investors, between Jan. 2016 and June 2019. This proprietary data set contains roughly 53 million accounts, and based on investor identities, they are grouped into 3 major categories: retail (RT), institutional (INST), and corporations (CORP). Retail investors are further stratified into 5 groups based on their account sizes at the beginning of each year, which is the average portfolio value (including equity holdings in all A-share listed firms, plus cash balance) over the previous 12 months.[4] As mentioned in the introduction, there are 5 subgroups: below 100,000 CNY (RT1), 100,000–500,000 CNY (RT2), 500,000–3 million CNY (RT3), 3 million–10 million CNY (RT4), and above 10 million CNY (RT5). Since the focus of this study is how retail trades are related to stock prices in the cross section, we sum up individual investors' trading information at 7 investor group levels (RT1–RT5, INST, and CORP) for each stock each day.

Panel A of Table 1 presents the summary statistics for investor accounts. During the sample period, the total number of active accounts for retail investors, institutions, and corporations is 53.4 million, 40,000, and 47,000, respectively. Within the retail investor category, there are 31.4 million, 15.3 million, 5.8 million, 0.7 million, and 0.2 million accounts for RT1 to RT5, respectively. Clearly, most of the retail investors have accounts less than 500,000 CNY. The overall trading volume on this exchange averages 201 billion CNY per day, with retail investors, institutions, and corporations accounting for 81%, 17%, and 2% of the total trading

[4]The annual calculation of account balances is designed by the exchange. There is a concern that investors might migrate to different groups after the initial grouping at the beginning of the year. We acknowledge this possibility. However, the fact that the grouping is redone each year eases the concern to some extent. The cutoff numbers to separate retail investors into different groups are also chosen by the exchange to comply with exchange regulations. For instance, derivatives and leverage trading are only allowed for investors with account size higher than 0.5 million CNY in order to protect investors with lower account balances.

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

TABLE 1

Summary Statistics

Table 1 reports summary statistics for trading and holdings by different investor groups. Our sample period is Jan. 2016 to June 2019, and our sample firms are A-share stocks listed on a major stock exchange with at least 15 trading days during the previous month. Panel A reports the number of accounts, aggregate trading, and holdings by different types of investors. Panel B reports the time series average of the cross-sectional distribution of stock characteristics. Panel C presents the time series average of the cross-sectional statistics of order imbalances for each investor group. Order imbalances (OIB) are computed as the buy share volume minus sell share volume divided by buy share volume plus sell share volume for each investor group, as specified in equation (1).

Panel A. Number of Accounts, Trading, and Holdings by Different Types of Investors

	RT1	RT2	RT3	RT4	RT5	INST	CORP
Account Value	<100,000 CNY	(100,000, 500,000) CNY	(500,000, 3 million) CNY	(3 million, 10 million) CNY	>10 million CNY		
No. of accounts (thousands)	31,410	15,282	5,827	735	235	40	47
Aggregate trading volume (Bil. CNY)	9	35	54	27	37	35	3
Aggregate trading volume (% of total)	5%	17%	27%	13%	19%	17%	2%
Aggregate holding value (Bil CNY)	336	951	1,566	840	1,794	4201	15,547
Aggregate holding value (% of total)	1%	4%	6%	3%	7%	17%	62%

Panel B. Stock Characteristics

	Mean	Std. Dev.	P25	P50	P75
Market capitalization (billion CNY)	20.1	80.3	2.9	5.6	12.1
Earnings-to-price ratio	0.0075	0.0155	0.0018	0.0060	0.0122
Daily stock return	−0.01%	2.17%	−1.09%	−0.22%	0.77%
Daily turnover (of tradable A-shares)	2.45%	4.37%	0.61%	1.15%	2.40%

Panel C. Order Imbalance in the Cross Section by Different Types of Investors

	Mean	Std. Dev.	AR1	Correlations						
				OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST	OIB_CORP
OIB_RT1	−0.021	0.187	0.243	1						
OIB_RT2	−0.011	0.171	0.259	0.802	1					
OIB_RT3	−0.006	0.166	0.216	0.610	0.710	1				
OIB_RT4	0.002	0.250	0.059	0.194	0.244	0.256	1			
OIB_RT5	0.019	0.352	0.102	−0.151	−0.158	−0.163	−0.091	1		
OIB_INST	−0.011	0.455	0.205	−0.315	−0.365	−0.380	−0.263	−0.188	1	
OIB_CORP	−0.004	0.720	0.088	0.022	0.029	0.021	−0.007	−0.043	−0.044	1

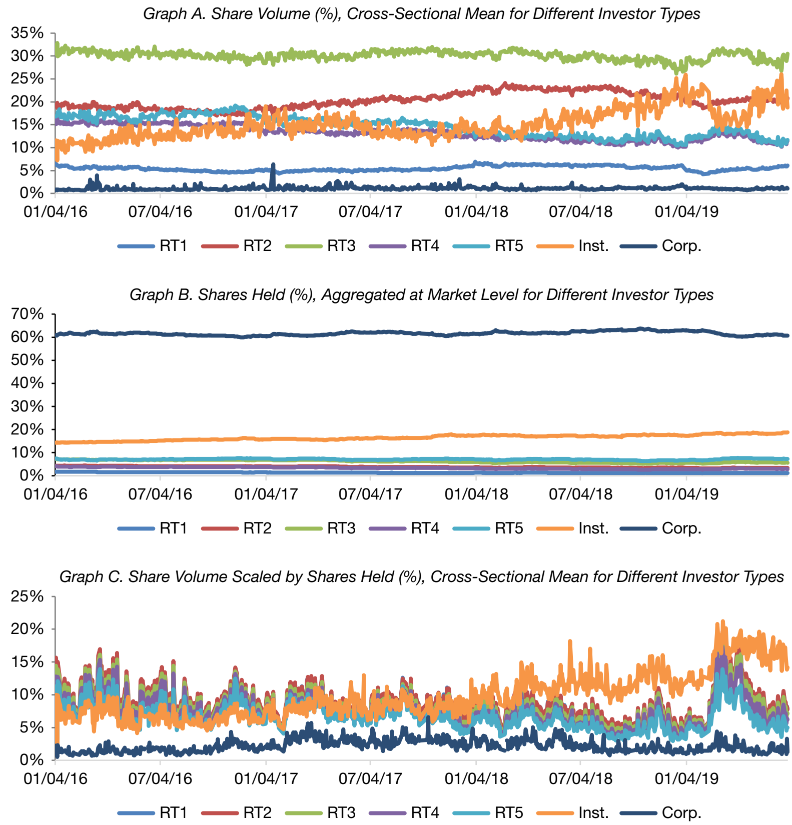
volume, respectively. Within the retail investor sector, the trading volumes for RT1 to RT5 are 5%, 17%, 27%, 13%, and 19% of the total trading volume, respectively, indicating that RT3 is the most active trading subgroup. Regarding stock holdings, retail investors' holdings account for 22%, institutions are 17%, and corporations are 62%. Within the retail investor sector, the account values for RT1 to RT5 are 1%, 4%, 6%, 3%, and 7% of the total tradable market cap, respectively. With low holdings and high trading volumes, Chinese retail investors trade quite actively; in some cases, they might even trade excessively.

To understand the relative importance of different investment groups' trading over time, we plot the time series of cross-sectional means of various investors' trading activities in Figure 1. Graph A presents each group's trading volume as a percentage of the total trading volume. Again, the RT3 group has the highest trading volume, accounting for about 30% of total trading. Interestingly, institutional trading gradually increased over time, from 10% in 2016 to over 20% in 2019.

FIGURE 1

Different Investor Type Order Flows Between Jan. 2016 and June 2019

Figure 1 reports the time-series plot of the different types of investor trading activity from Jan. 2016 to June 2019. Our sample firms are A-share stocks listed on a major stock exchange with at least 15 trading days during the previous month. Graph A presents time-series plot of the volume percentage for each investor type, reported as cross-sectional mean. Graph B shows the time-series plot of shares held by investor type, aggregated at market level. Graph C presents the time-series plot of each investor type's trading volume, scaled by the shares held by that investor type, and reported as the cross-sectional mean.



The corporations barely trade and account for a negligible amount of trading volume. Graph B displays the percentage of shares held by each group, and it can be seen that the time-series patterns in holdings are stable. Corporations hold a large chunk of shares, accounting for over 60% of total shares; institutions hold about 17% of the shares; large and medium retail investors hold about 3% to 6% of the shares; and the smallest retail investors RT1 hold around 1% of the shares. In Graph C, we compute relative trading activeness with respect to holdings for each group, by scaling each group's trading share volumes by each group's holdings at the stock level. The retail investors all follow similar time-series patterns, trading around 10% of their holdings each day on average, which is quite high. We observe a slow decrease in trading after 2018 to about 5%, but there is also a large but short-term surge of up to 15% at the beginning of 2019. Considering that the market return was −22.2% during 2018, and it reverts back to 21.2% during the first half of 2019, it is likely that the large negative returns reduce retail trading enthusiasm, while large positive returns reignite the enthusiasm. By the end of 2019, retail trading returns to the average level of 10% of holdings. We notice that among the 5 groups of retail investors, RT5 trades slightly less than the rest of the retail investors. We also observe a clear increasing pattern in the activeness of institutional trading, which increased from 7% in 2016 to 16% in 2019. Finally, corporates in general rarely trade, and their trading accounts for about 2% of their holdings.

Overall, in the Chinese stock market, retail investors dominate in trading, while corporations dominate in holdings. Retail investors' dominance in trading in the Chinese stock market is likely the joint result of market development, regulations, and investor preferences. This pattern is not particularly rare for emerging markets but is quite different from that of developed markets.[5]

B. Data on Stock Returns and Firm Characteristics

We obtain data on stock returns, volumes, and accounting information from Wind Information Inc. (WIND), the largest financial data provider in China. For consistency with the retail data, the sample period runs from Jan. 2016 to June 2019. We adopt the filters in Liu et al. (2019) and exclude stocks with less than 15 days of trading records during the most recent month. Liu et al. (2019) also eliminate stocks that became public within the past 6 months, stocks with fewer than 120 days of trading records during the past 12 months, and the smallest 30% of all firms listed in the Chinese A-share market. This study does not exclude these stocks from the main results because retail investors trade actively in small stocks and during a stock's initial public offering period. We present the results with all the filters from Liu et al. (2019) in our robustness checks, and the findings are similar to the additional filters. Starting from Mar. 31, 2010, margin buy and short sell are allowed on Chinese stock exchanges for subsets of stocks. We include these leveraged trades in our main results and provide an additional analysis that excludes leveraged trading in our robustness checks. We merge the exchange and WIND data by stock ticker and present account summary statistics. Our final sample covers over 1.1 million stock-day observations, and on each day, we have an average of around 1,200 firms.

Panel B of Table 1 presents several key variables for the Chinese capital market. Market capitalization is the product of the previous month's closing price and total A-shares outstanding. The average Chinese firm capitalization is 20.1 billion CNY, or 3 billion USD, about half of the cross-sectional average in the U.S. stock market during the same period, which is 6.9 billion USD. The earnings-to-price ratio (EP ratio) is the ratio of the most recently reported quarterly net profit, excluding nonrecurrent gains/losses over last month end's market capitalization.

[5]We present additional summary statistics on retail trading volumes and holdings in Appendix Table III in the Supplementary Material. The results show that small retail investors prefer to trade and hold small, low earning/price ratio, and high turnover firms, while the largest retail investors' trading and holdings are tilted toward larger, high earnings-to-price ratio firms. In terms of sectors, the small retail investors prefer to trade and hold the alternative energy sector and prefer not to trade and hold banks and life insurance firms, while the institutions and corporates behave in opposite ways. Finally, small retail investors tend to use small order sizes, while large retail investors tend to use large order sizes and institutions use all order sizes.

According to Liu et al. (2019), the EP ratio captures the value effect. The average EP ratio is 0.0075 in China, while the average EP ratio is 0.0272 in the U.S. stock market, which indicates high valuation ratios in China. The average daily stock return is −0.01% for Chinese stocks, while the average daily stock return is 0.04% in the U.S. stock market over the same sample period.[6] Finally, we compute daily turnover as daily share trading volume divided by tradable shares outstanding. The average daily turnover in China is 2.45%, indicating that it takes 1÷2.45% = 41 days to turn around all shares, which is much larger than the daily turnover of 1.11% in the U.S. during the same period. These summary statistics show that an average Chinese firm has smaller capitalization, higher valuation ratio, lower returns, and higher turnover than an average U.S. firm.

C. The Order Imbalance Measure

The order flows from different groups of investors are measured using order imbalances, as in Chordia and Subrahmanyam (2004). For stock *i,* day *d*, and investor group *G*, we compute

$$\text{(1)} \qquad \text{OIB}_{i,d,G} = \frac{\sum_{j \in G} \text{BUY\_VOL}_{i,d,j} - \sum_{j \in G} \text{SELL\_VOL}_{i,d,j}}{\sum_{j \in G} \text{BUY\_VOL}_{i,d,j} + \sum_{j \in G} \text{SELL\_VOL}_{i,d,j}},$$

where the numerator is the *difference* between buy and sell share volumes[7] for stock *i* on day *d*, summed over all individual *j*'s within group *G*, and the denominator is the *sum* of buy and sell share volumes for stock *i* on day *d*, summed over all individual *j*'s within group *G*. The data allow us to directly observe each trade's direction. When a set of investors buy more than they sell, the order imbalance is positive and vice versa. We compute the order imbalance measure for each investor group as OIB_RT1 to OIB_RT5, OIB_INST, and OIB_CORP.

Panel C of Table 1 reports the summary statistics for the order imbalance measures. The average order imbalances for RT1 to RT5, institutions, and corporations range between −0.021 and 0.019 and are all close to 0. The small magnitude of these average order imbalance measures indicates that most buys and sells within each investor group cancel each other out. The standard deviations of order imbalances are larger for large retail investors (0.352) and institutions (0.455), as compared to small and medium retail investors (between 0.166 and 0.250), which indicate that there is more cross-stock variation in large retail investor and institutional trading activity. The 1-day autocorrelation coefficient, AR1, for these OIB measures is 0.243, 0.259, 0.216, 0.059, and 0.102 for RT1 to RT5, respectively,

[6]Previous literature using the U.S. data shows that microstructure frictions can generate noise in daily return measures. For instance, Blume and Stambaugh (1983) show that daily returns computed from the end-of-day closing prices can have an upward bias due to bid–ask bounce. We compute this bias measure using the closing bid and ask prices for all A-share stocks listed on this major stock exchange. The average bias measure is generally below 0.0002% across all stocks, which is negligible compared to the bias computed in Blume and Stambaugh (1983). Therefore, we compute daily returns using daily close prices without Blume and Stambaugh's (1983) adjustments.

[7]In this study, "share volume" refers to the number of shares traded on that day, and "cash volume" refers to the product of share volume and the execution price.

suggesting that small and medium retail order imbalances are generally more persistent than large retail imbalances.

In terms of order flow correlations across the 7 groups, the order flows from smaller retail investors, OIB_RT1, OIB_RT2, and OIB_RT3, are highly correlated, with correlation coefficients primarily higher than 0.60. OIB_RT4 is still positively correlated with OIB_RT1–OIB_RT3, but with lower correlation coefficients of around 0.20. The largest retail investors' order imbalance, OIB_RT5, is negatively correlated with all 4 other groups, with correlations around −0.15, indicating that this group of retail investors might have different trading patterns from the others. Institutional order imbalances are negatively correlated with all 5 retail groups, with correlations between −0.380 and −0.188, again implying different trading patterns from retail investors, even the largest retail investors. As we find earlier, corporations barely trade and their correlations with the rest of the investor categories are all lower than 10%.[8]

In this data section, we include OIB_INST and OIB_CORP to ensure the completeness of the summary statistics. Given that corporations are long-term investors and rarely trade and that our study focuses on trading behavior, we exclude corporations from the remaining empirical results. In terms of institutional investors, given that retail investors are commonly assumed less sophisticated in comparison with institutional investors, we retain institutions in our main results for comparison purposes.

III. Can Retail Order Flows Predict Future Stock Returns?

A. Predicting Next-Day Stock Returns Using Retail Order Flows

To investigate the roles different retail investors play in the price discovery process, we first examine retail order flow's predictive power for the next day's return. We adopt the 2-stage Fama and MacBeth (1973) regression to examine the predictive patterns of order flows for next-day returns. At the first stage, we estimate the following cross-sectional regression for each day d within each investor group G:

$$(2) \qquad \mathrm{RET}_{i,d} = a0_{d,G} + a1_{d,G}\mathrm{OIB}_{i,d-1,G} + a2_{d,G}{}'\mathrm{CONTROLS}_{i,d-1} + u1_{i,d},$$

where the dependent variable $\mathrm{RET}_{i,d}$ is the stock return for firm i on day d, and the independent variables include order imbalance measures from investor group G from the previous day, $\mathrm{OIB}_{i,d-1,G}$, and control variables, $\mathrm{CONTROLS}_{i,d-1}$. We follow previous literature for the choice of control variables. To control for potential momentum/reversal from past returns, we include returns from the previous day, RET(−1), returns from the previous week, RET(−6, −2), and returns from the previous month, RET(−27, −7). For size, value, and liquidity effects, we include log market size (SIZE), EP_RATIO, and TURNOVER as controls, all computed from the previous month end.

[8]Appendix Figure I in the Supplementary Material presents the time series plot of the order imbalance measures for each investor group, and we find no evidence of time trends or breaks over our sample period.

TABLE 2

Predicting Next-Day Stock Returns Using Order Imbalances from Different Investor Groups

Table 2 reports the estimation results of whether trading activity by different investor groups can predict the cross section of future stock returns. The coefficients are estimated from Fama and MacBeth's (1973) regressions, as specified in equation (2). The dependent variable is the next-day stock return (RET), and the independent variable is the previous day's order imbalance OIB(−1). The control variables are the previous day's return RET(−1), previous week's return RET(−6,-2), previous month's return RET(−27,-7), previous month's log market cap (SIZE), earnings-to-price ratio (EP), and monthly turnover (TURNOVER). The interquartile range for the relevant explanatory order imbalance is to compute the difference in predicted future returns for the interquartile range (INTERQUARTILE_RETURN). To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using the Newey and West (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

		Dependent Variable: RET					
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
OIB(−1)	Estimate	−0.0093***	−0.0091***	−0.0065***	−0.0009***	0.0012***	0.0016***
	[*t*-stat]	[−17.36]	[−17.10]	[−15.74]	[−6.95]	[10.48]	[18.06]
RET(−1)	Estimate	−0.0027	−0.0091**	0.0006	0.0189***	0.0190***	0.0132**
	[*t*-stat]	[−0.61]	[−2.09]	[0.12]	[3.68]	[3.74]	[2.49]
RET(−6,−2)	Estimate	−0.0149***	−0.0132***	−0.0124***	−0.0120***	−0.0115***	−0.0113***
	[*t*-stat]	[−8.17]	[−7.15]	[−6.66]	[−6.41]	[−6.17]	[−6.08]
RET(−27,−7)	Estimate	−0.0039***	−0.0036***	−0.0034***	−0.0033***	−0.0032***	−0.0034***
	[*t*-stat]	[−4.35]	[−4.08]	[−3.84]	[−3.70]	[−3.61]	[−3.84]
SIZE	Estimate	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	[*t*-stat]	[0.36]	[0.17]	[−0.16]	[−0.32]	[−0.18]	[−0.21]
EP	Estimate	0.0147***	0.0150***	0.0145***	0.0144***	0.0146***	0.0140***
	[*t*-stat]	[2.88]	[2.89]	[2.75]	[2.77]	[2.82]	[2.72]
TURNOVER	Estimate	−0.0007***	−0.0007***	−0.0007***	−0.0007***	−0.0007***	−0.0007***
	[*t*-stat]	[−3.50]	[−3.65]	[−3.64]	[−3.79]	[−3.79]	[−3.79]
INTERCEPT	Estimate	−0.0012	−0.0006	0.0002	0.0005	0.0001	0.0002
	[*t*-stat]	[−0.47]	[−0.25]	[0.06]	[0.20]	[0.06]	[0.10]
Adjusted R^2		8.83%	8.68%	8.43%	8.10%	8.11%	8.25%
INTERQUARTILE_OIB		0.2222	0.1827	0.1678	0.2868	0.4536	0.6740
INTERQUARTILE_RETURN		−0.2062%	−0.1668%	−0.1089%	−0.0247%	0.0523%	0.1046%

From the first-stage estimation, we obtain the daily time series of coefficients for each investor group G, $\{a0_{d,G}, a1_{d,G}, a2'_{d,G}\}$. For the second-stage estimation, we conduct statistical inference based on the means, $\{a0_G, a1_G, a2'_G\}$, and standard errors of the first-stage coefficients. To be specific, we compute standard errors following the method in Newey and West (1987), with the optimal bandwidths chosen by using Newey and West's (1994) approach.[9] If the order flow variable from a specific investor group G predicts future returns in the correct direction, defined as an occasion when more past purchases are associated with higher future returns and more past sales are associated with lower future returns, we expect the coefficient $a1_G$ to be significantly positive and vice versa.[10]

The estimation results for equation (2) are reported in Table 2, which displays distinctive predictive patterns across different groups of retail investors.

[9]We also consider robustness check using BIC to select the optimal lag number. Results are similar and available from the authors.

[10]As an alternative to the Fama–Macbeth regression, we also adopt the portfolio sorting approach, using order imbalance measures from different groups as sorting variables. Results are similar to the results of the Fama–Macbeth regression and are discussed in robustness check in Section VI.D.

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

For the smallest retail investor group, RT1, the coefficient on the retail order flow variable is −0.0093, with a significant *t*-statistic of −17.36. The negative coefficient indicates that if retail investors RT1 buy more than they sell on a given day, the next-day return on that stock is significantly negative. To understand the economic magnitude of the coefficients, we report the interquartile range for OIB_RT1 at the bottom of the table. Multiplying the interquartile range, 0.2222, by the regression coefficient of −0.0093 generates an interquartile daily return difference of −21 basis points (more than 50% annualized!). The predictive patterns are qualitatively similar for retail investors in groups RT2–RT4. All coefficients are negative and statistically significant, and the daily interquartile return differences are −17, −11, and −2 basis points for RT2, RT3, and RT4, respectively. That is to say, the first 4 groups of investors trade in the incorrect direction versus future price movements: The stocks they buy have lower returns, and the stocks they sell have higher returns. It is possible that these retail investors are short of finance literacy, or their trading displays behavioral biases, which lead to the wrong return predictions. Interestingly, when we move from smaller account sizes to larger ones, the negative coefficients become smaller, indicating that larger retail investors trade more correctly than smaller ones. Indeed, for the largest retail investors, RT5, the coefficient on the past day's order imbalance is 0.0012, which is positive and significant with a *t*-statistic of 10.48. The interquartile daily return difference is 5 basis points per day (over 12% per year). It appears that the trading of the largest retail investors predicts the cross section of future stock price movements in the correct direction.

For comparison, the coefficient on the previous day's order imbalance is 0.0016 for institutions, with a *t*-statistic of 18.06. That is to say, institutional order flows predict future stock price movements in the correct direction, and the interquartile return difference is 10 basis points per day (over 25% per year), which is approximately twice the magnitude of the RT5 estimate. This finding is consistent with many previous studies showing that institutional investors are generally more informed than retail investors.[11]

For the control variables, the coefficients on the previous day's return have mixed signs, while the coefficients on the previous week's and previous month's returns are all negative and significant, indicating strong reversals over the weekly and monthly horizons. Size is mostly insignificant, whereas the EP variables are always positive and significant, indicating a strong value effect. The coefficients on turnover are always negative and significant, suggesting that higher turnover leads to lower returns in the future. These findings are consistent with those of previous studies on the Chinese stock market, such as Liu et al. (2019). These results also confirm that the predictive power of the various order flow variables for future stock returns is not a manifestation of the size, value, liquidity, or momentum/reversal effect.

[11]Readers might wonder whether the RT5 are corporate insiders, which allow them to have inside information and positive prediction for future returns. China has strict rules on the windows during which insiders can trade related stocks. For instance, they cannot trade before significant corporate announcements, such as earnings announcements. Later results show that RT5 trades actively around earnings announcements, which support the notion that they are unlikely to be corporate insiders.

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

B. Predictive Patterns with Different Counterparties

Trading dynamics among counterparties is an interesting and important concept that helps to understand how different groups of investors interact with each other. However, most existing studies cannot offer much insight into this because it is difficult to pin down each trade's counterparties due to data limitation.[12] We are fortunate to have the whole trading history from the exchange, which makes it possible for us to identify the counterparties' group for each trade. Therefore, in this subsection, we closely examine the trading dynamics among different groups of investors and whether different trading pairs are associated with different return predictive patterns.

We group trades by counterparties from the buy and sell sides of each trade in 3 steps. First, because corporations rarely trade and account for less than 2% of daily trading volume, we exclude trades with corporations. Second, we regroup the remainder of the 6 groups into 3 larger groups to reduce overall dimensions: i) RT1 to RT4 are bundled together as one group because their order imbalances are positively correlated, and they share similar predictive patterns and ii) RT5 and INST are considered as 2 separate groups because their order flows are negatively correlated with those from smaller retail investors, and the order flows from RT5 and INST are also negatively correlated. For the final step, with 3 groups of investors and 2 sides of trades (buy or sell), we divide all the remaining trades into 6 bins: BBS, BSB, BSS, SBB, SBS, and SSB. The first letter indicates the trade direction of RT1–RT4, the second indicates the trade direction of RT5, and the third indicates INST. For instance, "BBS" means RT1–RT4 "buy," RT5 "buy," and INST "sell." Panel A of Table 3 presents the proportion of each type of trade. The highest proportions are "BSS" and "SBB," both above 20%, where RT5 is on the same side as institutions. The lowest proportions are for "BSB" and "SBS," both lower than 13%, indicating that RT1–RT4 are less likely to be on the same side as institutions.

To address whether the return predictive patterns change when the counterparties are different, we modify equation (2) to create equation (3), as follows:

$$\text{(3)} \qquad \text{RET}_{i,d} = b0_{d,G} + \left(\sum_{k=1}^{6} b1_{k,d,G} I_{k,d} \right) \text{OIB}_{i,d-1,G} + b2'_{d,G} \text{CONTROLS}_{i,d-1} + u2_{i,d}.$$

According to the 6 bins, we define the corresponding indicator variable $I_{k,d}$, which takes the value of 1 when the trade belongs to the kth bin, and 0 otherwise. We estimate equation (3) for each of the 3 investor groups, RT1–RT4, RT5, and INST, using the Fama–MacBeth regression. The first-stage coefficient $b1_{k,d,G}$ captures the predictive power for the kth combination on day d for investor group G, and the second-stage coefficient $b1_{k,G}$ (the mean of the first-stage time series of coefficient) measures the average predictive power for the kth combination for group G.

Panel B of Table 3 reports the estimation results. Our prior is that given that the smaller investors, RT1–RT4, predict returns negatively, they might be at a

[12]Exceptions include Boehmer, Sang, and Zhang (2020), which examines the trading patterns of retail investors following insider trading using retail trading trades identified from TAQ and Thomson Reuters Insiders data, and Van Kervel and Menkveld (2019), which examines the high-frequency trading around large institutional orders. Unlike the rest of the studies, these 2 papers can identify the counterparties of the trades in their sample, but their samples only cover limited subsets of the market.

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

TABLE 3

Predictive Patterns for Next-Day Returns with Different Counterparties

Table 3 reports the estimation results for the predictive patterns of next-day returns with different counterparties. For the counterparties, we put RT1 to RT4 together as 1 group and RT5 and INST as 2 separate groups. With the 3 groups of investors and 2 sides of trades (buy or sell), we divide the remaining trades into 6 bins: BBS, BSB, BSS, SBB, SBS, and SSB. The first letter indicates the trade direction of RT1–RT4, the second indicates the trade direction of RT5, and the third indicates INST. Panel A reports the coverage of each bin in the total sample. Panel B reports the coefficients estimated from Fama and MacBeth's (1973) regressions, as specified in equation (3). The control variables are the same as those in Table 2 and are not reported for abbreviations. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using the Newey and West (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Distribution of Counterparties

	RT1–RT4	RT5	INST	Proportions of Total Sample
BBS	Buy	Buy	Sell	17%
BSB	Buy	Sell	Buy	11%
BSS	Buy	Sell	Sell	21%
SBB	Sell	Buy	Buy	22%
SBS	Sell	Buy	Sell	13%
SSB	Sell	Sell	Buy	17%

Panel B. Predictive Patterns with Different Counterparties

		Dependent Variable: RET		
$b1_{k,G}$		OIB_RT1–OIB_RT4	OIB_RT5	OIB_INST
		I	II	III
OIB(−1) × BBS	Estimate	−0.0020***	−0.0007***	0.0021***
	[*t*-stat]	[−2.85]	[−3.21]	[11.39]
OIB(−1) × BSB	Estimate	−0.0041***	0.0004**	−0.0014***
	[*t*-stat]	[−3.24]	[2.26]	[−4.53]
OIB(−1) × BSS	Estimate	−0.0079***	0.0036***	0.0036***
	[*t*-stat]	[−12.10]	[13.91]	[15.11]
OIB(−1) × SBB	Estimate	−0.0145***	0.0032***	0.0019***
	[*t*-stat]	[−17.23]	[12.42]	[8.88]
OIB(−1) × SBS	Estimate	−0.0255***	0.0023***	−0.0013***
	[*t*-stat]	[−10.57]	[7.59]	[−3.33]
OIB(−1) × SSB	Estimate	−0.0093***	−0.0027***	0.0015***
	[*t*-stat]	[−13.20]	[−11.11]	[10.10]
Controls		Yes	Yes	Yes
Adjusted R^2		9.18%	8.41%	8.60%

disadvantage while trading against the large retail investors and the institutional investors. We focus on the order flows from RT1 to RT4 in the first specification. The coefficients on the order flows from RT1 to RT4 are consistently negative, implying that small retail investors' negative predictive power is prevalent and does not depend on counterparties. We notice that the coefficient magnitudes are relatively large when RT1–RT4 are on the opposite side of both RT5 (BSS) and INST (SBB), which supports our prior that RT1–RT4 might be at a disadvantage when trading against both RT5 and INST. In specification II, we focus on the predictive pattern of RT5 with different counterparties. The coefficients are mostly positive, consistent with RT5's positive predictive power for returns in general. Interestingly, the coefficients turn to negative for the cases of BBS and SSB, when RT5 is on the same side of RT1–RT4 but on the opposite side of INST. This suggests that when RT5 sides with smaller retail investors and trades against INST, their prediction becomes erroneous. Similar patterns also apply to INST in regression III, which has positive coefficients when INST trades on the same side of RT5, but the coefficients become negative when INST trades on the opposite side of RT5. That is, when RT5 and INST agree with each other and trade against the smaller retail investors, their prediction for future return is correct, but when RT5/INST agrees with the smaller retail investors, but disagree with each other, they have the incorrect signs for predicting future returns.[13]

C. Predictive Patterns with Different Holding Horizons

Another key variable for investment is the holding horizon. The statistics in Table 1 show that Chinese retail investors have high turnover ratios, which lead to short holding horizons. Does the predictive power of order flows change when holding horizons vary? Conventional wisdom suggests that shorter holding horizons are likely to be associated with excessive trading and possibly poorer returns. Due to data limitation, it is difficult for most existing studies to directly verify how holding horizons are associated with predictive power for a large sample. Taking advantage of our data, we compute holding horizons for each investor group and examine how holding horizons are related to predictive patterns of retail order flow.

The first step is to find a measure to capture holding horizons. Direct and accurate holding horizon calculation is quite challenging, as there are partial orders, cross trading, and we need to make assumptions regarding the in and out orders. To overcome this, we borrow the simple and intuitive days-to-cover ratio (DTCR) idea from the short-selling literature, which measures the number of days it takes for all short shares to be covered, as in Diether, Lee, and Werner (2009) and Boehmer and Wu (2013). That is, we define the holding period of average type G investors for stock i on day d as $\text{DTCR}_{i,d,G} = \frac{\text{TOTAL\_HOLDING\_SHARES}_{i,d,G}}{\text{DAILY\_SHARE\_VOLUME}_{i,d,G}}$, or the total shares held by investor group G divided by the daily shares traded on stock i for type G investors. For example, if 20% of the total shares are held by RT5, and the RT5 daily trading volume for this stock is 1%, then it takes 20 days for the entire holding stock by group G to be covered, and the DTCR for RT5 would be 20 days. Panel A of Table 4 reports the cross-sectional distribution of DTCRs for each investor group. Overall, the DTCRs are similar among RT1–RT4, with means of approximately 50 days, whereas RT5 and INST have longer horizons of approximately 200 days.

To examine how DTCR, a rough measure for holding horizons, is associated with the return predictive pattern, we simplify our analysis by selecting 3 benchmark horizons: 10, 20, and 60 days, which roughly correspond to the 25th, 50th, and 75th DTCRs, across different groups. Next, we separate all stock-day investor group observations, depending on the DTCR calculated from the previous day, into 4 bins: (0,10], (10,20], (20,60], and above 60 days. We modify equation (2) to create equation (4), which incorporates the differences in DTCR as follows:

$$(4) \qquad \text{RET}_{i,d} = c0_{d,G} + \left(\sum_{k=1}^{4} c1_{k,d,G} I_{k,d} \right) \text{OIB}_{i,d-1,G} + c2'_{d,G} \text{CONTROLS}_{i,d-1} + u3_{i,d}.$$

[13]We also form portfolios based on order imbalance and trading counterparties and present the results in Panel A of Appendix Table IV in the Supplementary Material. The results are qualitatively similar to those in Table 3.

TABLE 4

Predictive Patterns for Next-Day Return for Different Holding Horizons

Table 4 reports the estimation results on the predictive patterns for the next-day return for different holding horizons. We measure the holding days using the days-to-cover ratio (DTCR), defined as the total shares held by investor group *G* divided by daily shares traded on stock *i* for this investor group. Panel A reports the cross-sectional distribution of each investor group's holding horizons. Panel B reports the coefficients estimated from Fama and MacBeth's (1973) regressions, as specified in equation (4). We control the relative trading volume (RVOL), defined as the daily trading volume divided by the average daily trading volume from previous 49 days. Other control variables are the same as those in Table 2 and are not reported for brevity. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey and West's (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Distribution of Holding Horizon (Days-to-Cover Ratio)

	Mean	Std. Dev.	P5	P25	P50	P75	P95
RT1	57	691	4	11	24	49	116
RT2	43	402	4	10	20	36	74
RT3	46	458	4	12	22	38	73
RT4	64	833	4	13	26	48	114
RT5	251	4,479	4	17	41	95	420
INST	199	3,750	3	12	31	72	265

Panel B. Predictive Patterns for Different Holding Horizons

		Dependent Variable: RET					
		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
OIB(−1) × (0,10] days	Estimate	−0.0171***	−0.0191***	−0.0172***	−0.0023***	0.0056***	0.0017***
	[*t*-stat]	[−29.84]	[−28.77]	[−23.05]	[−4.28]	[13.68]	[12.51]
OIB(−1) × (10,20] days	Estimate	−0.0086***	−0.0078***	−0.0054***	−0.0006**	0.0015***	0.0012***
	[*t*-stat]	[−21.46]	[−19.74]	[−12.75]	[−2.43]	[7.07]	[9.62]
OIB(−1) × (20,60] days	Estimate	−0.0064***	−0.0052***	−0.0033***	−0.0004***	0.0007***	0.0013***
	[*t*-stat]	[−18.24]	[−15.42]	[−11.82]	[−2.81]	[6.37]	[11.49]
OIB(−1) × above 60 days	Estimate	−0.0074***	−0.0007	−0.0086*	0.0008	0.0002	0.0008***
	[*t*-stat]	[−6.26]	[−0.12]	[−1.72]	[0.35]	[1.52]	[7.80]
RVOL(−1)	Estimate	−0.0007***	−0.0008***	−0.0011***	−0.0014***	−0.0013***	−0.0013***
	[*t*-stat]	[−5.91]	[−6.97]	[−9.03]	[−11.31]	[−10.82]	[−10.88]
Controls		Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		10.03%	9.94%	9.72%	9.24%	9.31%	9.26%

Here, the indicator $I_{k,d}$ takes the value of 1 if the trade belongs to the *k*th bin of DTCR, and 0 otherwise. If conventional wisdom is correct, shorter holding horizons may be associated with poorer predictive power. Since the DTCR measure is typically correlated with trading volumes, which can have predictive power for returns by itself, as in Gervais, Kaniel, and Mingelgrin (2001), we include a volume variable as an additional control variable. In particular, we follow Gervais et al. (2001) and construct the relative trading volume (RVOL) measure, which is the trading volume on the stock day divided by the average daily stock trading volume from the previous 49 days.[14] Other control variables are the same as those in equation (2).

The estimates of $c1_{k,G}$ are presented in Panel B of Table 4. Take RT1 as an example. For the short horizon of less than 10 days, the coefficient is −0.0171 and highly significant. When the holding horizon becomes longer, the coefficient slightly decreases in magnitude to −0.0086 for 10 to 20 days, −0.0064 for 20 to 60 days, and −0.0074 for longer than 60 days. This decay shows that mistakes in the trading direction of the smallest retail investors slowly diminish when holding

[14]We thank our referee for this suggestion.

horizons become longer. The changes are more dramatic for RT2 to RT4, which begin with negative and significant coefficients for shorter horizons, and eventually become mostly insignificant for horizons longer than 60 days. These findings are consistent with the conventional wisdom that shorter holding horizons are likely associated with poorer returns, while longer horizons improve performances. Interestingly, for RT5, the patterns are different. For horizons shorter than 10 days, the $c1$ coefficient is 0.0056 and highly significant, but the $c1$ coefficient quickly reduces to 0.0015 (still highly significant) for holding horizons of 10 to 20 days and eventually becomes insignificant for holding horizons longer than 60 days. One possibility for the decreasing coefficients and statistical significance is that RT5 might trade on time-sensitive information, which is gradually incorporated into price, leading to the pattern of strong predictive power of shorter holding horizons than longer holding horizons.

The pattern for INST is somewhat similar to that for RT5, whereas the rate of reduction for the coefficients is much slower. In fact, for horizons shorter than 10 days, the $c1$ coefficient is 0.0017, and it slowly decreases to 0.0008 for holding horizons longer than 60 days. In this case, it is possible that the INST's order flow contains some time-sensitive information, which explains the reduction; it is also possible that the INST order flow contains information related to firm fundamentals, which takes longer for the information to be revealed and thus *slow* decay. Of course, most of these coefficients are positive and significant, indicating that, even for longer holding horizons, the order flows from RT5 and INST still contain price-relevant information in the right direction.[15]

D. Predicting Long-Term Stock Returns Using Retail Order Flows

Previous exercises focus on next-day return prediction, and it is natural to ask whether the predictive patterns continue for longer terms. If the predictive pattern quickly vanishes or reverses, the return predictability may be driven by short-term noise. If the predictive pattern persists over longer horizons, the return predictability is more likely to be linked to firm fundamentals or persistent biases. Therefore, we extend the Fama–MacBeth specification in equation (2) to create equation (5), where longer horizons are up to 12 weeks:

$$(5) \qquad \mathrm{RET}_{i,w} = d0_{w,G} + d1_{w,G}\mathrm{OIB}_{i,d-1,G} + d2'_{w,G}\mathrm{CONTROLS}_{i,d-1} + u4_{i,w},$$

that is, we use the previous day's order imbalance from group G, $\mathrm{OIB}_{i,d-1,G}$, to predict weekly returns over the next wth week. To be more specific, $\mathrm{RET}_{i,w}$ is calculated as a 5-day return from the beginning to the end of week w, where $w = 1, \ldots, 12$. For instance, when $w = 1$, $\mathrm{RET}_{i,w}$ is the cumulative return over days d to $d + 4$, and when $w = 12$, $\mathrm{RET}_{i,w}$ is the cumulative return over days $d + 55$ to $d + 59$. If order imbalances have only short-lived predictive power for future returns, we should observe the coefficient $d1$ quickly reverses. Alternatively, if

[15]As an alternative to the Fama–MacBeth regression, we construct double-sort portfolios based on order imbalance and holding horizons. The results are reported in Panel B of Appendix Table IV in the Supplementary Material and are qualitatively similar to those in Table 4.

TABLE 5

Predict Returns over the Next 12 Weeks

Table 5 reports the estimation results of whether trading activity by different investor groups can predict the cross section of returns over the next 12 weeks. We report the coefficient estimates from the Fama and MacBeth (1973) regressions specified in equation (5). The main independent variables are the previous day's order imbalance, OIB(−1). The control variables are the same as those in Table 2 and are not reported for brevity. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey and West's (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

*w*th Weeks		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
1	Estimate	−0.0226***	−0.0220***	−0.0144***	−0.0019***	0.0027***	0.0044***
	[*t*-stat]	[−18.84]	[−16.98]	[−13.43]	[−6.70]	[10.17]	[17.98]
2	Estimate	−0.0065***	−0.0060***	−0.0037***	0.0001	0.0010***	0.0012***
	[*t*-stat]	[−9.83]	[−8.56]	[−5.89]	[0.50]	[6.01]	[5.64]
3	Estimate	−0.0038***	−0.0031***	−0.0015**	0.0001	0.0007***	0.0007***
	[*t*-stat]	[−5.37]	[−4.09]	[−2.31]	[0.45]	[3.77]	[3.17]
4	Estimate	−0.0024***	−0.0021**	−0.0012	−0.0001	0.0007***	0.0005**
	[*t*-stat]	[−3.46]	[−2.57]	[−1.63]	[−0.37]	[3.94]	[2.09]
5	Estimate	−0.0014***	−0.0014**	−0.0011**	−0.0005**	0.0004**	0.0002
	[*t*-stat]	[−2.69]	[−2.44]	[−2.07]	[−2.00]	[2.25]	[1.39]
6	Estimate	−0.0029***	−0.0024***	−0.0016***	−0.0003	0.0001	0.0005***
	[*t*-stat]	[−4.69]	[−3.55]	[−2.71]	[−1.22]	[0.72]	[2.66]
7	Estimate	−0.0027***	−0.0025***	−0.0018***	−0.0002	0.0001	0.0007***
	[*t*-stat]	[−5.02]	[−4.43]	[−3.65]	[−0.68]	[0.39]	[4.25]
8	Estimate	−0.0015**	−0.0010	−0.0007	−0.0002	0.0003**	0.0004**
	[*t*-stat]	[−2.57]	[−1.52]	[−1.08]	[−0.73]	[2.04]	[2.39]
9	Estimate	−0.0010*	−0.0005	−0.0004	0.0002	0.0004**	0.0001
	[*t*-stat]	[−1.94]	[−1.00]	[−0.73]	[1.03]	[2.43]	[0.43]
10	Estimate	−0.0007	−0.0004	−0.0004	0.0002	−0.0001	0.0002
	[*t*-stat]	[−1.22]	[−0.60]	[−0.81]	[0.94]	[−0.44]	[0.90]
11	Estimate	−0.0006	−0.0007	−0.0001	0.0000	−0.0002	0.0002
	[*t*-stat]	[−1.02]	[−1.01]	[−0.12]	[−0.04]	[−1.10]	[1.10]
12	Estimate	−0.0005	−0.0001	0.0005	0.0001	0.0001	0.0000
	[*t*-stat]	[−0.90]	[−0.12]	[0.76]	[0.41]	[0.39]	[0.05]

the specified retail order imbalance has a longer predictive power, the coefficient $d1$ may slowly decrease, rather than quickly reverse.

We present the estimates of coefficient $d1$ in equation (5) in Table 5. For the smallest retail investor RT1, the coefficients on OIB_RT1 are negative and monotonically increase from −0.0226 at week 1 to −0.0005 at week 12, while the coefficients are statistically significant until week 9. Reverse patterns are not observed, which implies that the predictive power is long-lasting. Similar patterns are observed for OIB_RT2 and OIB_RT3. For OIB_RT4, the coefficient is −0.0019 and highly significant at week 1 but quickly reverses and becomes insignificant, indicating that its predictive power for future returns might be temporary. The positive predictive power of OIB_RT5 and OIB_INST also significantly persists for approximately 8–9 weeks, and there are no significant reversals. The general persistence of the positive predictive pattern for RT5 and INST indicates that the predictive power is likely rooted in information related to fundamentals or against persistent noise/behavioral biases.[16]

[16]We also consider an alternative specification for the long-term predictability. To match the weekly returns on the left-hand side, we use the previous week's order imbalance $OIB_{i,w-1,G}$, to predict future

IV. What Drives the Order Imbalance Predictive Power for Future Returns?

A. Alternative Hypotheses for Explaining Retail Order Flow and Its Return Predictive Power

Given the large differences in the predictive power of future returns for different investor groups' order flows, it is important to understand the driving forces behind these order flows. Previous literature provides several hypotheses for explaining investor order flows, which may help to explain the heterogeneous predictive patterns of different investor groups for future returns. Below, we outline 4 main hypotheses, and these hypotheses are not mutually exclusive.

First, Chordia and Subrahmanyam (2004) state that order flows tend to be persistent and that persistent buying/selling pressure could directly lead to the predictability of future returns. Here, we adopt the previous day's order imbalance measure, $\text{OIB}_{i,d-1,G}$, as the proxy for order persistence.

Second, Kaniel et al. (2008) argue that retail traders in the U.S. are mostly contrarian, providing liquidity to the market and afterward receiving a positive premium for liquidity provision. Following this logic, if the retail trades are momentum trades that demand liquidity, it is then possible that momentum trades will have a negative relationship with future returns, and if the retail trades are contrarian and provide liquidity, these contrarian trades will have a positive relationship with future returns. For this hypothesis, we choose returns from the previous day, week, and month as proxies for momentum or contrarian trading.

Third, Liu et al. (2021) connect retail trading motives to behavioral biases and find that overconfidence and gambling preferences are the 2 dominant behavioral biases that affect the trade of Chinese retail investors. For the overconfidence measure, we follow Barber et al. (2009) and Liu et al. (2021) and proxy it with the average of the daily investor group's stock turnover from the previous 20 days, $\text{OVERCONF}_{i,d-1,G}$.[17] For gambling preferences at the stock level, $\text{GAMBLE}_{i,d-1}$, we follow Bali, Cakici, and Whitelaw (2011) and compute the maximum daily returns from the previous 20 days.[18]

Finally, Kelley and Tetlock (2013) find that retail investors, especially the aggressive ones, may have valuable information about fundamental firm news, and thus, their trading could correctly predict the direction of future returns. We measure firm-level fundamental news by the cumulative abnormal returns (CAR) over the

---

weekly returns. The results reported in Appendix Table V in the Supplementary Material are similar to those in Table 5.

[17]For the overconfidence proxy, Barber and Odean (2000) use each household's portfolio's turnover. Our data have a different structure and cannot be used to construct household-level proxies. Considering our study focuses on investor groups rather than individual households, we carry the spirit from the previous literature and use the stock-level turnover from each investor group.

[18]An alternative measure for gambling preference proxy is introduced by Liu et al. (2021), which relies on events when the stock return hits a 10% price limit. However, the 10% price limit only accounts for 0.07% of the total sample and is not suitable for our purposes using on daily stock returns. Thus, we choose the maximum daily return as our main gambling measure. We also consider other alternative proxies for gambling preferences, such as idiosyncratic volatility and skewness. These proxies deliver similar results to those using maximum daily returns and are available from the authors.

earnings announcement period. Unlike the proxies for order persistence, liquidity provision, and behavioral biases, which can be computed for each stock on each day, the news proxies are only available on earnings news days, which account for 1.58% of stock days, rendering our 2-stage estimation (to be introduced in Section IV.B.) imprecise. To cope with this missing data issue for the news hypothesis, we first consider the order persistence, liquidity provision, and behavioral bias hypotheses in Section IV.B and focus on the news hypothesis using an event-day approach in Section IV.C.

B. A 2-Stage Decomposition for Order Flows' Return Predictive Power

To find out whether the above hypotheses explain the trading behavior of different retail investor groups and their predictive power for future stock returns, we adopt a 2-stage decomposition method as in Boehmer et al. (2021).[19] For the first stage, we use the above hypotheses to explain the retail flow measures to find out which ones are important drivers for the order flows. This step also helps to decompose the retail order flows into hypothesis-implied components for each hypothesis. For the second stage, we investigate which of the hypothesis-implied components contributes to the predictive pattern of different investor order flow measures.

Panel A of Table 6 reports the first-stage estimation results. In the first row, the coefficients of the lagged order flow variables are always positive and significant, indicating that order persistence is an important driver of order flow. For the next 3 rows, order flows are connected with returns from the previous day, week, and month. The order imbalances of RT1, RT2, and RT3 load positively and significantly on the previous day's return, indicating that these investors buy more if the previous day's return is positive and sell more if the previous day's return is negative. This corresponds to a daily momentum trading strategy, which demands immediate liquidity. For larger retail investors in RT4 and RT5, order imbalances load negatively and significantly on returns from the previous day, indicating that they are contrarian investors who buy low, sell high, and possibly provide immediate liquidity. If we extend the horizon to the previous 1 week or 1 month, the coefficients on all returns are negative and significant, indicating that all retail investors become contrarian and buy losers and sell winners over the longer term.[20]

[19]A step-by-step description of the 2-stage decomposition is provided in Appendix A in the Supplementary Material.

[20]Our finding that large retail investors are contrarian and smaller ones are momentum traders over daily horizon differs from some previous studies. For instance, contrarian patterns are documented in Kaniel et al. (2008) using monthly horizons in the U.S. and Barrot, Kaniel, and Sraer (2016) using daily and weekly horizons in France. Using U.S. data, Kelley and Tetlock (2013) and Boehmer et al. (2021) both find that retail trades follow momentum over daily horizons but are contrarian at weekly horizons. In our setting, we find the trading patterns from investors with smaller account sizes are similar to those in Kelley and Tetlock (2013) and Boehmer et al. (2021), while investors with the largest account sizes behave similarly to the patterns documented in Kaniel et al. (2008) and Barrot et al. (2016). We also examine whether the momentum/contrarian patterns are potentially related to average holding periods of stock. Interacting the return of the previous day with holding horizons, we find small retail investors' momentum is stronger for short holding periods, while large retail investors' contrarian is similar for both short and long holding periods. The results are available from the authors.

TABLE 6

Two-Stage Decomposition for Understanding the Predictive Patterns of Retail Order Flows

Table 6 reports the estimation results of the decomposition of the predictive power of different investor groups' order imbalances for the cross section of future stock returns. We estimate 2-stage Fama and MacBeth (1973) regressions. Panel A reports the first-stage estimation results, where the order imbalance measures are decomposed into 5 components as specified in equation (A1) of the Supplementary Material. Variable OVERCONF(−1) is measured as the corresponding investor group's average turnover on the stock from the previous 20 days as a proxy for each investor's overconfidence. The variable GAMBLE(−1) is the maximum daily return from the previous 20 days as a proxy for gambling preference. Panel B reports the second-stage decomposition of the order imbalance's predictive power, as specified in equations (A2) and (A3). As an independent variable, the variable OIB(−1, PERSISTENCE) is estimated in the first stage using the past order imbalance and reflects price pressure. The variable OIB(−1, LIQUIDITY) is estimated in the first stage using past returns over different horizons and is connected to the liquidity provision or liquidity demand hypothesis. The variable OIB (−1, OVERCONF) is estimated in the first stage, reflecting overconfidence. The variable OIB(−1, GAMBLE) is estimated in the first stage using the maximum daily returns from the previous 20 days and reflects a preference for gambling. The residual part of the previous day's order imbalance from the first-stage estimation is denoted "other," which can be attributed to private information about future returns. The control variables are the same as those in Table 2 and are not reported for brevity. For each regression, we also report the difference in predicted day-ahead returns for observations at the 2 ends of the interquartile range (INTERQUARTILE_RETURN) in Panel B. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey and West's (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A. First Stage of Projecting the Order Imbalance on Persistence, Past Returns, Overconfidence, and Gambling Proxies

		Dependent Variable				
		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5
OIB(−1)	Estimate	0.1870***	0.1967***	0.1711***	0.0499***	0.1036***
	[*t*-stat]	[24.24]	[32.92]	[35.03]	[15.03]	[27.85]
RET(−1)	Estimate	0.5159***	0.7269***	0.4482***	−0.2205***	−1.2968***
	[*t*-stat]	[9.17]	[16.36]	[14.52]	[−7.57]	[−28.84]
RET(−6,−2)	Estimate	−0.4214***	−0.2196***	−0.1063***	−0.0804***	−0.0485***
	[*t*-stat]	[−19.83]	[−15.14]	[−7.71]	[−4.22]	[−2.74]
RET(−27,−7)	Estimate	−0.0350***	−0.0196***	−0.0235***	−0.0399***	−0.0203***
	[*t*-stat]	[−6.49]	[−4.89]	[−6.72]	[−8.36]	[−3.74]
OVERCONF(−1)	Estimate	0.0894***	0.0611***	0.0657***	0.0418***	−0.0881***
	[*t*-stat]	[4.53]	[4.57]	[7.28]	[3.64]	[−8.80]
GAMBLE(−1)	Estimate	0.0330	0.0784***	0.1731***	0.2423***	−0.0671**
	[*t*-stat]	[1.49]	[5.06]	[11.69]	[11.90]	[−2.53]
INTERCEPT	Estimate	−0.0214***	−0.0120***	−0.0115***	−0.0078***	0.0231***
	[*t*-stat]	[−5.70]	[−5.09]	[−7.63]	[−4.00]	[9.41]
Adjusted R^2		7.08%	5.60%	3.89%	0.73%	2.00%

Panel B. Second-Stage Decomposition of Order Imbalance's Predictive Power

		Dependent Variable: RET				
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5
OIB(−1, PERSISTENCE)	Estimate	−0.0301***	−0.0276***	−0.0205***	−0.0086***	0.0060***
	[*t*-stat]	[−16.73]	[−15.39]	[−12.22]	[−4.37]	[7.87]
OIB(−1, LIQUIDITY)	Estimate	−0.0135***	−0.0205***	−0.0261***	0.0140	0.0042*
	[*t*-stat]	[−4.48]	[−6.18]	[−4.69]	[0.69]	[1.69]
OIB(−1, OVERCONF)	Estimate	−0.1006***	−0.2158***	−0.2413***	−0.3730***	0.1134***
	[*t*-stat]	[−4.00]	[−5.97]	[−6.05]	[−5.85]	[5.90]
OIB(−1, GAMBLE)	Estimate	−0.2877***	−0.0981***	−0.0474***	−0.0389***	0.1918***
	[*t*-stat]	[−4.89]	[−4.08]	[−4.35]	[−4.98]	[6.35]
OIB(−1, OTHER)	Estimate	−0.0086***	−0.0084***	−0.0060***	−0.0008***	0.0011***
	[*t*-stat]	[−15.54]	[−15.51]	[−14.16]	[−6.19]	[9.59]
Adjusted R^2		10.50%	10.35%	10.07%	9.59%	9.52%
INTERQUARTILE_RETURN						
OIB(−1, PERSISTENCE)		−0.1177%	−0.0962%	−0.0591%	−0.0125%	0.0281%
OIB(−1, LIQUIDITY)		−0.0290%	−0.0351%	−0.0261%	0.0091%	0.0097%
OIB(−1, OVERCONF)		−0.0337%	−0.0475%	−0.0484%	−0.0428%	0.0257%
OIB(−1, GAMBLE)		−0.0314%	−0.0254%	−0.0271%	−0.0312%	0.0425%
OIB(−1, OTHER)		−0.1778%	−0.1493%	−0.1008%	−0.0233%	0.0490%

The next 2 rows present the results on how behavioral biases are related to order flows. The coefficients on the overconfidence proxy are all positive and significant for RT1–RT4, indicating that overconfidence might be a strong driver of trading by these retail investors. For the largest retail group, RT5, the coefficient becomes −0.0881, with a significant t-statistic of −8.80. In other words, the largest retail investors' trades are in the opposite direction of the overconfidence proxy. In terms of gambling preference, for RT2–RT4, the coefficients are always positive and significant, indicating that these retail investors prefer to buy stocks with lottery features.[21] When we move on to RT5, the coefficient is −0.0671 with a significant t-statistic of −2.53, which indicates that the largest retail investors' trades are in the opposite direction of the gambling motive.

We report the second-stage decomposition results in Panel B of Table 6. We consider the first retail group, RT1, as an example. The coefficient estimate on OIB(PERSISTENCE) is −0.0301, with a t-statistic of −16.73, which implies that order persistence significantly and negatively contributes to the predictive power of the RT1 trading flow. The coefficient estimate on OIB(LIQUIDITY) is −0.0135, with a t-statistic of −4.48, which suggests that daily momentum trading probably significantly and negatively contributes to the predictive power of the RT1 trading flow. The coefficient of OIB(OVERCONF) is −0.1006, with a t-statistic of −4.00, and the coefficient of OIB(GAMBLE) is −0.2877, with a t-statistic of −4.89. These 2 significant coefficients imply that the 2 behavioral biases significantly and negatively contribute to the predictive power of RT1 trading flow. For the OIB(OTHER) component, the coefficient is −0.0086, with a significant t-statistic of −15.54, indicating that there is residual information other than those incorporated in the 3 hypotheses, which significantly contributes to RT1's negative predictive pattern for future returns. In terms of economic magnitude, we compute the interquartile range of all 5 components of the order imbalance measure. For the smallest retail group RT1, if we move from the 25th percentile to the 75th percentile in the distribution, the interquartile differences in future 1-day stock returns are −0.1177%, −0.0290%, −0.0337%, −0.0314%, and −0.1778% for OIB(PERSISTENCE), OIB(LIQUIDITY), OIB(OVERCONF), OIB(GAMBLE), and OIB(OTHER), respectively. In other words, order persistence, liquidity demand, overconfidence, and gambling preferences all contribute to the negative predictive power of RT1 for next-day returns, with the first term having the largest magnitude. Similar patterns are observed for other smaller retail investor groups RT2–RT4.

Regarding the largest retail investors, RT5, the patterns are quite different. In terms of the coefficient estimates, we find that order persistence, overconfidence, gambling preferences, and others are all positive and significant. In terms of economic magnitude, if we move from the 25th percentile to the 75th percentile in the distribution, the interquartile differences in future 1-day stock returns are 0.0281%, 0.0097%, 0.0257%, 0.0425%, and 0.0490% for OIB(PERSISTENCE), OIB(LIQUIDITY), OIB(OVERCONF), OIB(GAMBLE), and OIB(OTHER), respectively. This indicates that order persistence, liquidity provision, and trading against behavioral biases all contribute to RT5's positive predictive pattern for future returns.

[21]The increasing coefficients from RT1 to RT4 should not be interpreted as RT4's gambling preference being stronger than RT1, because these variables are not standardized and cannot be compared directly. If we standardize them and make them comparable across investor groups, RT3 has the strongest gambling preference. The results are available from the authors.

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

Overall, our decomposition exercise shows that a substantial part of the negative predictive power of retail investors with smaller account sizes comes from order persistence, liquidity demand, and behavioral biases, while the positive predictive power of retail investors with larger account balances mostly comes from order persistence and trading against overconfidence and gambling preferences.[22] Across all investor groups, the significance and large magnitude of the "other" component indicate that existing hypotheses cannot fully explain trading behaviors and their predictive power for returns. Then, what does "other" stand for? One possibility is information, which we take a close look at in the next subsection.

C. A Close Look at the Information Channel

It is important to understand how various retail investors participate in the information discovery process. As mentioned earlier, the most influential information at the firm level is earnings news; hence, we follow Kelley and Tetlock (2013) and measure firm-level information by the CAR over the earnings announcement period. Notice that earnings news only happens quarterly rather than daily, so the daily Fama and MacBeth (1973) estimation adopted for the 2-stage estimation might not be proper for understanding how Chinese retail investors process information. As an alternative, in this section, we focus on event days to study this issue. We capture each retail investor group's participation in the information discovery process in 3 steps.

First, we examine whether different retail investors can predict earnings news the next day. A positive answer indicates that these investors anticipate the information before it becomes public, either because they have access to private information or because they have better skills in prediction. We estimate the following cross-sectional specification for each quarter q:[23]

$$\text{CAR}_{i,d-1,d} = e0_q + e1_q \text{OIB}_{i,d-1} + e2_q' \text{CONTROLS}_{i,q-1} + u5_{i,q}. \tag{6}$$

Assuming that the earnings announcement day is day d, we compute the cumulative returns over day $d-1$ and day d and subtract the market returns over the same period to obtain CAR for each stock, $\text{CAR}_{i,d-1,d}$.[24] The main predictive variable on the right-hand side of the equation is the order imbalance measure from day d-2. Notice that each firm has only 1 earnings day per quarter, and equation (6) is estimated for each quarter in the cross section to ensure that we cover all firms in each quarter. Statistical inferences are based on the quarterly time series of the estimated coefficients, and standard errors are computed using Newey and West's (1987) method, with the optimal bandwidths chosen by using the Newey and West (1994) approach. If retail order flows can predict earnings surprises in the right direction, coefficient $e1$ should be significantly positive and vice versa. Panel A of Table 7 presents the estimation results. For retail investors RT1–RT3, the

[22] In Appendix Table VI in the Supplementary Material, we examine the robustness of the 2-stage decomposition by adding additional 2 weeks lag of order flow in the first stage, construct the new "persistence" component, and reestimate the second-stage decomposition. The results are robust.

[23] All estimations in this study are estimated within each investor group. Beginning in this section, we omit the subscripts G to make the formula more readable.

[24] We also examine wider windows such as CAR(−1,1) and CAR(−3,3), and the results are similar and available from the authors.

coefficients $e1$ are −0.0251, −0.0234, and −0.0166, respectively, with highly significant t-statistics. These negative and significant coefficients indicate that these investors incorrectly predict earnings surprises. Coefficient $e1$ for RT4 is close to 0 and insignificant. In contrast, the coefficient $e1$ for RT5 is 0.0023, positive and statistically significant, implying that these investors can correctly predict future earnings surprises.

Second, we examine whether different retail groups can process contemporaneous public news. Here, the dependent variable is retail order flow, $\text{OIB}_{i,d}$, which we connect to contemporaneous earnings news, $\text{CAR}_{i,d-1,d}$:

$$\text{OIB}_{i,d} = f0_q + f1_q\text{CAR}_{i,d-1,d} + f2_q'\text{CONTROLS}_{i,d-1} + u6_{i,q}. \tag{7}$$

If a particular type of retail order imbalance can process contemporaneous and public earnings news in the right direction, the associated coefficient $f1$ is significantly positive and vice versa. Panel B of Table 7 reports these results. For retail

TABLE 7

A Closer Look at the Relationship Between Investor Order Flows and Earnings News

Table 7 reports the estimation results on the relationship between different investor groups' order flows and public earnings announcements. Panel A reports whether investor order flow can predict earnings surprises, as specified in equation (6). The dependent variable is the cumulative abnormal return from day $d-1$ to day d, CAR[−1,0], and the independent variable is order imbalance from day d-2, OIB(−2). Panel B reports whether trades from different retail groups can process contemporaneous news, as specified in equation (7). The dependent variable is the order imbalance OIB(0), and the independent variable is the cumulative abnormal return from day $d-1$ to day d, CAR[−1,0]. Panel C reports the effect of earnings news days on the return predictability of different investor group trades, as specified in equation (8). The dependent variable is the return on day d, and the independent variables include the previous day's order imbalance OIB(−1), the news dummy EVENT(−1), and the interaction terms OIB(−1) × EVENT(−1). The EVENT(−1) dummy equals 1 if there is an earnings announcement for that firm day, and 0 otherwise. The other control variables are the same as those in Table 2, and these coefficients are not reported. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey and West's (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Investor Order Flow Predicting Future Earnings Announcement News Events

		Dependent Variable: CAR[−1,0]				
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5
OIB(−2)	Estimate	−0.0251***	−0.0234***	−0.0166***	−0.0003	0.0023***
	[t-stat]	[−8.12]	[−5.81]	[−4.56]	[−0.50]	[3.67]
Adjusted R^2		6.33%	5.98%	5.57%	5.19%	5.15%

Panel B. Investor Order Flow Regressed on Contemporaneous Earnings Announcement News Events

		Dependent Variable				
		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5
CAR[−1,0]	Estimate	−1.9225***	−1.8291***	−1.4349***	−0.8781***	0.1583**
	[t-stat]	[−22.65]	[−40.09]	[−14.53]	[−5.40]	[2.12]
Adjusted R^2		13.66%	14.60%	10.14%	1.85%	0.60%

Panel C. Return Predictive Power of Investor Order Flow Interacted with Earnings Announcement News Events

		Dependent Variable: RET				
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5
OIB(−1)	Estimate	−0.0079***	−0.0070***	−0.0038***	−0.0008*	0.0005**
	[t-stat]	[−9.33]	[−7.67]	[−5.55]	[−1.93]	[3.08]
OIB(−1) × EVENT(−1)	Estimate	−0.0080***	−0.0093***	−0.0071**	−0.0007	0.0014**
	[t-stat]	[−4.22]	[−3.90]	[−2.52]	[−1.54]	[2.07]
EVENT(−1)	Estimate	0.0011***	0.0008****	0.0006***	0.0005**	0.0005**
	[t-stat]	[3.94]	[2.89]	[2.80]	[2.55]	[2.40]
Adjusted R^2		7.36%	7.16%	6.82%	6.38%	6.35%

investors RT1–RT4, the coefficient $f1$ are -1.9225, -1.8291, -1.4349, and -0.8781, respectively, with highly significant t-statistics. These negative and significant coefficients indicate that these retail investor groups process contemporaneous public earnings news in the wrong direction. By contrast, the coefficient $f1$ for RT5 is 0.1583, with a t-statistic of 2.12, implying that RT5 might be able to correctly process contemporaneous public earnings news.

Third, we examine whether the predictive power of retail order flows for future returns improves or deteriorates on event days to understand how much the information hypothesis helps explain the return predictive patterns observed in Section III. Here, we add the event-day dummy and interaction term:

$$\text{RET}_{i,d} = g0_d + (g1_d + g2_d\text{EVENT}_{i,d-1})\text{OIB}_{i,d-1} + g3_d\text{EVENT}_{i,d-1} + g4'_d\text{CONTROLS}_{i,d-1} + u7_{i,d}. \quad (8)$$

The event dummy, $\text{EVENT}_{i,d-1}$, is equal to 1 if firm i has earnings news on day $d-1$, and 0 otherwise. For non-news days, the predictive power of retail trades is measured by the coefficient $g1$; for news days, the predictive power is measured by $(g1 + g2)$. If coefficient $g2$ is significantly different from 0, the group of retail investors anticipates future stock returns differently on these news days. In the U.S., firm earnings announcements are chosen by firms and are scattered throughout the year. In China, all firms are required to report their financial statements to regulators before the 4 preset deadline dates each year. Consequently, firms mostly announce their earnings within a short period before these deadline dates, and there would be 0 announcements outside these short periods. To ensure that we have enough observations to estimate the Fama and Macbeth (1973) coefficients in equation (8), we only include days with at least 5% of the total number of firms with earnings announcements, which gives us 68 days, or 8% of the total days in our sample.

Panel C of Table 7 present the results. We consider the smallest retail investor, RT1, as an example. The coefficient on order imbalance, $g1$, is -0.0079 and statistically significant, indicating that, on average, the trades from RT1 negatively predict future returns. When there is earnings announcement news, the coefficient on the interaction of the event dummy and the order imbalance is -0.0080 and is statistically significant, implying that the negative prediction of RT1 for future stock returns doubles on earnings news days. This is consistent with our earlier finding that smaller retail investors fail to predict and process earnings news, which leads to more negative predictions of returns on event days. Similar patterns are observed in RT2, RT3, and RT4. For the largest retail investor, RT5, the coefficients $g1$ and $g2$ are 0.0005 and 0.0014, respectively, both of which are statistically significant. In other words, large retail investors' predictive power for future returns quadruples on earnings news days, possibly because these retail investors can correctly predict and process earnings news, which enhances their ability to predict future stock returns.

Overall, our results reveal interesting heterogeneous patterns in how retail investors predict and process public information. Smaller retail investors are unable to predict future news and lack the skills to correctly process public news, while the largest retail investors and institutions can correctly anticipate future earnings news and incorporate contemporaneous news into their trading. The differences in the information processing abilities of different retail investors clearly contribute to the differences in their predictive power for future returns.[25]

V. Trading Performances Using Tracking Portfolios of Retail Trading

A. The BLLO Method

With the short- and long-term predictive patterns documented in the previous sections, it is natural to ask how good the trading performance would be if we tracked the order flows from different groups of retail investors. To answer this question, we follow Barber et al. ((2009), BLLO hereafter), who design a simple and intuitive method to compute the trading performances of different types of investors in the Taiwan Stock Exchange over 1995 to 1999. A step-by-step description of BLLO's method is provided in Appendix B of the Supplementary Material. Here, we provide a brief description of BLLO's method, which has 4 steps.

First, BLLO separate all investors into "individuals," "corporations," "dealers," "foreigners," and "mutual funds," using identities provided by the exchange. Second, within each investor group, BLLO compute the aggregate daily net buy and net sell positions and create daily matching trading portfolios, the "net buy" and "net sell" portfolios. Third, BLLO track these "net buy" and "net sell" portfolios over a holding horizon of n days, compute cumulative cash flows and returns from this tracking strategy, and treat them as proxies for trading performances of different investor groups. After computing the total performance, BLLO decompose the total into 3 intuitive components, stock selection, market timing, and trading cost.

We adopt the BLLO method in our study for 3 reasons. First, the data structure of our sample is quite similar to that of BLLO's method, therefore, it is relatively straightforward to apply to our data. Second, the BLLO method is simple and easy to interpret, and it allows us to compare our results with their results. Third, the decomposition of total performance into stock selection, market timing, and trading cost helps us to connect our earlier results on the predictive power of order flows in the cross section. Notice that the performance estimates from the BLLO method are not realized gains and losses from investors. Rather, they are estimates of a trading strategy that follows signals from retail trading, namely the net positions of each investor group, which reflects the investment skills of different groups of investors.

We make 3 parameter choices within the BLLO framework. First, BLLO assume that the net buy and sell portfolios are held for n days and sets n to different horizons. As the main discussion of BLLO focuses on the holding horizon of 140 days (n = 140), we choose the same horizon for ease of comparison. Second, BLLO assume that the net buy (sell) portfolios are independent for each day d and annualize the performance by aggregating the daily net buy and sell portfolios over each year. Here, we relax the independent assumption and compute the standard error estimates using the Newey and West (1987) method, with the optimal

[25]As an alternative, we use the Financial News Database of Chinese Listed Companies (CFND) to investigate whether the results from earnings news can be extended to other news. The results confirm our findings in Section IV and are available from the authors.

bandwidths chosen by using the Newey and West (1994) procedure.[26] Third, to have a perspective of total performance as a percentage of the total investment of these tracking portfolios, which is similar to the idea of return on investments, BLLO measure total investment as the aggregate holding value for each group of investors. For the ease of comparison, we choose the aggregate holding values in Panel A of Table 1 and present the return on investment as the total performance over aggregate holding.

B. Performances of Chinese Investors' Trading Using the BLLO Method

Panel A of Table 8 reports the performance of the tracking portfolios. The annualized return to investment ranges between −5.61% and −2.80% for a 140-day holding horizon for RT1 to RT4, and RT5 has a close to 0, −0.29%, return to investment, while the institutions have a return to investment of 1.15%. It might not be too surprising to find that BLLO's net buy- and sell-tracking portfolio approach provides negative returns on investments for RT1–RT4, given that our earlier results show that RT1–RT4 predicts future returns negatively and significantly. Similarly, positive returns on investments for INST are also expected because earlier evidence shows that institutional order flows predict future returns positively and significantly. It is intriguing to find that the net buy and sell portfolios tracking the order flows of RT5 deliver negative returns on investments, while earlier results show that RT5 predicts future returns positively and significantly.

The answer is in the next 3 columns, which provide the 3 components of total performance. For RT5's annualized return on investment of −0.29%, the stock selection component contributes to a significant and positive coefficient of 1.05%, which is consistent with the earlier positive cross-sectional return predictive power; the market timing component contributes an insignificant coefficient of −0.30%, indicating that RT5 might not be able to time the market; and the trading cost component contributes a significant coefficient of −1.03%. That is, even though RT5 is capable of stock selection, its low market timing ability and significant transaction cost cause the tracking portfolio to have a close-to-zero total return on investment.

The decomposition of RT1–RT4 is different from that of RT5 in the sense that their groups of investors do not possess either stock selection or market timing ability, and so all 3 components of total returns on investment are all negative, mostly attributable to poor stock selection and significant trading cost. For instance, the trading cost component ranges between −1.38% (RT1) and −1.80% (RT2), which suggests that RT2 pays the most for trading relative to their holding. In the case of institutional investors, the results are quite similar to RT5, except that they have superior stock selection ability (STK_SELECT = 1.38%), market timing ability (MKT_TIMING = 0.19%), and lower trading costs (TRD_COST = −0.41%), and the total return on investment is positive and significant.[27]

[26]We conduct robustness check using optimal lag numbers from the Bayesian information criterion (BIC). The results are similar and available from the authors.

[27]We compare Panel A of Table 8 with BLLO's results for the Taiwan stock market and find that the general patterns are similar. In Appendix Figure II in the Supplementary Material, we also conduct a similar exercise for shorter holding periods. The results display similar patterns as in Panel A of Table 6.

TABLE 8

Annual Performance Using Tracking Portfolios of Retail Trading, from Stock Selection and Market Timing

Table 8 reports the annual performance for different retail investor groups using Barber et al.'s (2009) approach between Jan. 2016 and June 2019. The net buy and sell portfolios are constructed each day, as specified in equation (A4) in the Supplementary Material, and the average construction prices are defined in equation (A6). The daily portfolio's total performance of holding n days is the portfolio cumulative excess return minus the trading cost, as specified in equation (A5), and we choose n equals 140 days following Barber et al. (2009). We then annualize the daily portfolio's total performance by timing the 240 trading days of each year and dividing by aggregate holding value in Panel A of Table 1 to present the investor group's annual performance in percentage. The total performance could be decomposed into stock selection, market timing, and trading cost, as specified in equations (A7)–(A10). Panel A presents the total annual performance. Panel B reports annual performance conditional on the momentum and contrarian. If the investor group G's order imbalance times the previous day's return, $\mathrm{OIB}(i,d,G) \times \mathrm{RET}(i,d-1) > 0$, we classify the stock i on day d for investor group G to be momentum; if $\mathrm{OIB}(i,d,G) \times \mathrm{RET}(i,d-1) \leq 0$, we classify the stock i on day d for investor group G to be contrarian. Panel C reports the annual performance of the earnings announcements. Earnings announcement events include the stock on announcement day d and the previous day $d-1$. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey and West's (1987) method, with optimal bandwidths chosen by following Newey and West (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Annual Trading Performance Using Tracking Portfolios

		Total	Stock Selection	Market Timing	Trading Cost
RT1	Performance	−5.61%***	−4.03%***	−0.19%	−1.38%***
	[*t*-stat]	[−10.29]	[−8.39]	[−0.56]	[−23.47]
RT2	Performance	−4.61%***	−2.55%***	−0.27%	−1.80%***
	[*t*-stat]	[−8.96]	[−6.21]	[−0.93]	[−23.07]
RT3	Performance	−3.60%***	−1.60%***	−0.29%*	−1.72%***
	[*t*-stat]	[−9.49]	[−5.51]	[−1.71]	[−20.89]
RT4	Performance	−2.80%***	−0.90%***	−0.33%*	−1.58%***
	[*t*-stat]	[−8.51]	[−3.98]	[−1.78]	[−19.53]
RT5	Performance	−0.29%	1.05%***	−0.30%	−1.03%***
	[*t*-stat]	[−0.92]	[5.27]	[−1.39]	[−18.84]
INST	Performance	1.15%***	1.38%***	0.19%	−0.41%***
	[*t*-stat]	[3.57]	[4.85]	[0.93]	[−20.49]

Panel B. Annual Trading Performance Using Tracking Portfolios: Separate Momentum and Contrarian Trading

		Total		Stock Selection		Market Timing		Trading Cost	
		Contrarian	Momentum	Contrarian	Momentum	Contrarian	Momentum	Contrarian	Momentum
RT1	Performance	−1.04%**	−4.57%***	0.07%	−4.10%***	−0.41%	0.22%	−0.70%***	−0.68%***
	[*t*-stat]	[−2.57]	[−8.47]	[0.25]	[−8.51]	[−1.39]	[1.03]	[−23.33]	[−21.29]
RT2	Performance	−0.95%**	−3.67%***	0.43%	−2.97%***	−0.46%*	0.19%	−0.91%***	−0.89%***
	[*t*-stat]	[−2.29]	[−7.88]	[1.46]	[−7.63]	[−1.79]	[1.12]	[−23.61]	[−21.15]
RT3	Performance	−1.14%***	−2.47%***	0.17%	−1.77%***	−0.42%**	0.14%	−0.88%***	−0.83%***
	[*t*-stat]	[−4.04]	[−8.00]	[0.85]	[−7.48]	[−2.47]	[1.02]	[−21.06]	[−19.65]
RT4	Performance	−0.95%***	−1.85%***	0.33%**	−1.23%***	−0.44%***	0.12%	−0.84%***	−0.74%***
	[*t*-stat]	[−4.09]	[−7.23]	[2.01]	[−6.66]	[−3.32]	[0.65]	[−19.59]	[−18.95]
RT5	Performance	1.32%***	−1.61%***	2.45%***	−1.40%***	−0.56%***	0.26%	−0.57%***	−0.46%***
	[*t*-stati]	[3.76]	[−4.13]	[7.96]	[−5.32]	[−3.03]	[1.18]	[−18.58]	[−18.66]
INST	Performance	0.64%***	0.51%**	0.89%***	0.49%***	−0.05%	0.23%	−0.19%***	−0.22%***
	[*t*-stat]	[4.49]	[2.04]	[6.71]	[3.04]	[−0.62]	[1.32]	[−18.23]	[−19.52]

Panel C. Annual Trading Performance Using Tracking Portfolios: Around Earnings Announcements

		Total		Stock Selection		Market Timing		Trading Cost	
		EA	EA/All Days	EA	EA/All Days	EA	EA/All Days	EA	EA/All Days
RT1	Performance	−0.39%***	6.89%	−0.36%***	8.96%	0.02%	−8.63%	−0.04%***	3.06%
	[*t*-stat]	[−3.68]		[−4.14]		[0.57]		[−5.91]	
RT2	Performance	−0.36%***	7.71%	−0.31%***	12.27%	0.01%	−4.85%	−0.06%***	3.12%
	[*t*-stat]	[−4.08]		[−3.51]		[0.47]		[−5.90]	
RT3	Performance	−0.24%***	6.61%	−0.18%***	11.46%	0.00%	0.03%	−0.05%***	3.18%
	[*t*-stat]	[−4.24]		[−2.95]		[−0.01]		[−5.86]	
RT4	Performance	−0.14%***	5.02%	−0.07%	8.36%	−0.01%	4.32%	−0.05%***	3.26%
	[*t*-stat]	[−2.70]		[−1.10]		[−1.32]		[−5.86]	
RT5	Performance	0.10%	−33.28%	0.14%**	13.81%	−0.01%	4.77%	−0.03%***	3.37%
	[*t*-stat]	[0.91]		[2.56]		[−1.30]		[−5.87]	
INST	Performance	0.11%*	9.80%	0.12%**	9.01%	0.00%	1.92%	−0.01%***	3.56%
	[*t*-stat]	[1.81]		[2.01]		[0.26]		[−5.86]	

C. Performances for Momentum and Contrarian Trading Using the BLLO Method

The previous section shows that different retail investors have different momentum and contrarian patterns, which contribute to their different predictive patterns for future returns. For instance, smaller retail investors display daily momentum patterns, which contribute to their negative return predictive pattern, while the largest retail investors display daily contrarian trading patterns, which contribute to their positive predictive pattern for future returns. Does the momentum or contrarian trading pattern also play a significant role in the performance of net buy- and sell-tracking portfolios?

To answer this question, we separate the total performances in Table 8, Panel A into two scenarios, "daily momentum" and "daily contrarian." If the investor group G's order imbalance times the previous day's return, $\mathrm{OIB}(i,d,G)\times\mathrm{RET}(i,d-1)$, is positive, we classify the stock i on day d for investor group G to be "daily momentum" because they buy when the price rises and they sell when the price drops; if $\mathrm{OIB}(i,d,G)\times\mathrm{RET}(i,d-1)$ is negative, then we classify the stock i on day d for investor group G to be "daily contrarian." We then calculate the performance for each scenario.

Panel B of Table 8 reports the results. Taking RT1 as an example, following the daily momentum pattern, the annualized return on investment for RT1 is −4.57% with *t*-statistic of −8.47, which is significantly worse than the annualized return on investment of RT1 at −1.04% with *t*-statistic of −2.57 when they follow contrarian strategy. The results show that momentum trading results in worse performance than that of contrarian trading. Similar patterns exist for RT2–RT4. For RT5, the annualized return on investment is −1.61% when their trades display daily momentum patterns and 1.32% when they display daily contrarian patterns. For INST, the annualized performances are positive in both momentum (0.51%) and contrarian (0.64%) conditions, while the contrarian conditions have slightly higher returns on investments. Regarding the 3 components of the total performances, the stock selection performances are in general better under the "daily contrarian" column than the "daily momentum" column; the market timing component is mostly insignificant under "daily momentum" but significantly negative under "daily contrarian," and the trading costs are similar for momentum and contrarian columns. In other words, investors generally perform better when they pursue daily contrarian rather than daily momentum trading strategies.

D. Performances around Earnings Announcements using the BLLO Method

Given that earnings announcements are the most informative firm-level events, we are also curious to understand how they contribute to the overall performance of the tracking portfolios. Our prior finding is that if investors have an information advantage around earnings announcements, they would have better performance around earnings announcements, while if investors have disadvantages around these event days, they would have worse losses on these days. There are 15,702 quarterly earnings announcement events during our 3.5-year sample.

As in Section IV.C, we consider both the day before earnings announcements and the announcement day as event days, accounting for 2.98% (=15,702 × 2/1,053,148) of the total stock-day observations. For this analysis, we compute the total performance on these event days and compare it with the overall performance on all days.

We present the results in Panel C of Table 8. For ease of comparison, we present the estimates over the earnings event days under "EA" and its proportion to all days under "EA/all days." The annualized total performance over earnings days for RT1 is −0.39%, which accounts for 6.89% of the total performance of all days. The stock selection, market timing, and trading cost components are −0.36%, 0.02%, and −0.04%, accounting for 8.96%, −8.63%, and 3.06% of all days, respectively. The coefficients on total performance, stock selection, and trading costs are all statistically significant. Considering that earnings days only account for 2.98% of all days, the total performance and stock selection component of the tracking portfolio following RT1 order flows is particularly bad on earnings days, which echoes our findings in Section IV. The trading cost on EA days accounts for approximately 3% of all days, which is close to the overall number of EA days. Similar patterns are observed for RT2–RT4. For RT5, the annualized EA total performance of RT5 is insignificant. For the stock selection component, RT5's annualized return on investment is 0.14%, accounting for 13.81% of RT5's total stock selection, with a t-statistic of 2.56. This finding is consistent with our earlier result that RT5 might be better informed around EA days. For INST, the total performance over EA days is 0.11%, accounting for 9.80% of the total performance, with a significant t-statistic. The stock selection component shows a similar pattern. These findings indicate that institutional investors perform better on EA days than on all days.[28]

VI. Robustness

A. Subperiods: Predictive Patterns for Different Years

Do the predictive patterns of retail order flows differ when market conditions change? Our 3.5-year sample period provides a good setting with different market conditions. The market return is −8.4% in 2016, 12.1% in 2017, −22.2% in 2018, and 21.2% in 2019. With 2 positive and negative returns, our sample period provides settings for both market ups and downs. To examine the predictive patterns of order flows for different years, we modify the Fama and MacBeth (1973) specification in equation (2) and interact the order flow variables with year dummies. The coefficients on the interaction terms provide information on whether the predictive pattern changes on each year.

From Panel A of Table 9, the negative predictive pattern of order flow from RT1 to RT4 for the next-day return, as observed in Table 2, is robust for different

[28]Previous literature (Kaniel et al. (2008), Barrot et al. (2016)) argues that retail investors provide liquidity to the market and are compensated for their liquidity provision. We further decompose retail investors' performances from the perspective of liquidity provision and provide the results in Appendix Table VII in the Supplementary Material. The performance of both small and large retail investors is significantly negative if the trades demand liquidity. But if they provide liquidity, small retail investors' performance becomes insignificantly different from 0, and large retail investors are compensated for liquidity provision, which generally supports the hypothesis that liquidity provision is compensated.

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

years. Moreover, it is interesting to notice that the negative magnitudes are generally larger for 2016 and 2018, indicating that the negative predictive pattern is even larger in a downward market. For the largest retail investor, RT5, the positive predictive pattern is also larger during 2016 and 2018, indicating that their information advantage, if any, may also be stronger when the market is down. Similar patterns are observed for institutional investors.

B. Subgroups: Predicting Patterns Across Firms with Different Characteristics

Previous studies show that stock returns can be significantly affected by firm and stock characteristics, such as size, EP ratio, and liquidity. Do the predictive patterns of retail order flows differ across firms with different characteristics? To answer this question, we modify the Fama and MacBeth (1973) specification in equation (2) and allow different coefficients for firms with different characteristics, by including interactions with characteristic dummies. Consider the size as an example. We first divide all firms on day d into 3 groups based on the previous month-end firm market capitalization. The dummy variable, $\text{SMALL\_SIZE}_{i,d-1}$, takes the value 1 if firm i belongs to the smallest one-third of firms, and 0 otherwise; $\text{MEDIAN\_SIZE}_{i,d-1}$ takes the value 1 if firm i belongs to the medium one-third of firms, and 0 otherwise; and $\text{LARGE\_SIZE}_{i,d-1}$ takes the value 1 if firm i belongs to the largest one-third of firms, and 0 otherwise. The coefficients provide information on whether the predictive pattern changes for firms of different sizes.

Panel B of Table 9 reports the results. In the first 3 rows, we separate the firms based on their market capitalization. The negative predictive pattern of order flow from RT1 to RT4 for the next-day return, as shown in Table 2, is quite robust for firms with different sizes. However, it is interesting to notice that the magnitudes generally decrease from the smallest to the largest firms, indicating that the negative predictive pattern is the strongest for smaller firms. For large retail investors, RT5, the positive predictive pattern remains for small- and medium-sized firms, but not for large firms, indicating that their information about future returns, if any, might be concentrated in smaller firms. In comparison, order flows from institutions significantly predict next-day returns in all 3 rows, and more so for large firms, suggesting that their information about future returns, if any, might be more prominent for larger firms. When we separate firms by EP, turnover, and stock prices, we observe similar patterns. In other words, the predictive patterns in Table 2 are generally robust across firms with different characteristics, and the negative (positive) predictive power of smaller (larger) retail investors is stronger for small, low-EP, and higher-turnover firms, while the positive predictive power of institutional investors is stronger for large and high EP firms.[29]

[29]We obtain a 3-month sample from Jan. 2019 to Mar. 2019 with investors' gender and age information and examine whether there are significant differences between male/female and young/old. The results are reported in Appendix Table VIII in the Supplementary Material. Across all gender–age groups, older male investors trade the most. For return prediction, male investors across all ages significantly and negatively predict returns, especially for older males, while they are insignificant and different from 0 for female retail investors. These patterns are mostly consistent with previous findings in Barber and Odean (2001).

TABLE 9

Robustness

Table 9 reports the robustness results. The sample period covers Jan. 2016 to June 2019, and the sample firms are A-share stocks merged with the proprietary data and have at least 15 trading days in the previous month. Panel A reports the return prediction on each year. Panel B reports the return prediction by different stock characteristics. Panel C reports return predictions by including all filters in Liu et al. (2019)). Panel D reports return predictions by excluding leveraged trading, which includes margin buys, short sales, and collateral trading. Panel E reports return predictions by excluding days that hit the price limits (+10% and −10%). Panel F reports the alternative method using Liu et al.'s (2019)) 3-factor adjusted alphas for long–short portfolios. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using the Newey and West (1987) method, with optimal bandwidths chosen by following Newey and Wess (1994). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Cross-Sectional Return Predictions by Different Years

				Dependent Variable: RET					
Year	Market Return	OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
2016	−8.4%	OIB(−1) × Y2016	Estimate	−0.0027***	−0.0025***	−0.0020***	−0.0004***	0.0004***	0.0003***
			[*t*-stat]	[−3.46]	[−3.41]	[−3.45]	[−3.37]	[3.43]	[3.26]
2017	12.1%	OIB(−1) × Y2017	Estimate	−0.0023***	−0.0022***	−0.0015***	−0.0002***	0.0003***	0.0004***
			[*t*-stat]	[−3.83]	[−3.70]	[−3.71]	[−3.11]	[3.39]	[3.87]
2018	−22.2%	OIB(−1) × Y2018	Estimate	−0.0026***	−0.0027***	−0.0018***	−0.0002**	0.0002***	0.0006***
			[*t*-stat]	[−3.75]	[−3.80]	[−3.70]	[−2.17]	[3.45]	[3.81]
2019	21.2%	OIB(−1) × Y2019	Estimate	−0.0018**	−0.0018**	−0.0013**	−0.0001*	0.0003**	0.0003**
Jan.–June			[*t*-stat]	[−2.25]	[−2.26]	[−2.29]	[−1.84]	[2.13]	[2.35]
		Control		Yes	Yes	Yes	Yes	Yes	Yes
		INTERQUARTILE_RETURN_Y2016		−0.0569%	−0.0434%	−0.0289%	−0.0088%	0.0136%	0.0248%
		INTERQUARTILE_RETURN_Y2017		−0.0478%	−0.0377%	−0.0240%	−0.0053%	0.0127%	0.0255%
		INTERQUARTILE_RETURN_Y2018		−0.0603%	−0.0527%	−0.0341%	−0.0060%	0.0129%	0.0375%
		INTERQUARTILE_RETURN_Y2019		−0.0414%	−0.0339%	−0.0215%	−0.0033%	0.0123%	0.0160%

Panel B. Cross-Sectional Return Predictions by Different Stock Characteristics

OIB Var.	OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
OIB(−1) × SMALL_SIZE	−0.0119***	−0.0131***	−0.0090***	−0.0006***	0.0020***	0.0014***
OIB(−1) × MEDIAN_SIZE	−0.0096***	−0.0093***	−0.0066***	−0.0011***	0.0009***	0.0015***
OIB(−1) × LARGE_SIZE	−0.0067***	−0.0061***	−0.0044***	−0.0012***	0.0001	0.0021***
OIB(−1) × LOW_EP	−0.0122***	−0.0131***	−0.0095***	−0.0010***	0.0020***	0.0015***
OIB(−1) × MEDIAN_EP	−0.0097***	−0.0098***	−0.0069***	−0.0009***	0.0012***	0.0015***
OIB(−1) × HIGH_EP	−0.0062***	−0.0056***	−0.0039***	−0.0007***	0.0001	0.0017***
OIB(−1) × LOW_TURNOVER	−0.0061***	−0.0057***	−0.0041***	−0.0007***	0.0003***	0.0013***
OIB(−1) × MEDIAN_TURNOVER	−0.0091***	−0.0092***	−0.0066***	−0.0010***	0.0011***	0.0014***
OIB(−1) × HIGH_TURNOVER	−0.0159***	−0.0176***	−0.0128***	−0.0009***	0.0026***	0.0019***
OIB(−1) × LOW_PRICE	−0.0088***	−0.0086***	−0.0067***	−0.0012***	0.0009***	0.0014***
OIB(−1) × MEDIAN_PRICE	−0.0098***	−0.0095***	−0.0068***	−0.0007***	0.0013***	0.0016***
OIB(−1) × HIGH_PRICE	−0.0094***	−0.0094***	−0.0060***	−0.0007***	0.0014***	0.0016***

(continued on next page)

TABLE 9 (continued)

Robustness

Panel C. Cross-Sectional Return Predictions, Including all Filters from Liu et al. (2019)

		Dependent Variable: RET					
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
OIB(−1)	Estimate	−0.0075***	−0.0071***	−0.0052***	−0.0011***	0.0004***	0.0017***
	[*t*-stat]	[−21.47]	[−19.65]	[−16.85]	[−8.09]	[5.07]	[16.98]
	Interquartile return	−0.17%	−0.14%	−0.09%	−0.03%	0.02%	0.11%

Panel D. Cross-Sectional Return Predictions, Excluding Leveraged Trading

		Dependent Variable: RET					
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
OIB(−1)	Estimate	−0.0092***	−0.0088***	−0.0056***	−0.0003***	0.0008***	0.0016***
	[*t*-stat]	[−17.26]	[−17.16]	[−15.54]	[−3.37]	[10.09]	[18.28]
	Interquartile return	−0.20%	−0.16%	−0.11%	−0.01%	0.05%	0.11%

Panel E. Cross-Sectional Return Predictions, Excluding Days That Hit the Price Limits (+10%, and −10%)

		Dependent Variable: RET					
OIB Var.		OIB_RT1	OIB_RT2	OIB_RT3	OIB_RT4	OIB_RT5	OIB_INST
OIB(−1)	Estimate	−0.0087***	−0.0087***	−0.0064***	−0.0010***	0.0009***	0.0017***
	[*t*-stat]	[−19.60]	[−19.57]	[−17.95]	[−8.38]	[10.41]	[18.10]
	Interquartile return	−0.19%	−0.16%	−0.11%	−0.03%	0.04%	0.11%

Panel F. Risk-Adjusted Alphas for Long–Short Portfolios over the Next 12 Weeks

	OIB_RT1		OIB_RT2		OIB_RT3		OIB_RT4		OIB_RT5		OIB_INST	
	Alpha	*t*-Statistic	Alpha	*t*-Statistic	Alpha	*t*-Statistic	Alpha	*t*-Statistic	Alpha	*t*-Statistic	Alpha	*t*-Statistic
1 day	−0.0042***	−20.55	−0.0036***	−17.53	−0.0027***	−15.13	−0.0007***	−6.71	0.0017***	12.71	0.0025***	19.28
1 week	−0.0089***	−15.34	−0.0068***	−12.97	−0.0038***	−9.22	−0.0004	−1.63	0.0034***	11.29	0.0056***	13.67
2 weeks	−0.0115***	−11.96	−0.0091***	−10.50	−0.0048***	−6.84	−0.0001	−0.16	0.0046***	9.08	0.0075***	11.53
3 weeks	−0.0130***	−10.52	−0.0103***	−9.37	−0.0051***	−5.68	0.0001	0.26	0.0054***	8.71	0.0086***	9.73
4 weeks	−0.0130***	−8.64	−0.0104***	−7.94	−0.0052***	−4.94	0.0003	0.81	0.0056***	7.87	0.0098***	9.19
5 weeks	−0.0130***	−7.45	−0.0099***	−6.53	−0.0050***	−4.01	0.0006	1.13	0.0062***	7.51	0.0098***	8.39
6 weeks	−0.0148***	−7.84	−0.0110***	−6.49	−0.0056***	−3.73	0.0005	0.98	0.0064***	7.29	0.0103***	7.99
7 weeks	−0.0173***	−8.76	−0.0130***	−6.74	−0.0068***	−3.61	0.0001	0.49	0.0064***	7.25	0.0119***	8.29
8 weeks	−0.0183***	−8.43	−0.0136***	−6.37	−0.0069***	−3.36	0.0001	0.50	0.0063***	6.63	0.0128***	6.95
9 weeks	−0.0191***	−7.62	−0.0142***	−5.59	−0.0070***	−2.97	0.0003	0.64	0.0068***	7.45	0.0135***	6.18
10 weeks	−0.0187***	−6.98	−0.0136***	−5.27	−0.0063***	−2.98	0.0005	0.87	0.0065***	7.68	0.0133***	6.38
11 weeks	−0.0187***	−6.24	−0.0141***	−4.66	−0.0070***	−2.77	−0.0005	0.27	0.0061***	6.83	0.0138***	6.00
12 weeks	−0.0183***	−5.75	−0.0139***	−4.52	−0.0072***	−2.89	−0.0010	0.03	0.0057***	5.78	0.0137***	6.03

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109024000085 Published online by Cambridge University Press

C. Different Filters

We consider 3 different sets of filters in the robustness check. In the main results, we discard stocks with fewer than 15 days of trading during the most recent month. In addition to this filter, Liu et al. (2019) also eliminate stocks that have become public within the past 6 months, stocks with less than 120 days of trading during the past 12 months, and the smallest 30% of all firms listed in China A-share market. We add all these additional filters and check the robustness of our results. In Panel C of Table 9, the order imbalance prediction directions are similar to those in Table 2 and the economic magnitude of institutions remains large. Thus, our main results are robust to the stricter filters from Liu et al. (2019).

Our second set of filters relates to leverage trading. Our trade-level data identify investors' margin buys, short sales, and collateral trades. Leveraged trading may differ from nonleverage trading. Each day, margin buys account for 10% of the trading volume, short sales account for 0.2%, and collateral trading accounts for 15% during our sample period. We exclude leverage trades and reestimate equation (2). Panel D of Table 9 shows that the order imbalance prediction directions are similar to the results in Table 2 and the economic magnitudes are quantitatively similar. That is, our results are also robust to whether or not we include these leverage trades.

Finally, we consider the price limits. One institutional feature of the Chinese market is price limit restrictions. That is, investors can buy and sell stocks freely when the stock's price is within ±10% of the previous day's closing price. If the price moves out of the ±10% range, trading stops until the next day opens. Chen et al. (2019) focus on price limit days and find that large investors tend to trade differently on these days. Here, we examine whether our results hold when we exclude price limit days. We reestimate equation (2) without the price limit days and present the results in Panel E of Table 9. Our results are robust to whether or not we include these price limit days.

D. Alternative Portfolio Forming Method

Our main results on return predictive patterns are based on the Fama and MacBeth (1973) regressions, which assume linear relationship between future returns and order flow variables. In this section, we adopt an alternative portfolio approach and examine whether our results still hold. More specifically, we sort firms into 5 groups each day based on the previous day's order imbalance from a particular investor group, buying and selling 20% of stocks with the highest and lowest order imbalance measures for that particular investor group. We report the risk-adjusted returns (alphas) on this long–short strategy for 1 to 60 days, where we conduct risk adjustment using the Liu et al.'s (2019) 3-factor model.

From Panel F of Table 9, the 1-day long–short portfolio alpha, using the previous day's order imbalance from RT1, is −0.0042 and highly significant. From 1 week to 12 weeks, the cumulative alphas for the long–short portfolio decrease from −0.0089 to −0.0183, and they are all highly significant. That is, the cumulative alphas using OIB_RT1 are consistently negative and significant, and there are no signs of reversal within 12 weeks, which echoes our earlier results in Tables 2 and 5. Similar patterns exist for RT2 and RT3. For RT4, the 1-day alpha is negative at −0.0007 but quickly becomes insignificant when we extend the holding horizon to 1 week, indicating RT4 for horizons longer than 1 day. For the largest retail investors in RT5, the 1-day alpha is 0.0017, which is positive and significant. The 12-week cumulative alpha is 0.0057, still positive and significant, confirming the results in Tables 2 and 5 that RT5 has both short- and long-term predictive power for future returns. The results for OIB_INST are similar to those for OIB_RT5.

VII. Conclusion

Using comprehensive retail trading and holding data from 2016 to 2019, we divide tens of millions of retail investors into 5 groups based on their account sizes and examine their roles in the price discovery process in terms of return predictive power and the driving forces of these predictive patterns. Retail investors with account sizes of less than 3 million CNY buy and sell stocks in the wrong direction. The prices of stocks they buy experience negative returns the next day, whereas those they sell experience positive returns. For retail investors with large account balances, trading predicts returns correctly. By tracing their differences in predicting future returns, we provide evidence that the negative predictive power of retail investors with smaller account sizes is mostly related to their order persistence, daily momentum trading, behavioral biases, and failures in processing earnings news. By contrast, the positive predictive power of large retail investors is mostly associated with order persistence, contrarian trading, trading against behavioral biases, and advantages in processing earnings news. Following BLLO (2009), we construct net buy and sell portfolios for each group of investors and track their trading flows. This performance generates results consistent with the predictive patterns, in the sense that smaller retail investors have low stock selection skills and market timing skills, while large retail investors have better stock selection skills.

Our study of the heterogeneous trading behavior of Chinese retail investors provides many unique insights into this large group of investors. In addition, the exchange acknowledges the heterogeneity of retail investors and focuses on adopting policies on investor education and suitability that restrict some kinds of trading for the smallest accounts. For example, retail investors are required to have at least 500,000 CNY holdings of stocks for at least 20 trading days to open a leverage trading account or to trade on a riskier Science and Technology Innovation Board (or STAR Market). These policies effectively exclude the smallest retail investors from leverage trading and trading on riskier start-ups, which could help protect them from even worse losses. Our study also raises many interesting questions. For example, why do retail investors dominate trading in the Chinese stock market? What role do institutional investors play? We leave these interesting and important questions to future research.

Supplementary material

To view supplementary material for this article, please visit http://doi.org/10.1017/S0022109024000085.

References

An, L.; D. Lou; and D. Shi. "Wealth Redistribution in Bubbles and Crashes." *Journal of Monetary Economics*, 126 (2022), 134–153.

Bach, L.; L. E. Calvet; and P. Sodini. "Rich Pickings? Risk, Return, and Skill in Household Wealth." *American Economic Review*, 110 (2020), 2703–2747.

Bali, T. G.; N. Cakici; and R. F. Whitelaw. "Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns." *Journal of Financial Economics*, 99 (2011), 427–446.

Barber, B. M.; X. Huang; T. Odean; and C. Schwarz. "Attention Induced Trading and Returns: Evidence from Robinhood Users." *Journal of Finance*, 77 (2022), 3141–3190.

Barber, B. M.; Y.-T. Lee; Y.-J. Liu; and T. Odean. "Just How Much Do Individual Investors Lose by Trading?" *Review of Financial Studies*, 22 (2009), 609–632.

Barber, B. M., and T. Odean. "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *Journal of Finance*, 55 (2000), 773–806.

Barber, B. M., and T. Odean. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics*, 116 (2001), 261–292.

Barber, B. M., and T. Odean. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *Review of Financial Studies*, 21 (2008), 785–818.

Barber, B. M.; T. Odean; and N. Zhu. "Do Retail Trades Move Markets?" *Review of Financial Studies*, 22 (2009), 151–186.

Barrot, J.-N.; R. Kaniel; and D. A. Sraer. "Are Retail Traders Compensated for Providing Liquidity?" *Journal of Financial Economics*, 120 (2016), 146–168.

Blume, M. E., and R. F. Stambaugh. "Biases in Computed Returns: An Application to the Size Effect." *Journal of Financial Economics*, 12 (1983), 387–404.

Boehmer, E.; C. M. Jones; X. Zhang; and X. Zhang. "Tracking Retail Investor Activity." *Journal of Finance*, 76 (2021), 2249–2305.

Boehmer, E.; B. Sang; Z. Zhang "Can Retail Investors Learn from Insiders?" Working Paper, Singapore Management University and University of Bristol (2020).

Boehmer, E., and J. Wu. "Short Selling and the Price Discovery Process." *Review of Financial Studies*, 26 (2013), 287–322.

Chen, T.; Z. Gao; J. He; W. Jiang; and W. Xiong. "Daily Price Limits and Destructive Market Behavior." *Journal of Econometrics*, 208 (2019), 249–264.

Chordia, T., and A. Subrahmanyam. "Order Imbalance and Stock Returns: Theory and Evidence." *Journal of Financial Economics*, 72 (2004), 485–518.

Diether, K. B.; K. H. Lee; and I. M. Werner. "Short-Sale Strategies and Return Predictability." *Review of Financial Studies*, 22 (2009), 575–607.

Dorn, D.; G. Huberman; and P. Sengmueller. "Correlated Trading and Returns." *Journal of Finance*, 63 (2008), 885–920.

Eaton, G. W.; T. C. Green; B. Roseman; and Y. Wu. "Retail Trader Sophistication and Stock Market Quality: Evidence from Brokerage Outages." *Journal of Financial Economics*, 146 (2022), 502–528.

Fama, E. F., and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (1973), 607–636.

Gervais, S.; R. Kaniel; and D. H. Mingelgrin. "The High-Volume Return Premium." *Journal of Finance*, 56 (2001), 877–919.

Grinblatt, M., and M. Keloharju. "The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set." *Journal of Financial Economics*, 55 (2000), 43–67.

Grinblatt, M.; M. Keloharju; and J. T. Linnainmaa. "IQ, Trading Behavior, and Performance." *Journal of Financial Economics*, 104 (2012), 339–362.

Hu, C.; Y.-J. Liu; and X. Xu. "The Valuation Effect of Stock Dividends or Splits: Evidence from a Catering Perspective." *Journal of Empirical Finance*, 61 (2021), 163–179.

Jiang, L.; J. Liu; L. Peng; and B. Wang. "Investor Attention and Asset Pricing Anomalies." *Review of Finance*, 26 (2022), 563–593.

Kaniel, R.; S. Gideon; and T. Sheridan. "Individual Investor Sentiment and Stock Returns." *Journal of Finance*, 63 (2008), 273–310.

Kelley, E. K.; and P. C. Tetlock. "How Wise are Crowds? Insights from Retail Orders and Stock Returns." *Journal of Finance*, 68 (2013), 1229–1265.

Li, X.; Z. Geng; A. Subrahmanyam; and H. Yu. "Do Wealthy Investors Have an Informational Advantage? Evidence Based on Account Classifications of Individual Investors." *Journal of Empirical Finance*, 44 (2017), 1–18.

Liao, J.; C. Peng; and N. Zhu. "Extrapolative Bubbles and Trading Volume." *Review of Financial Studies*, 35 (2022), 1682–1722.
Linnainmaa, J. T. "Do Limit Orders Alter Inferences About Investor Performance and Behavior?" *Journal of Finance*, 65 (2010), 1473–1506.
Liu, H.; C. Peng; W. A. Xiong; and W. Xiong. "Taming the Bias Zoo." *Journal of Financial Economics*, 143 (2021), 716–741.
Liu, J.; R. F. Stambaugh; and Y. Yuan. "Size and Value in China." *Journal of Financial Economics*, 134 (2019), 48–69.
Liu, Y.; G., Zhou; and Y. Zhu. "Trend Factor in China: The Role of Large Individual Trading." *Review of Asset Pricing Studies*, forthcoming (2024).
Newey, W. K., and K. D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.
Newey, W. K., and K. D. West. "Automatic Lag Selection in Covariance Matrix Estimation." *Review of Economic Studies*, 61 (1994), 631–653.
Titman, S.; C. Wei; and B. Zhao. "Corporate Actions and the Manipulation of Retail Investors in China: An Analysis of Stock Splits." *Journal of Financial Economics*, 145 (2022), 762–787.
Van Kervel, V., and A. J. Menkveld. "High Frequency Trading Around Large Institutional Orders." *Journal of Finance*, 74 (2019), 1091–1137.
Welch, I. "The Wisdom of the Robinhood Crowd." *Journal of Finance*, 77 (2022), 1489–1527.