Are Foreign Investors Informed?
Trading Experiences of Foreign Investors in China

Christian T. Lundblad, Donghui Shi, Xiaoyan Zhang, and Zijian Zhang*

December 1, 2022

ABSTRACT

Using a proprietary dataset from 2016 to 2019, we find that order flows from foreign investors, facilitated by regulatory liberalization through several channels, present strong predictive power for future stock returns in the Chinese market. Most surprisingly, foreign investors possess the ability to process local firm-level public news, and their predictive power is particularly strong on large price movement days when the implications of firm-level information are likely most pronounced. Evidence also shows that foreign investors have some capability in processing global market-level information. Finally, regulatory reforms that generally relax investment access requirements further improve foreign investors' predictive power.

Keywords: Foreign investors, the Chinese stock market, public information, market liberalization.
JEL classification: G12, G14, G15, G18.

* Christian T. Lundblad is with Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Donghui Shi is with Fanhai International School of Finance, Fudan University, Xiaoyan Zhang is with PBC School of Finance, Tsinghua University, and Zijian Zhang is with PBC School of Finance, Tsinghua University.

Are Foreign Investors Informed?
Trading Experiences of Foreign Investors in China

This version: December 1, 2022

ABSTRACT

Using a proprietary dataset from 2016 to 2019, we find that order flows from foreign investors, facilitated by regulatory liberalization through several channels, present strong predictive power for future stock returns in the Chinese market. Most surprisingly, foreign investors possess the ability to process local firm-level public news, and their predictive power is particularly strong on large price movement days when the implications of firm-level information are likely most pronounced. Evidence also shows that foreign investors have some capability in processing global market-level information. Finally, regulatory reforms that generally relax investment access requirements further improve foreign investors' predictive power.

Keywords: Foreign investors, the Chinese stock market, public information, market liberalization.
JEL classification: G12, G14, G15, G18.

I. Introduction

Many studies show that foreign capital plays a significant and positive role in spurring the development of emerging stock markets. For example, foreigners help to lower firms' capital costs (Bekaert and Harvey, 2000), spur economic growth (Bekaert, Harvey and Lundblad, 2005), facilitate cross-border mergers and acquisitions (Ferreira, Massa and Matos, 2010), promote corporate governance (Ferreira and Matos, 2008; Aggarwal et al., 2011), expedite global information transmission (Bae et al., 2012), and improve price efficiency (Kacperczyk, Sundaresan and Wang, 2021). Over the past 20 years, regulators from China, clearly recognizing these benefits, consistently invited foreign investors to participate in the development of the Chinese stock market.

To allow foreign capital access to domestic Chinese equity A-shares, three major channels were created. First, the Qualified Foreign Institutional Investors (QFII) program, launched in November 2002, allowed foreign institutional investors to trade equities and other financial instruments by converting foreign currencies into onshore RMB. Second, the Renminbi QFII (RQFII) program, introduced in December 2011, permitted qualified overseas institutional investors to invest directly in the domestic capital market using offshore RMB. Third, and most recently, the Hong Kong Stock Connect (HKC) programs, linking the Hong Kong stock market with the Shanghai Stock Exchange and Shenzhen Stock Exchange, were launched in November 2014 and December 2016, respectively. HKC enables Hong Kong and overseas individual and institutional investors to trade eligible stocks listed on the mainland exchanges. By the end of 2021, foreign investors held around RMB 3.67 trillion in A-shares through these various channels, collectively accounting for 4.97% of A-share aggregate market capitalization.

Despite the rapidly growing presence of foreign investors in China, physical and language barriers remain. Hence, it is natural to ask whether these investors can process Chinese local market information and whether their trades can predict future price movements. In other words, we are interested in two questions: first, whether foreign investors are informed about future local stock returns, and second, if they are informed, what types of information - firm-specific, market-level, or global - are they able to process? Given that the Chinese equity markets, collectively as the second largest in the world, play an increasingly important role in global asset allocation,[1] and direct trading data on foreign investors is scarce, answers to our research questions are interesting and particularly important for the international investment community.

We are grateful that the Shanghai Stock Exchange, the largest stock exchange in China, for permitting access to a comprehensive sample of investors' daily trading records from 2016 to 2019, otherwise not available to the public. For compliance purposes, the exchange identifies each buy and sell order with the originators, domestic or foreign, and their access channels such as QFII, RQFII or HKC. Based on this trade level information, we aggregate foreign order flows at the stock level each day. For comparison purposes, we also collect order flows from local institutions, such as mutual funds, hedge funds, and others, to serve as a benchmark.

We first examine whether foreign investors are informed by measuring the link between their order flows and future stock price movement in the China A-share market. Given the language and culture differences and the distances between foreign and local markets, several studies show that foreign investors are at a disadvantage when trading in local markets. For instance, Kang and Stulz (1997) find that foreign ownership in Japan does not predict future stock returns. In our case, the foreign order flows from all three channels into China have significant predictive power for future stock returns. Taking QFII as an example, an interquartile increase in daily QFII order flow is

[1] As an example, the evolution in MSCI index construction to include Chinese A-share is fundamentally transforming global passive fund allocations. https://www.msci.com/msci-china-a-inclusion.

associated with an 11.88 bps increase in the next day's stock return (or 29.94% annualized), with a highly significant *t*-statistic of 17.02. When we turn to RQFII and HKC, an interquartile increase in daily RQFII and HKC order flows is associated with 3.05 bps and 7.57 bps increases in the next day's return (or 7.69% and 19.08% annualized), respectively. For comparison, an interquartile increase in daily local institutional order flow is associated with a 9.33 bps increase in the next day's return (or 23.51% annualized). Taking these figures together, foreign investors' trading activity significantly predicts future local stock returns, and thus they seem to be informed about relevant fundamental information in China. Further, their predictive power is on par with their local institutional counterparts. When we extend the prediction window from days to weeks, foreign investors still significantly predict cumulative stock returns over, at least, the next 12 weeks, implying that the information they have is not transient.[2]

Given our evidence on return predictability, a natural next question is: what types of information drive foreign investors' predictive power for future returns? Notice that information can be separated into different categories: firm-level vs. market-level, and local vs. global. The prior literature shows that the physical distances and language barriers make it difficult for foreign investors to process local firm-level news. However, this does not seem to be the case for foreign investors in China. We first collect data on firm-level events, including earnings announcements, analyst recommendations and media news. There is clear evidence that foreign investors can process local firm-level information, in the sense that foreign investor order flows can directly predict future earnings news and media news, and the predictive power of foreign investor order

[2] Readers might wonder if both foreign and local and foreign institutional investors positively predict returns, who are their counterparties. Our data show that more than half of the counterparties for foreign investors are local retail investors, and the predictive power of foreign investors' trades is stronger when they trade against retail investors than local institutions.

flows for returns is, relative to non-event days, significantly higher on the event days. We connect the predictive patterns of foreign investors to their cross-border business, and provide anecdotal evidence that foreign investors are stationed in or close to China, and they hire professionals to overcome language and culture barriers, which seem to be effective.

Do foreign investors have an advantage in processing global market news, possibly because foreign investors have better access to global market news? Here, we use local and global market returns as proxies for market-level new information. We first show that the aggregate order flows for QFII can significantly predict next-day market returns, consistent with earlier findings that foreign institutions possess advantages in processing market-level information. We also provide suggestive evidence that the predictive power of foreign order flows in the cross-section is higher when the global market experiences large price movements, indicating that they, to some extent, may have abilities in processing the global market level information. However, the magnitude and significance of foreign investors' cross-sectional predictive power on market-level news days are much lower than those for firm-level news days, indicating that market-level information has limited influence in helping foreign investors' ability in selecting stocks in the cross-section.

Finally, for our four-year sample, Chinese regulatory authorities gradually relax the restrictions on foreign capital, allowing better access for foreign investors to participate in the Chinese stock market. For instance, they increase investment quotas in 2016, relax capital flow controls in 2018, and lift asset allocation limitations in 2019, etc. While these measures increasingly permit inflows of foreign capital, does the broader participation of foreign investors improve their overall predictive power regarding future stock returns? This is an interesting empirical question. If the friendlier regulatory environment attracts more informed and active investors, then the predictive power of foreign order flows would increase. However, if the liberalization attracts less informed investors (say, index funds at one extreme), then the predictive power that we document might decrease. Our empirical results show that expanding investment quotas and capital flows, on average, improve foreign investors' return predictive power.

For these research questions, the previous literature largely provides evidence against foreign investors' informational advantages. In the context of the rapidly growing Chinese stock market, are foreign investors capable of processing information so that their trading activity predicts future Chinese stock returns? Will the patterns be similar to those findings for other emerging markets? If they are different, what types of information are particularly relevant?

Our study is related to three strands of the previous literature. The first examines whether foreign investors face informational disadvantages in the local equity market. Similar to Kang and Stulz (1997), Choe, Kho and Stulz (2005) and Dvořák (2005) show that foreign investors are at informational disadvantage than locals in South Korea and Indonesia, respectively. More recently, Froot and Ramadorai (2008) suggest that a positive relation between international portfolio flows and closed-end fund performance. Ferreira et al. (2017) find that foreign institutional ownership predicts local stock returns, and advocate a price pressure explanation rather than a firm fundamentals explanation.[3] A second strand of the literature contains studies on how foreign investors behave in the Chinese stock market. Chen, Wang and Zhu (2019) and Bian et al. (2020) both focus on HKC investors, which has publicly available data, and find HKC investors have some predictive power for returns, and their trades reduce volatility.[4] Finally, the third strand of literature concerns institutional investors' informational advantages over public information. For

[3] Other studies investigating the performance of foreign investors in local market include Brennan and Cao (1997), Dvořák (2005), Agarwal et al. (2009), Baik et al. (2013), Seasholes (2000), Grinblatt and Keloharju (2000), Froot, O'Connell and Seaholes (2001), and Bailey, Mao and Sirodom (2007).

[4] Existing studies also examine other aspects of foreign investors in China, such as information asymmetry (Chan, Menkveld and Yang, 2008), corporate governance (Huang and Zhu, 2015), reactions to analysts' recommendation (Jia, Wang and Xiong, 2017), firm disclosure (Yoon, 2021) and corporate activity (Ma, Rogers and Zhou, 2021).

example, Irvine, Lipson, and Puckett (2007) find that institutional trades before analyst recommendation releases earn abnormal profits. Campbell, Ramadorai, and Schwartz (2009) show that institutional trades predict earnings surprises. Hendershott, Livdan, and Schürhoff (2015) show that institutional investors are informed about news content. Huang, Tan, and Wermers (2020) find that institutions can trade correctly on news tone after the earliest news release.

Compared with previous studies, our paper makes three distinct contributions. First, relying on proprietary trading data, we are one of the first studies that provide comprehensive evidence on the trading behaviors of foreign investors, QFII, RQFII and HKC, and whether their trades contain information on Chinese stocks. Second, we provide an in-depth analysis of whether and how foreign investors' order flows are related to many layers of public information, firm or market and local or global. Early studies, mostly using quarterly institutional ownership data, are unable to provide direct evidence on how these investors anticipate and process the information in their trading and holding patterns. Third, we provide evidence that regulatory changes, which facilitate foreign investors' access, improve their predictive power on local stock returns. Our findings on the predictive patterns of various foreign investors and their information processing skills are important for academic researchers, industry practitioners and regulators alike.

II. Hypothesis Development

To guide our empirical analysis, we develop four main hypotheses regarding foreign investors' informativeness, their abilities to process public information and the influence of government regulations.

The first hypothesis is about whether foreign investors' trading predicts future local stock price movements. We measure foreign investors' behavior by their trading order flows, which are widely used in studies on retail investors (Kelley and Tetlock, 2013; Barrot, Kaniel and Sraer, 2016; Boehmer et al., 2021) and institutions (Hendershott, Livdan, and Schürhoff, 2015). If order flows from a particular group of investors significantly predict future stock returns, we infer that this group of investors is informed, and vice versa. The literature provides mixed evidence on the degree to which foreign investors are informed in the local market. On the one hand, in comparison with local investors, foreign investors are physically further away from local firms and might possess poorer information sources. It is also harder to maintain relationships with local firms and analysts. Therefore, foreign investors might not be informed about local firms, or at least they are less informed than local investors. On the other hand, foreign investors are generally institutions from more developed markets. These high-powered and well-resourced institutions might have considerable advantages in information collection and processing skills. Thus, they are likely to be informed, even in an overseas market. Based on previous findings, we propose our first hypothesis regarding the relation between foreign investors' order flows and future stock returns:

Hypothesis 1: Foreign investors from the QFII, RQFII, and HKC programs are informed about stock prices in the Chinese stock market; that is, their order flows can predict future stock returns.

If investors are informed about future stock price movements, it is normally the case that they are informed about certain information in a manner better than the general market such that they trade in a way that benefits them when that information is released. Since we don't have the means to measure private information, we focus on available public information data. Notice that public information is "private" before its public release, and one could be informed about eventually released public information, which is still a "private" information advantage. We separate public information into two categories: firm-level information, and market-level information.

Since all firms in our sample are Chinese firms, we first restrict our attention to local firm-level information. Previous literature, such as Savor and Wilson (2016), shows that firm events such as earnings announcements and analyst activities contain valuable fundamental information, and stock prices normally exhibit a strong reaction to these releases. Therefore, if investors' return predictive power is related to their access and ability to process firm-level information, their order flows should be able to predict this news, and also predict returns to a greater degree on firm-level news days. Meanwhile, geographic distance can cause information asymmetry among investors, and investors located near their investments possess informational advantages. Due to the physical proximity and the potential language and cultural barriers, it might be challenging for foreign investors to process local firm information. If foreign investors are not able to process local firm news, we expect their order flows to fail to predict the news and they have lower return predictive power on firm news days than on non-news days. We establish our second hypothesis regarding local firm-level information:

Hypothesis 2: Foreign investors are able to process local firm information. That is, foreign investors' order flows can predict firm-level news, and the predictive power of future returns is higher on local firm news days than on non-news days.

We next turn to market-level information, such as stock market movements and essential macroeconomic indicators announcements. Macro information has a significant impact on asset prices. Given that many foreign investors in China are affiliated with the best investment institutions in the global market, it is possible that these investors can process market-level news, especially global news, better than their local counterparts. Following the rationale developed in Bae et al. (2012), if foreign investors are capable of processing market-level information, their order flows would predict market-level returns in the time series, and their cross-sectional return predictive power would be higher on market news days than on non-news days. On the other hand, market-level information, such as key economic indicator releases, is highly confidential and hard to predict. It may be much more difficult for foreign investors to predict either market returns or firm-level returns on important market news days. Therefore, we establish our third hypothesis:

Hypothesis 3: Foreign investors are able to process market-level news, especially global news. That is, foreign order flows can predict market returns, and the cross-sectional predictive power of foreign order flows is stronger on market-level news days relative to non-news days.

Finally, we associate foreign investors' informativeness with market liberalization. Even though Chinese regulators generally welcome foreign capital, they cautiously design the regulations through investment quotas, eligible stock pools and currency transfers to gradually facilitate foreign investors' participation. How does this evolution of relaxed regulation relate to the degree to which foreign investors possess informational advantages in the local market? Fewer restrictions on foreign capital may lower the potential cost of foreign investment and attract more sophisticated overseas investors, thereby enhancing foreign investors' overall return prediction capacity. In contrast, we acknowledge that a friendlier investment environment could make it easier for less informed or passive investors to access the domestic stock market. We let the empirical results uncover which hypothesis fits the data better. We propose our fourth and final hypothesis regarding regulations:

Hypothesis 4: The relaxation of restrictions on foreign investors improves their return predictive power in the Chinese stock market.

III. Institutional Background and Data

A. Foreign Investors in the Chinese Stock Market

Foreign investors invest in the Chinese onshore stock market mainly through three programs: QFII, RQFII, and HKC. As investment channels for foreign capital, the three programs share common goals yet differ in several aspects, such as investor eligibility, investment scope and capital control, which may lead to distinctive performance patterns in the Chinese stock market. [5]

We summarize the key differences in Table I. First, in terms of investor eligibility, QFII and RQFII include only foreign institutional investors, whereas HKC includes both individual and institutional investors from both Hong Kong and oversea areas. It is worth noting that foreign investors through QFII must meet certain thresholds on assets under management and operational durations. As a result, most of the QFIIs are large and renowned institutions in global capital markets, such as Barclays Bank and Goldman Sachs. In contrast, RQFII was created in 2011 to expedite offshore RMB business, and it was only available to Hong Kong subsidiaries of domestic financial institutions (such as Huaxia Fund Management HK) and foreign institutional investors (such as Fidelity HK). [6] Therefore, especially at its early stages, the RQFIIs include many institutions intending to attract offshore RMBs, rather than pursuing superior investment performance. The HKC program provides access for both institutional investors and retail investors; while international asset management companies (e.g. J.P. Morgan China A-share Funds) and overseas brokers backed by hedge funds are the main HKC investors, retail trading accounts for only a small portion of the HKC program.[7] Given the small proportion of retail investors through the HKC program, we treat the HKC order flows as representative of institutional investors in later discussions.

Second, foreign investors from different channels are subject to different capital control regulations. To promote long-term involvement, QFIIs and RQFIIs are subject to a 3-month lock-up period, and QFIIs can only repatriate investment principal and profit monthly, up to 20% of the

[5] It is possible that some foreign institutions access the Chinese stock market through multiple programs. How sophisticated institutions strategically optimize over the three programs is not the focus of our study.
[6] Out of the 230 RQFIIs, 152 are subsidiaries of foreign institutions.
[7] From the speech of Fang Xinghai, the vice chairman of China Security Regulation Commission, on April 19, 2021.

previous year's total assets. Capital inflow and outflow, however, are not a concern for HKC investors, meaning that they can more easily enter and exit the Chinese domestic market in short periods. Therefore, QFIIs and RQFIIs presumably have lower turnovers and focus more on long-term returns than HKC investors. In addition, there are investment quotas on individual QFII/RQFII/HKC investors, as well as certain aggregate restrictions across all program participants. To lower the regulation costs on foreign investors, the quotas are generally set at relatively high numbers (and are often not binding).

Third, the eligible stocks are different across the QFII, RQFII and HKC programs. QFIIs and RQFIIs are both allowed to invest in all A-share stocks listed on exchanges, fixed-income securities, and other financial products. In contrast, HKC investors can only trade the constituent stocks of the SSE 180 Index and the SSE 380 Index, as well as all the SSE-listed A shares with H shares listed on the Hong Kong Stock Exchange. The broad scope of financial instruments available for QFII and RQFII may attract large asset management companies who have multi-asset investment demand, as well as institutions that use derivatives to control risks or perform complex strategies. To ensure that A-share stocks are not primarily owned by foreigners, there is an upper limit, in the sense that all three types of foreign investors combined cannot hold more than 30% of a firm's total shares outstanding.

Because of these differences, foreign investors in the three programs may have different trading patterns and investment skills. Given the stricter eligibility requirements, tighter restrictions on capital flows, and wider investment scope, QFIIs are likely to be sophisticated investors, focusing on long-term performance and fundamentals. In comparison, given the capital controls, RQFIIs are also likely to be long-term investors but may be less sophisticated because many are Hong Kong subsidiaries whose primary goal is the absorption of offshore RMB. The

HKC investors are not subject to strict capital controls and can trade more freely over short horizons, so their investment horizon could be shorter.

Over our sample period of 2016 to 2019, regulators gradually remove restrictions on quotas and capital controls. We summarize these changes in Figure 1. For instance, restrictions on capital repatriation for QFII and RQFII were removed in June 2018, and the investment quota was gradually increased and eventually lifted in May 2020. The process of market liberalization offers us a unique opportunity to examine the impact of liberalization on the evolution of foreign investors' behavior.

B. Data

Our sample period spans January 1, 2016 to June 30, 2019. Shanghai Stock Exchange generously provide data on foreign investors' daily trading and holding. We obtain other stock trading data and financial accounting information from WIND, a widely used Chinese financial database. As in Liu, Stambaugh and Yuan (2019), we exclude stocks with less than 15 non-zero volume trading days in the past month to eliminate the influence of long-trading suspensions. After merging the SSE data with the WIND data, we obtain a sample of approximately 1.1 million stock-day observations for over 1,200 stocks and 849 trading days.

For each stock each day, we collect buy and sell data for different groups of investors. Given that most of the foreign investors are institutional investors, we also collect information on local institutional investors to serve as a comparison benchmark. For our purposes, local institutional investors include mutual funds, hedge funds, insurance companies, security companies, trust companies, and other institutional investors. We rely on investor order imbalances data to measure their trading activities. Following Boehmer et al. (2021), we compute investor group G's order imbalance for stock i on day d as follows:

$$Oib(i,d,G) = \frac{Buyvol(i,d,G) - Sellvol(i,d,G)}{Buyvol(i,d,G) + Sellvol(i,d,G)}, \tag{1}$$

where $Buyvol(i,d,G)$ and $Sellvol(i,d,G)$ represent the total number of shares bought and sold by all investors within group G. The variable $Oib(i,d,G)$ captures the trading direction of the investor group G for this stock, and its value varies between -1 and 1. A positive number means that investors buy more than sell, and a negative number means that investors sell more than buy. The order imbalance variable is set to missing when there is no stock trading on that day.

C. Summary Statistics

We present summary statistics in Table II. Panel A reports the trade and holding data of foreign investors and local institutions. One special feature of the Chinese stock market is that retail investors contribute 80% of daily trading volumes over our sample period, so institutional investors, foreign and local, only account for about 20% of daily trading volumes. For average daily trading volumes, QFII, RQFII, and HKC investors account for 0.79%, 0.08%, and 2.24% of market daily volume, respectively, while local institutional investors account for 14.80%. These statistics indicate that QFII and HKC investors, as well as local investors, are relatively more active in trading, whereas RQFII investors tend to trade less frequently. The QFII, RQFII and HKC investors trade 946, 174 and 561 stocks per day, and the lower number of stocks traded by HKC is mostly a result of the investment constraints imposed by the regulators. The holdings of the QFII, RQFII and HKC investors account for 0.95%, 0.23%, and 1.20% of market floating capitalization, and the local investors account for 14.19%.[8] We report the time series of foreign investors' aggregate trading and holding in Figure 2. The trading volume and holdings of QFIIs and RQFIIs

[8] Table IA.I of Internet Appendix provides more details on foreign investors' investment preferences. The foreign investors, as well as local institutions, tend to trade and hold stocks with large size, high earnings-to-price ratio, and low turnover. Additionally, manufacturing sector has the largest holding and trading for foreign investors. The education sector has the lowest holding and trading for both foreign and local investors.

are relatively stable. HKC becomes considerably more important over time, with trading volume and holdings steadily increasing.

Since the focus of our study is on the cross-sectional trading behaviors of foreign investors, Table II Panel B reports the time-series average of cross-sectional statistics on the order imbalance measure. The means of order imbalance for QFII, RQFII, and HKC are, respectively, -0.01, 0.02, and 0.02, with standard deviations at 0.86, 0.82, and 0.58. It is possible that for one particular stock, the trades are likely concentrated in one direction, causing the order imbalance measure to take values close to 1 or -1, which leads to the relatively large cross-sectional variation in QFII and RQFII order flows. In comparison, the mean of the order imbalance for local institutions is -0.01 with a standard deviation of 0.47, indicating that domestic investors' trading dispersion across stocks is smaller than those of foreign investors. The last column reports the cross-sectional mean of the first-order autocorrelation of the order imbalance measure. The coefficients are 0.09, 0.44, 0.12, and 0.18 for QFII, RQFII, HKC, and local institutions, respectively, which suggests that RQFIIs display a more persistent trading propensity than other investors. In the last three columns, we report the time-series average of the cross-sectional correlation coefficients for order imbalance measures across four investor groups. The order imbalances of all investor groups are positively correlated, implying that trades from different types of investors may overlap to some extent. However, the correlations are generally lower than 0.14, indicating that investors' trading behaviors are quite different across groups.

IV. Empirical Results

A. Predicting Future Stock Returns Using Foreign Order Flows

To investigate the return predictive power of foreign order flows in the cross-section as in Hypothesis 1, we adopt the two-stage Fama-MacBeth (1973) regression method. We use daily horizons as an example, and extend to longer horizons using modifications. At the first stage, we estimate the following specification for each group G on each day d:

$$Ret(i,d) = a0(d,G) + a1(d,G)Oib(i,d-1,G) + a2(d,G)'Controls(i,d-1) + \epsilon(i,d,G). \tag{2}$$

The dependent variable $Ret(i,d)$ is the dividend and split-adjusted daily return for stock i on day d, which is expressed as a percentage in our dataset. The main independent variable is investor type G's order imbalance from the previous day, $Oib(i,d-1,G)$. For control variables, we follow the previous literature and include the previous day's stock return $Ret(i,d-1)$, the previous weekly cumulative return $Ret(i,d-6,d-2)$, the previous monthly cumulative return $Ret(i,d-27,d-7)$, log firm size ($Lnsize$) from the previous month-end, firm earnings to price ratio (EP) as the ratio of most recently reported quarterly earnings to the market capitalization from the previous month-end, and turnover ($Turnover$) as the ratio of monthly trading volume to floating A shares from the previous month-end.

From the first stage estimation, we obtain a time-series of the cross-sectional coefficients $\{\widehat{a0}(d,G), \widehat{a1}(d,G), \widehat{a2}'(d,G)\}$. In the second stage, we compute means and standard errors and conduct inference using the time series of these coefficients. The standard errors are calculated using the Newey-West (1987) methodology with five lags, the optimal lag number under the Bayesian information criterion. If a particular group G of foreign investors' order flow correctly predicts future stock returns, we expect a significantly positive average coefficient $\widehat{a1}(G)$. An insignificant coefficient of $\widehat{a1}(G)$ indicates no predictive power, and a significant and negative coefficient $\widehat{a1}(G)$ implies that the foreign investors' trades are, on average, opposite to future stock price movements.

A.1. Predictive Power over Daily Horizons

Table III Panel A presents the estimation results of equation (2). For QFII, the coefficient on Oib is 0.0649 (*t*-statistic=17.02), implying that QFII's order flow significantly and correctly predicts future stock returns. In terms of the magnitude, given the interquartile of QFII order flow is 1.8295, when we move from the 25th to the 75th percentile, the next day's return increases by 1.8295*0.0649*0.01=0.1188% (29.94% annualized).[9] In terms of RQFII and HKC, the coefficients on Oib are 0.0247 and 0.0783, both with significant *t*-statistics, corresponding to daily interquartile returns of 0.0305% and 0.0757% (7.69% and 19.08% annualized), respectively. The coefficient on Oib for local institutions is 0.1330 (*t*-statistic=18.57), and the daily interquartile return is 0.0933% (23.51% annualized). These results provide support to Hypothesis 1 that, on average, all three types of foreign investors' order flows correctly predict the next day's stock returns with interquartile returns comparable to one another.

We also examine whether the predictive power of foreign investors that we document is comparable with that exhibited by local institutions. Specifically, we compute the time series of the interquartile returns for each group of investors and compare whether their differences are significantly different from zero. That is, we multiply the time-series of coefficients $\widehat{a1}(d, G)$ by investor *G*'s interquartile range of order flow and obtain the time-series of interquartile returns. At the bottom of Panel A, we report the mean of time-series interquartile return differences between different foreign investors and the local institutions (the benchmark), with the *t*-statistics adjusted following Newey and West (1987) with five lags. The time-series average of the interquartile return difference between QFII and local institutions is 0.0255% per day (or 0.0255%*252day =

[9] As an alternative to an interquartile return, we also consider a normalized order imbalance measure as an independent variables. The results are similar and available on request.

6.43% per year), with a t-statistic of 3.29. That is, the predictive power of the QFII order flows seems to be higher than the local institutions. For the RQFII and HKC order flows, their predictive power is significantly lower than local institutions, with daily differences in interquartile returns being -0.0626% and -0.0184%, respectively. This simple comparison shows that QFII has the highest interquartile returns, local institutions the second, HKC the third, and RQFII the lowest.

In comparison with findings in the literature that foreign investors have limited informational advantages in Emerging East Asia, this result is surprising. For example, Froot, O'Connell and Seaholes (2001) find that foreign portfolio flows have insignificant predictive power on future equity returns at short and long horizons in markets such as Hong Kong, Indonesia, and Korea. Using Chinese market data from 2000-2010, Ferreira et al. (2017) show a portfolio sorted by local institutional ownership earns 0.65% (t-statistic=1.53) higher monthly excess returns than a portfolio sorted by foreign institutional ownership. In sharp contrast to these studies, we find, using comprehensive trading records over daily horizons, that foreign investors such as QFII are not at disadvantage in the local market, suggesting that they may, in fact, possess informational advantages in the Chinese stock market.

For the control variables, we find significantly negative coefficients on $Ret(d-6, d-2)$ and $Ret(d-27, d-7)$ in most specifications, suggesting strong reversal patterns in stock returns over the weekly and monthly horizons. In our sample period, while the size effect is insignificant, we find that stocks with high earnings-to-price ratios exhibit larger future returns, consistent with the value effect. While the coefficients on turnover are most negative, consistent with the hypothesis that high trading volume might be driven by speculation and lower future lower returns. The average adjusted R^2s from the first-stage OLS regressions range from 8.83% to 14.75%.[10]

A.2. Predictive Power over Longer Horizons

Given the strong one-day prediction for stock returns, we examine whether the predictive power remains over longer horizons. We modify the benchmark regression in equation (2), by using weekly returns, $Ret(i, w)$, as the dependent variable, with w ranging from 1 to 12. For instance, when w equals 1, $Ret(i, w)$ represents the cumulative stock return from $d + 1$ to $d + 5$; when w equals 2, $Ret(i, w)$ represents the cumulative return from $d + 1$ to $d + 10$, and so on. The independent variable and the control variables are the same as those in equation (2). Standard errors are adjusted following Newey and West (1987) with five lags. If foreign investors' predictive power extends to longer horizons, a positive and significant $\widehat{a1}(G)$ is expected.

Table III Panel B presents the estimation results. To save space, we only report the coefficients on *Oib*, and the implied interquartile cumulative returns. The statistical significance levels are denoted by asterisks, with ***, **, and * indicating significance at 1%, 5%, and 10%, respectively. Take QFII as an example; the coefficient for *Oib* is 0.1123 at week 1, and gradually increases to 0.2507 at week 12, with no sign of reversal. All coefficients differ from zero at the 1% significance level, indicating that order flows from QFIIs significantly predict returns over longer horizons, and possibly their return predictive power is related to long-term information, such as firm fundamentals, rather than short-term information, such as price pressure. The patterns are similar for RQFII and local institutions. For HKC, the coefficient climbs from 0.0985 at week 1 to

[10] In Figure IA.1, we show the time-series coefficients $\widehat{a1}(d, G)$ to ensure that there are no outliers in the cross-sectional regressions over time. The time series are stable and do not display extreme values.

0.1874 at week 8, then declines to 0.1677 at week 12, indicating a slight price reversal. Information contained in HKC order flows can likely be relatively short-term.

To compare the economic magnitude of the predictive power of various investors over longer horizons, we present the cumulative interquartile returns over the next 12 weeks at the bottom of Panel B. For a heuristic understanding of the magnitudes and trends, we also directly plot the interquartile return differences predicted by the order flows from different investors in Figure 3. We observe the following three patterns. First, all four lines generally trend up and do not present major reversals over 12 weeks (except there is a slight flattening pattern for HKC order flows), suggesting that the predictive power of foreign and local institutions' order flows is lasting rather than transient. Second, the interquartile returns for QFII and local institutions are quite close to each other throughout the 12 weeks, and both are larger than that of RQFII and HKC. Our daily results in Table II show that QFII has stronger return predictive power than local institutions over the next day. From the bottom of Table III and Panel B, this advantage of QFII over local institutions remains over week 1 but becomes statistically insignificant. For the next 11 weeks, the performances of QFII and local institutions are similar and do not exhibit differences with statistical significance. That is, the predictive power of QFII and local institutions are comparable over the 12-week horizon. Third, RQFII and HKC have lower predictive power than QFII, and between the two, HKC has stronger predictive power than RQFII over week 1; but starting from week 2, RQFII performs better than HKC over the next 11 weeks. If we look across all three foreign capital channels, it is clear that foreign investors' performance differences are related to their institutional background. As QFII has the strictest eligibility requirements, the tightest restrictions on capital flows over longer periods, and the widest investment scope, they may disproportionately be large international institutions focusing on long-term investments. RQFIIs face similar regulation settings to QFIIs, suggesting they may too largely be long-term investment institutions. However, RQFIIs may be somewhat less sophisticated because many are local institutions' Hong Kong subsidiaries whose primary goal is to absorb offshore RMB. For HKC, cross-border flows are much easier and less restricted, which may attract more short-term investors and lead to lower long-term return predictive power of order flows.[11]

A.3. Counterparties

Given the strong predictive power of foreign investors' trading for future stock returns in Chinese stock market, an intriguing and important question arises: against whom do they trade? Local retail investors or local institutions? Over a similar sample period, Jones et al. (2022) investigate the stock trading behavior of Chinese retail investors, and find that smaller retail investors' trading activity *negatively* predict future stock returns. Is it true that foreign investors disproportionately trade against local retail investors? To answer this question, we need to identify the counterparties in each and every trade; given obvious data limitations, this is difficult to achieve in the previous literature.[12] For this study, we are fortunate to have access to a level of detail from the exchange that facilitates the identification of trades from different groups of investors.

We separate counterparties into three groups: foreign investors, local institutions and retail investors (RT). Next, we separate investors' daily trade directions into buy (B) and sell (S), based on the end-of-day order imbalance measure. With the three groups of investors and the two sides of each trade, all stock-day observations are separated into six bins: BBS, SSB, BSS, SBB, BSB and SBS, with the first letter indicating the trade direction of foreign investors, the second for local institutions, and the third for retail investors. For example, the first bin "BBS" means foreign

[11] We provide a robustness check for the long-term predictive patterns using risk adjusted returns. The empirical findings are quite similar and are reported in the Internet Appendix Table IA.II.

[12] Early studies with transaction data such as Bailey, Wang and Sirodom (2007), Dvořák (2005) and Grinblatt and Keloharju (2000) do not investigate the counterparties of foreign investors as we do in this section.

investors and local institutions are both net buyers while retail investors are net sellers. In Table IV Panel A, we present the proportion of trades in the six groups. Take QFII as an example. The proportions of trades in "BBS", "SSB", "BSS", "SBB" "BSB" and "SBS" are 27%, 26%, 4%, 4%, 19% and 19% respectively, implying that QFII trade against local retail investors 61%(= 27%+26%+4%+4%) of the time, while QFII trade against local institutions 47%%(=4%+4%+19%+19%) of the time. Similar patterns hold for RQFII and HKC.

To examine whether foreign investors' return predictive patterns change with different counterparties, we modify equation (2) as follows:

$$Ret(i,d) = a0(d,G) + [\sum_{k=1}^{6} a1(k,d,G) * I(k,d,G)]Oib(i,d-1,G) + a2(d,G)'Controls(i,d-1) + \epsilon(i,d,G). \tag{3}$$

The indicator variable *I(k,d,G)* is equal to one if trading from foreign investor group *G* for stock *i* on day *d-1* falls in the *k*-th counterparty bin, otherwise it is zero. We interact the six dummy variables with investors' order imbalances in the previous day and estimate Fama-MacBeth regressions. The estimated coefficients $\widehat{a1}(1,G)$ and $\widehat{a1}(2,G)$ are assumed to be significantly positive because in the two groups both foreign investors and local institutions are on the opposite side of uninformed retail investors.

Table IV Panel B presents the estimation results. Take QFII as an example. First, when they trade on the same side with local institutions and on the opposite side of local retail investors, as in the bins of BBS and SSB, the coefficients are positive and significant, indicating that both foreign and local institutions have an informational advantage over local retail investors. Second, when QFII trade against both local institutions and local retail investors, as in BSS and SBB, their predictive power is a more mixed positive or negative. Similarly, when QFII trade on the same side of retail investors but against local institutions, as in BSB and SBS, their predictive power is also more mixed than the first case. These findings imply that when foreign investors are on the same side of local institutions, their predictive power is, on average, more consistent and significant. There are also cases that when foreign investors trade differently from local institutions, they can still significantly and positively predict future returns, as in SBB and SBS, which suggest foreign investors may possess specific informational advantages in the local market relative to local institutions. Similar patterns also hold for RQFII and HKC. [13]

A.4. Account Performance of Foreign Investors

Given the strong predictive power of foreign investors for local returns, we also examine whether foreign investors make money in the Chinese stock market. Given our detailed data on holding and trading, we design a methodology to obtain an estimate of aggregate performance for foreign investors and local institutions. On each trading day, we first compute the total cash flow for each investor group, by adding up the holdings and the day's trading:

$$\begin{aligned} Total(d,G) = \sum_{i=1}^{N} \Big[& HoldShares(i,d,G) * \big(close(i,d) + dividend(i,d)\big) \\ & - HoldShares(i,d-1,G) * close(i,d-1) \Big] \\ & + \sum_{i} [SellShare(i,d,G) * SellPrice(i,d,G) - BuyShares(i,d,G) \\ & * BuyPrice(i,d,G)] - \sum_{i} TrdCost(i,d,G). \end{aligned} \tag{4}$$

The first component is tied to the capital gain. Here $HoldShares(i,d,G)$ is the total holding shares of investor group G for stock i on day d; $close(i,d)$ is the close price for stock i on day d ; and $dividend(i,d)$ is the cash payouts. So, the capital gain is the difference between day d's holding

[13] As estimating regressions for RQFII, we find in 24 cross-sectional regressions, the six interaction variables and the constant are highly colinear, leading to extreme estimations. Therefore, we drop the 24 estimated cross-sectional coefficients when calculating the time-series average.

value and day d-1's holding value, which is $Shares(i, d-1, G) * close(i, d-1)$. The second component is tied to active trading over day d, where the cash inflow is computed as the selling proceeds, $SellShares * SellPrice$, less the purchase cost, $BuyShares * BuyPrice$. The last component is the transaction cost, $TrdCost$, which include the commission cost (0.05% imposed on both the buy and sell side, with a minimum of 5 CNY for each trade), the stamp tax (0.10% of the sales amount), and the transfer fee (0.002% imposed on both sides).

Table V presents the results. The annual gain for QFII, RQFII, HKC and local institutions are 42.82, 12.53, 63.67 and 411.12 billion RMB, respectively. Clearly, foreign investors groups earn positive cash flows in the Chinese A-share market during our sample period, consistent with our earlier results regarding their predictive power for returns. If we compare the magnitudes, the average annual cash flow to QFII, at 42.82 billion RMB, is roughly 10% of what local institutions earn, at 411.12 RMB, consistent with the relative holding and trading magnitudes that we observe in the market. To have a sense of a percentage return on investment, we calculate the account percentage performance by dividing the aggregate cash flows by the daily average of investors' aggregate holdings as presented in Table II Panel A. The annual returns for QFII, RQFII, HKC and local institutions are 17.83% (42.82/240.23), 21.61% (12.53/58.01), 20.46% (63.67/311.14) and 11.45% (411.12/3590.2), respectively, suggesting notable return performance in each case. [14]

A.5. Separating Order Flows Among Different Groups of Investors

Our results thus far show that foreign investors, such as QFII, perform similarly to local institutions. It is possible that foreign and local investors share overlapping information, so that they trade similarly, leading to similar predictive patterns. It is also possible that they possess

[14] Following Barber et al (2009), we also take an alternative perspective by decomposing the total cash flow into three parts: market timing, stock selection and transaction costs. Interestingly, using their methodology, we find that stock selection ability contributes to the majority of our documented return performance. The results are available on request.

different information, and they have similar magnitudes of predictive power by coincidence. To find out whether the information is mostly overlapping or largely unique among different groups of investors, we orthogonalize each group's order flow with respect to another group's order flow and examine the residual's predictive power for future returns. For instance, for each day d, we project foreign investors' order flows onto local institutions' order flows as follows,

$$Oib\left(i, d, G_{Foreign}\right) = b0(d, G) + b1(d, G)Oib(i, d, G_{Local}) + \epsilon(i, d, G). \tag{5}$$

After we obtain the time-series of $\widehat{b1}(d, G)$, we decompose the foreign order flow into two parts,

$$Oib_{i,d,G_{Foreign}}^{overlap} = \widehat{b1}(d, G)Oib(i, d, G_{Domestic}), Oib_{i,d,G_{Foreign}}^{specific} = \widehat{b0}(d, G) + \hat{\epsilon}(i, d, G), \tag{6}$$

with the first term being the overlapping component with local institutions, and the second term being the foreign-specific component. After the decomposition, we re-estimate equation (2) by including both components. Similar procedures are followed to decompose local institutions' order imbalance with respect to order flows from all three foreign investor groups.

Table VI reports the estimation results. For QFII, the coefficients on $Oib^{overlap}$ and $Oib^{specific}$ are 0.3553 and 0.0593, with *t*-statistics of 0.18 and 15.67, respectively, indicating that QFII's predictive power mostly stems from foreign-specific information rather than from the overlapping component with the local institutional order flows. In terms of economic magnitude, the daily interquartile returns for overlapping and foreign specific order flows are 0.0399% and 0.1034%, respectively, indicating that the foreign specific information in order flow contributes more to QFII's performance. Similar patterns are observed for RQFII, HKC, in the sense that only the foreign specific order imbalance displays significant return predictive power. In contrast, the pattern is different for local institutions, where both the overlapping and local-specific components of order flows significantly predict future stock returns. In terms of economic magnitude, the interquartile return for the overlapping component is 0.0769%, somewhat smaller than the interquartile return of 0.1205% driven by the local-specific order flows. [15]

Our findings that foreign-specific order flows contribute more to foreign investors' return predictive power, especially for QFII, suggest that foreign investors may possess unique informational advantages in the local stock market. These may reflect foreign investors' ability to correctly process local or global information. Next, we examine the information contained in order flows.

B. Firm Information and Return Predictive Power

B.1. Foreign Order Flows and Earnings Information

The prior literature reveals that earnings announcements and analyst-related activities are related to firm fundamentals and have implications for future stock price movements (Bradley et al., 2014; Savor and Wilson, 2016). We obtain earnings announcement data from WIND. For analyst data, though CSMAR is a widely used analyst database (Dong et al., 2021; Chen et al., 2022), its coverage is incomplete, particularly in earlier periods. Following Li, Wong and Yu (2020), we construct a comprehensive analyst sample from four major data providers: CSMAR, WIND, RESSET, and SUNTIME. [16] For analysts' activities, we focus on forecast revisions and recommendation changes. Our sample includes 15,477 earnings announcements and 41,722 analyst-related events, totaling 50,331 event days for individual stocks, accounting for 4.94% of all stock days in our sample.

[15] We also include order imbalances from all four investor groups together in one regression to predict the next day's return. The caveat of this setup is that all order flow variables need to have a non-missing value for the stock on that day, which excludes more than 90% of our sample stocks. For this much smaller sample, results are presented in Table IA.III shows that all four Oib variables have positive and significant coefficients.

[16] The dataset construction details are provided in the Internet Appendix A.

We adopt two approaches to investigate whether foreign investors can process firm-level news. In the first approach, we examine whether the order imbalance of foreign investors from the previous trading day can predict earnings news on the next day. The earnings news is computed using the cumulative abnormal returns (CAR) over different horizons, by subtracting the same period market returns. Given that earnings are announced quarterly and earnings news is not evenly distributed over calendar days. We modify the daily Fama-MacBeth setup to a quarterly horizon, where we first estimate the cross-sectional regression for each quarter q,

$$CAR(i,d,d+k) = c0(q,G) + c1(q,G)Oib(i,d-1,G) + c2(q,G)'Controls(i,d-1) + \epsilon(i,d,G). \tag{7}$$

All controls are the same as in equation (2). After we obtain the quarterly coefficients, we calculate the time-series average and conduct inferences accordingly. If coefficient $c1$ is significantly positive, it implies that the previous day's order imbalance predicts the future earnings news correctly, which supports Hypothesis 2.

Table VII Panel A presents the estimation results for equations (7). For abnormal return on the event day 0, AR[0], the coefficient for QFII Oib is 0.0878, with a significant *t*-statistic of 3.20, suggesting that previous day QFII order flows predict earnings news and the market’s response to the news correctly and significantly. When we extend the horizon to two days using CAR[0,1], one quarter using CAR[0,61], and one year CAR[0,251], the corresponding coefficients are all significantly positive and the interquartile returns gradually increase to 1.2327%. We find similar but weaker results for RQFIIs and HKC investors. The local institutions in the last column display similar predictive patterns to those of QFIIs, but with larger magnitudes, indicating that local institutions seem to process local earnings news better. These results provide direct evidence that order flows from foreign investors contain long-term earnings information, which supports Hypothesis 2.

For the second approach, if investors can access, or anticipate, the information contained in earnings news, their order flows ahead of events should have greater return predictive power on event days than on non-event days. Notice that some news is fully expected, and hence leads to no or little reaction in realized returns, whereas other news items are unexpected and lead to large reactions in returns. For instance, Jiang and Zhu (2017) use large stock price jumps to identify significant information events. Therefore, for the second exercise, we first compute the 5th and 95th percentiles of event day returns across all firms and all days to separate the largest reactions of returns to the information, which also indicates that these events are most value relevant. We define an indicator $Bignews(i,d)$, which is equal to 1 if stock i's return on event day d is outside of these 5th and 95th percentiles, and otherwise it is zero. Similarly, we define another indicator, $NBignews(i,d)$, which is equal to 1 if on event day d, stock i's return is within the 5th and 95th percentiles, and otherwise it is zero. Empirically, we separately estimate the predictive power of order flows for future returns on the most and least value-relevant events in the following design for each quarter q:

$$\begin{aligned} Ret(i,d) = {} & e0(q,G) \\ & + [e1(q,G) + e2(q,G)\ Bignews(i,d) + e3(q,G) NBignews(i,d)] \\ & \times Oib(i,d-1,G) + e4(q,G)' Controls(i,d) + \epsilon(i,d,G). \end{aligned} \tag{8}$$

Here, investors' order flows interact with the two indicators to allow the predictive power to differ on the most and least value-relevant events. If the next day is a non-event day, $e1(G)$ captures the predictive relation between order flows and future returns. If the next day is an event-day with large movements in prices, $e1(G) + e2(G)$ captures the predictive relation between order flows and future returns. Similarly, if the next day is an event-day but not with large movements in prices, $e1(G) + e3(G)$ captures the predictive relation between order flows and future returns. Positive coefficient estimates, $e2(G)$ and $e3(G)$, indicate that investor group G has higher return predictive power on event days than on non-event days, suggesting that investors can process firm information regardless of the content quality. The differences in coefficients, $e2(G)$ and $e3(G)$, tells us whether the investors can process information related to large price movements or not.

Panel B provides results for the estimation of equation (8). For QFII, the $\widehat{e1}$, $\widehat{e2}$ and $\widehat{e3}$ coefficients are 0.0977, 0.5177 and -0.0342 respectively, all significant at the 99% confidence level. The interquartile return on non-event days is 0.0977*1.8295*0.01=0.1787%, and the interquartile return on event days with large price changes is (0.0977+0.5177)*1.8295*0.01=1.1259%, and the interquartile return on event days with small price changes is (0.0977-0.0342) *1.8295*0.01=0.1161%. That is, the predictive power of QFII for future stock returns is quite similar before non-event days and event days with no large price changes, while before event days with large returns, their predictive power is almost six times higher. These results show that QFIIs can anticipate or have access to firm information when the most value-relevant news becomes public the next day. Similar patterns are observed for order flows from RQFIIs, HKC and local institutions. In terms of economic magnitude, computed using interquartile returns, QFII has the strongest return predictive power on the most value-relevant news days across the three foreign investor groups.

Following Boehmer et al. (2020), we gauge the importance of firm events to investors' overall performance using the fact that 0.49% of the sample are events with large price changes and 4.45% of the sample are events with small price changes. Take QFII as an example. The overall performance is the sum of interquartile returns on event days multiplied by the percentage of event days in the total sample, 0.1787%*(1-4.94%)+1.1259%*0.49%+0.1161%*4.45%= 0.1806%. Thus, event days with large price changes account for (1.1259%*0.49%)/0.1806%=3.06% of the overall performance, and other event days account for (0.1161%*4.45%)/0.1806%=2.86%. Similarly, the contribution of the most valuable event days for RQFII, HKC and local institutions is 7.38%, 1.94% and 5.83%, and the contribution of least valuable event days is 6.73%, 7.87%, and 4.13%, respectively. Except for HKC, the most valuable events contribute much more to overall performance than the least valuable events. The results indicate that events with large price changes are important sources of investors' return predictive power.

Overall, these results support Hypothesis 2, as we find that foreign investors are capable of processing local firm information related to earnings announcements and analyst activity, especially regarding events leading to large price movements.

B.2. Media News

The previous literature also shows that press coverage contains uncovered content regarding a firm's fundamentals that can be used to predict future stock returns (Tetlock, Saar-Tsechansky and Macskassy, 2008). In this section, we collect news data from the Chinese Research Data Service Platform's Financial News Database of Chinese Listed Companies (CFND). CFND gathers financial news from over 400 websites and 600 newspapers, including reports from 20 mainstream online financial media outlets and China's eight largest national business newspapers, all written in Chinese. Using the same database, Ge and Zhang (2022) show that the news tone can correctly predict stock returns in both short and long horizons, implying that news contains valuable information on stock prices. Our sample contains 353,551 firm news days, accounting for 34.69% of total observations.

We estimate similar regressions as in equations (7) and (8) by focusing on media news days, rather than earnings news days. The results are presented in Table VIII Panel A and B. Take QFII as an example. In Panel A, the coefficient for AR[0] is 0.0798, with is positive and significant, implying an interquartile return of 0.1570%. That is, the previous day's order flow from QFIIs can correctly predict the next news day's return, suggesting that the QFIIs can anticipate media news or the market's reaction to the news, which supports Hypothesis 2. In Panel B, the $\widehat{e1}$, $\widehat{e2}$ and $\widehat{e3}$ coefficients are 0.0906, 0.3550 and -0.0085 respectively. All except $\widehat{e3}$ are significant at the 99% confidence level. The interquartile return on non-news days is 0.0906*1.8295*0.01=0.1657%, and the interquartile return on news days with large price changes is (0.0906+0.3550)*1.8295*0.01=0.8153%, and the interquartile return on news days with small price changes is (0.0906-0.0085)*1.8295*0.01=0.1502%. The results suggest that order flows from QFIIs have predictive power for future stock returns, especially for news days with large price movements. For the contribution of public firm news to QFII's performance, we calculate the overall performance for QFII as 0.1657%*(1-34.69%)+0.8153%*3.47%+0.1502*31.23%=0.1834% daily. News days with large price movements contribute (0.8153%*3.47%)/0.1834%=15.42% to overall performance, and news days with small price movements contribute (0.1502*31.23%)/0.1834%=25.57%.

For RQFII and HKC, we find smaller and less significant coefficients, indicating that their predictive powers are not significantly different on media news days. The predictive power of local institutions' order flows is quite similar to that of QFIIs. The results imply that financial media news associated with large price movements significantly contributes to the return predictive power that we document throughout the paper. [17] By comparing different groups of investors, we find the order flows from all foreign investors and local institutions have stronger predictive power when there is significant media news on the next day, indicating that they may have access to the news before it is announced, or they can anticipate the news.

B.3. How Foreign Investors Process Local Firm Level Information

Our results show that foreign investors can predict local firm level news, and can process local firm information, especially the most value-relevant information related to large stock price movements. Given the potential information asymmetry induced by geographic distance, this finding might be surprising. In our opinion, many factors might contribute to this finding. First, the information environment in China during our sample period is greatly improved and becomes accessible to foreign investors, and thus doesn't create a significant information barrier for foreign investors with skills and resources. Second, foreign institutions may establish offices in places nearby Mainland China, like Hong Kong. In fact, according to SAFE (State Administration of Foreign Exchange), around 25% of QFIIs and 43% for RQFII locate in Hong Kong. Or, more directly, foreign institutions hire managers of Chinese ethnicity, or China working experiences to overcome cultural and regulatory barriers. To corroborate this, we directly examine the hiring websites of several foreign institutional investors and collect anecdotal evidence. For instance, J.P. Morgan asks research associates for China-related roles to attain "all necessary CSRC licenses"; BlackRock asks searches for a VP for equity research with "5-10 years China equity market

[17] We consider four alternative specifications for firm level news and include them in the Internet Appendix. First, to maintain ease of interpretation in the main text, we do not add event dummy variables. Second, we separate large price movement firm event days into positive return days and negative return days. Third, we separate the earnings news and analyst news. Finally, we identify big news days as days with large returns. From results in Table IA.IV to Table IA.VII, we find event dummy variables does not change any findings; foreign investors tend to have higher return predictive power on days with negative returns; and both foreign investors and local institutions are more capable of processing analyst-related events rather than earnings announcements. Finally, foreign investors have stronger return predictive power on large return days.

experience, preferably in leading financial institutions"; and Goldman Sachs directly asks for "strong communication skills in Chinese (written and verbal)".

Third, Chinese firms are important participants in the global supply chain both as suppliers and, increasingly, consumers. If foreign investors are familiar with the global environment, they might be more familiar with firms with significant cross-border business activities, and thus possess important information about these firms. To examine this possibility, we investigate whether foreign investors' predictive power is related to firm-level overseas activities. In China, public-listed firms disclose overseas revenue in semiannual and annual financial statements. Here, we use the absolute value of the ratio of overseas revenue to total revenue, $|Overseas(i, d-1)|$, as a measure of firms' cross-border business activities. We modify the benchmark regression in equation (2) by interacting the order flow variables with $|Overseas(i, d-1)|$.

Table IX presents the estimation results. For foreign investors and local institutions, the return prediction coefficients on the interaction between investor order imbalance and a firm's overseas revenue are all positive, ranging between 0.0427 and 0.0720, with statistical significance for QFII and HKC. The results suggest that the predictive power of order flows is stronger for firms with more overseas activities. While this may suggest foreign investors are more familiar with these firms or they better understand a global component of their performance, these results are interestingly also true for local institutional investors. In terms of economic magnitudes, the interquartile return for QFII on hypothetical firms with a unit overseas ratio is 0.1925% per day, higher than 0.1327% for local institutions, indicating that foreign investors may possess informational advantages on firms with significant cross-border business.

Overall, regarding the strong and significant predictive power of foreign investors' order flow around firm-level news days, it is likely that the maturing information environment in China, together with the experience and diligence of sophisticated foreign investors, make up for otherwise potential disadvantages of foreign investors in the local market and lead to comparable performance between foreign and local institutions regarding firm-level news days.

C. Market-Level Information and Return Predictive Power

Besides firm-level information, market-level information is also important for asset price determination. Given foreign investors in China are mostly sophisticated international institutions, we anticipate that foreign investors may be able to process market-level news, especially global news. The most informative news at the market level is likely the market return itself, which presumably contains all information happening at the aggregate level. Therefore, we use returns on a value-weighted portfolio of all Chinese A-share stocks as a proxy for local market information, and returns on the MSCI World Index as a proxy for global market information.

To examine Hypothesis 3, we implement two approaches, one focusing on time-series predictive patterns and the other focusing on cross-sectional predictive patterns. For the first approach, we examine whether aggregate foreign order flows can directly predict market-level returns over time. That is, for investor group G on day *d-1*, we first compute the aggregate order flows, $AggOib(d-1, G)$, as the aggregate buy volumes in RMB across all stocks, minus the aggregate sell volumes, divided by the sum of aggregate buy and sell, and we estimate the following time-series predictive regression:

$$MktRet(d, d+k) = f0(G) + f1(G)AggOib(d-1, G) + \epsilon(d, G). \tag{9}$$

According to Hypothesis 3, if order flows contain information relevant for market level return, coefficient $f1$ would be significant and positive. Since market level returns change quickly, we focus on a one-day horizon and a five-day horizon [0, 4]. In addition, following the logic from the previous section, we assume that unexpected valuable market-level news leads to large price movements for market indices. Therefore, we separate days into big market news days and non-big-news days with an indicator, $BignewsMkt(d)$, which equals 1 if the stock market return on day d is outside the 5th and 95th percentiles of all market return days in the sample, and zero otherwise. We estimate a modification of equation (9),

$$MktRet(d) = f0(G) + [f1(G) + f2(G) * BignewsMkt(d)] \times AggOib(d-1, G) + \epsilon(d, G), \tag{9'}$$

where we use returns on the Chinese stock market and MSCI World Index to identify days with big Chinese market news and global market news, separately. In all time-series regressions, the standard errors are adjusted following Newey and West (1987) with five lags.

Table X Panel A presents estimation results of equations (9) and (9'). For market return prediction results of equation (9), QFIIs can correctly predict future market returns over the next day and the next five days. The predictive power is statistically significant over the next day, but becomes marginally significant over the next five days. Similar patterns exist for RQFIIs, HKC and local institutions, but with lower significance. Then we investigate whether investors' predictive power on market returns is stronger on market-level news days. Take QFII as an example. For Chinese stock market news, the $\widehat{f1}$ coefficient is positive but insignificant, indicating relatively weak predictive power on market returns when there is no big market news. However, the $\widehat{f2}$ coefficient on the interaction term is significantly positive, suggesting QFIIs can better predict the next day's market return when there are large Chinese stock market movements. Similar patterns exist for HKC. In terms of global market news, we also find stronger return predictive power on big global market news days for QFII, but not for the other groups of investors. These results provide direct evidence that order flows from foreign investors, especially QFII, can predict market-level returns, especially when there is big news.

For the second approach, we examine whether the cross-sectional predictive power of foreign investors' trade flows, as we document in Section IV.A.1, is stronger on market-level news days than on non-news days. That is, if foreign investors have an informational advantage on market-level news, it might help their cross-sectional predictive power because the market-level news would affect firm-level returns. Then, we examine foreign investors' cross-sectional return predictive power on days with and without large aggregate market price movements. Notice that from equation (2), we obtain the time-series of coefficient $\widehat{a1}(d,G)$ for each day. To understand the dynamics of $\widehat{a1}(d,G)$, we link it to the large market returns as follows:

$$\widehat{a1}(d,G) = h0(G) + h1(G)BignewsMkt(d) + \epsilon(d,G). \tag{10}$$

A positive coefficient $\widehat{h1}(G)$ suggests that investors potentially make use of market-level information to improve their cross-sectional predictive power of *Oib* for cross-sectional returns. Standard errors of the time-series coefficients are adjusted using the Newey-West (1987) methodology, with five lags.

The estimation results for equation (10) are reported in Table X Panel B. We first examine how the Chinese market-level news is related to the predictive power of foreign investors' order flows. The $\widehat{h0}$ coefficient is always positive and significant, showing significant predictive power of all order flow measures for next-day returns when there is no large movement in market returns. For the interaction term, the $\widehat{h1}$ coefficient is insignificant for all foreign investors and is marginally significant for local institutions, but with a negative sign. These results suggest that both foreign and local investors don't seem to have better cross-sectional predictive power on days with large local market movements. In terms of global stock markets, the patterns are somewhat different. First, the $\widehat{h1}$ coefficient is positive for all investor groups, indicating that they may have higher return predictive power on days when the global market experiences large movements. Second, at the 90% confidence level, the $\widehat{h1}$ coefficient is marginally significant for QFII on days with large global market movements. The interquartile returns for QFII on days with global market movement are 0.1566% separately, the highest among our investor groups, indicating they might be at advantage in processing global news. Overall, we find supportive evidence for Hypothesis 3 that foreign investors may be able to process global market information.[18]

When we compare the predictive results on firm-level and market-level news days, it seems that both foreign and local investors possess greater cross-sectional informational advantages regarding firm-specific information than market-level news, which is intriguing. It is possible that market-level news, especially news events related to strong price reactions, are unpredictable shocks and/or reflect highly confidential policies or macroeconomic data releases, and thus it is hard for investors to make precise predictions before announcements.

D. Stock Market Liberalization

During our sample period, China gradually relaxes the QFII, RQFII and HKC regulations to permit greater access for foreign investors. These reforms facilitate the entrance of foreign investors to the Chinese market, which also provides an opportunity for us to examine how foreign investors' return predictive power evolves along with a greater degree of regulatory access.

[18] In addition, in Internet Appendix Table IA.VIII, we use the Citigroup economic surprise index (CESI) as an additional proxy and do not find a significant relation between investors' return predictive power and large surprises. From unreported results, we collect major macro announcement days in the U.S. and China and examine whether order flows can predict returns on these days, and fail to find strong predictive power, which suggests that macroeconomic news are difficult to forecast.

There are three major policy changes for the QFII program in our sample. First, on February 3, 2016, SAFE announced to increase in the maximum basic investment quota for a single QFII from $1 billion to $5 billion. Second, on June 10, 2018, SAFE announced the removal of the 3-month lock-up period and the maximum 20% capital repatriation limitation for QFII. Third, on January 14, 2019, SAFE announced an increase in QFII's total investment quota from $150 billion to $300 billion. There are two major policy changes for the RQFII program. RQFIIs originally were not allowed to invest in stocks or stock investment funds at levels that exceeded 20% of their raised capital. CSRC verbally announced the lifting of that restriction at a press conference on September 30, 2016. Then, on June 11, 2018, SAFE announced the removal of the 3-month lock-up period for RQFII. Finally, there are two regulatory changes for the HKC program. First, on August 16, 2016, the RMB 300 billion aggregated quota was removed. Second, on May 1, 2018, the daily quota increased from RMB 13 billion to RMB 52 billion. Based on these regulations, we define seven regulation dummy variables, $Quota2016QFII(d)$, $FX2018QFII(d)$, $Quota2019QFII(d)$, $Invest2016RQFII(d)$, $FX2018RQFII(d)$, $Quota2016HKC(d)$ and $Quota2018HKC(d)$. Each dummy variable is equal to zero before the related event occurs and one afterward.

To examine the relationship between regulatory reform and foreign investors' return predictive power, we project the time-series of coefficient $\widehat{a1}(d,G)$ from equation (2) on regulation dummies,

$$\widehat{a1}(d,G) = l0(G) + l1(G)'Regulations(d-1) + \epsilon(d,G). \tag{11}$$

According to Hypothesis 4, if relaxing regulations provides greater access or lower transactions cost for foreign investors in China stock market, we expect that foreign investors' predictive power increases after a particular regulation change, implying positive values of the coefficient vector $\widehat{l1}(G)'$.

Table XI reports the estimation results. For QFII, we observe a significantly positive coefficient on *Quota2016QFII*, which indicates an increase in return predictive power since February 2016. We do not find a significant relationship between policy changes and RQFII's return predictive power. For HKC, the coefficients on *Quota2016HKC* and *Quota2018HKC* are 0.0234 and 0.0913 with *t*-statistics of 2.05 and 5.25, meaning that HKC better predicts stock returns after the expansion of investment quotas. Our results imply that to some extent, the relaxation of investment access can improve foreign investors' return predictive power, suggesting that more informed trading is reflected in foreign capital flows in aggregate after the relaxation of regulations, which supports Hypothesis 4.

V. Robustness and Further Discussion

A. Price Pressure

Ferreira et al. (2017) find that foreign investors' predictive power is better explained by a price pressure explanation rather than as an underlying fundamental information explanation. Here we decompose the order flow measures into two components, a 20-day moving average measure, $MAOib$, to capture a persistent component of the order imbalance variable, which is directly linked to price pressure, and the difference between the Oib and the $MAOib$, $ShockOib$, to capture the day-to-day fluctuations in order flows.[19] We then examine whether foreign investors' return

[19] Moving average technique is a simple and reasonable way to obtain persistent components of time series variables. We use a 20-day window because we want to avoid both short-term noises and long-term cycles in the data. Results using 15 days and 30 days are similar and are available on request.

predictive power is driven by the moving average component or the daily change component, by estimating equation (2) using $MAOib$ and $ShockOib$ in the place of Oib.

From Panel A of Table XII, the coefficients of $MAOib$ and $ShockOib$ are all positive and statistically significant, indicating that both persistent and temporary order flows contribute to investors' predictive power. To determine the relative importance of these two components, we separately calculate interquartile returns for the two measures. Take QFII as an example. The interquartile returns for $MAOib$ and $ShockOib$ are 0.0382% and 0.0924%, indicating that the latter contributes more to the predictive power of order flow for future returns than the price pressure component. Similar patterns are also observed for HKC and local institutions. For RQFII, $MAOib$ and $ShockOib$ have quite similar interquartile returns. Put together, the results show that the price pressure component of order flow can predict return significantly, but for most of the foreign and local order flows, it is the shock component of order flow that drives the predictive pattern for future returns.

B. Buy vs. Sell Orders

For our main results, we directly combine buy and sell orders to compute the order imbalance measure. Given that short selling is difficult in China's A-share market[20], buy orders are more likely to establish new long positions rather than closing out existing short positions, while sell orders are more likely to close out existing long positions rather than opening short positions. If that's true, the buy order might contain more information than the sell order. To examine this hypothesis, we separate order imbalance into buy and sell flow, $Buy(i, d, G)$ and $Sell(i, d, G)$,

[20] The average daily short volume in RMB only accounts for 0.47% of total volume in 2021. QFII/RQFII are allowed to do leverage trading only after December 29, 2020.

computed as the buy and sell volume divided by stock's outstanding A-shares, respectively. We then estimate equation (2) by using $Buy(i, d, G)$ and $Sell(i, d, G)$ in place of order imbalance.

Panel B of Table XII presents the results. Take QFII as an example. The coefficient on buy is 1.7667 with a significant t-statistic, and the interquartile return for buy is 0.0112%; while the coefficient on sell is -0.2000 and insignificant, and the interquartile return is -0.0012%. The positive and significant coefficient on investors' buy indicates more buys are associated with higher returns, and negative coefficients on investors' sell indicate more sells are associated with lower returns, both consistent with informative trading. But the interquartile return is much larger for buy than for sell, suggesting investors' buy indeed have stronger return predictive power. We also calculate the economic magnitudes. Similar patterns can be found for other investors, except that the sell side also significantly predicts returns for HKC and local institutions, possibly because they adopt more short-selling in their trades.

C. Eligible Stocks in HKC Program

Given that QFIIs and RQFIIs can invest in any A-share stocks, the results we report so far are based on the entire A-share universe. The restriction is different for HKC investors, who can only invest in constituent stocks of the SSE 180 Index and SSE 380 Index and A shares that have H shares listed on the Hong Kong Stock Exchange. Therefore, this section restricts our sample to the eligible stocks that HKC can invest in and examines whether our original results are robust. We estimate the Fama-MacBeth regression in equation (2) for each investor group separately in a subsample with only eligible stocks in the HKC program.

Panel C in Table XII presents the next day's return prediction results. Even for this smaller group of stocks, the coefficients on *Oib* are again positive and significant at the 99% confidence level for all investor groups. The interquartile returns are 0.1125%, 0.0430%, 0.0853% and

0.1087%, close to those in Panel A Table III. Panel D presents longer-horizon cumulative return prediction results. In this subsample, we continue to find that foreign investors significantly predict stock returns over longer periods.

D. Index Constituents

Passive investors tend to trade indices directly rather than individual stocks, which potentially creates differences in the information environment for index constituent firms vs. non-constituent firms. We consider two influential indices, the local Chinese Stock Index 300, or the CSI300, which tracks the performance of the top 300 A-share stocks and is widely used by local investors, and the MSCI Emerging Market Index, which include large-cap A-shares that are eligible for HKC investors and is the benchmark for many international funds. We replace $|Overseas(i,d-1)|$ in equation (2) by constituent dummies, $CSI(i,d-1)$ and $MSCI(i,d-1)$, which equal 1 for stocks belonging to each index, and zero otherwise.

For the CSI 300 and MSCI results in Panel E and F of Table XII, we find index inclusions do not affect the predictive power of foreign investors' order flows. However, the predictive power of local institutions is interestingly and significantly higher for firms in the CSI 300 and MSCI, implying that local institutions are better informed about these index constituent firms. [21] [22]

VI. Conclusion

We investigate whether foreign investors are informed in the Chinese stock market using a comprehensive account-level dataset covering January 1, 2016, to June 30, 2019. We find that

[21] In the Internet Appendix Table IA.IX, we examine investors' return predictive power on AH dual-listed stocks and SOEs. we find no significant differences of foreign investors predictive power for firms with and without dual listings or for SOEs and Non-SOEs.

[22] The Chinese A-share stock market imposes 10% limits on daily stock prices (5% for special treatment stocks). In unreported results, we remove observations (representing 1.71% of the total sample) where stocks hit the daily price limits as this may cloud inference. The return predictive patterns still hold.

QFII, RQFII, and HKC investors can predict future stock price movements over both short and longer horizons. When relating their predictive power to firm-level information, we find that foreign investors can successfully predict firm-level earnings news. Their return predictive power is significantly stronger on the most value-relevant firm news days with large price movements, and the magnitude is comparable across foreign and local institutions. The evidence suggests that foreign investors are not at an informational disadvantage for firm-level information to local institutions, contrary to most previous studies. We also find foreign investors can predict market-level news, local and global, but the magnitude is smaller, and the significance is lower. Finally, during the market liberalization process, we find that expanding investment quotas helps to improve foreign investors' return predictability.

These findings have important implications for policymakers and researchers. Regulators should promote the development of price discovery and financial market efficiency by further examining how to take advantage of foreign investors' abilities. Identifying how foreign investors traded during the COVID-19 period, relative to their local counterparts, also presents a promising avenue for future research.

REFERENCES

Agarwal, Sumit, Sheri Faircloth, Chunlin Liu, and S. Ghon Rhee, 2009, Why do foreign investors underperform domestic investors in trading activities? Evidence from Indonesia. *Journal of Financial Markets* 12, 32-53.

Aggarwal, Reena, Isil Erel, Miguel A. Ferreira, and Pedro Matos, 2011, Does governance travel around the world? Evidence from institutional investors, *Journal of Financial Economics* 100, 154-181.

Bae, Kee-Hong, Arzu Ozoguz, Hongping Tan, and Tony S. Wirjanto, 2012, Do foreigners facilitate information transmission in emerging markets?, *Journal of Financial Economics* 105, 209-227.

Baik, Bok, Jun-Koo Kang, Jin-Mo Kim, and Joonho Lee, 2013, The liability of foreignness in international equity investments: Evidence from the US stock market, *Journal of International Business Studies* 44, 391-411.

Bailey, Warren, Connie X. Mao, and Kulpatra Sirodom, 2007, Investment restrictions and the cross-border flow of information: Some empirical evidence, *Journal of International Money and Finance* 26, 1-25.

Barber, Brad M, Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean, 2009, Just how much do individual investors lose by trading? *The Review of Financial Studies* 22, 609-632.

Barrot, Jean-Noel, Ron Kaniel, and David Sraer, 2016, Are retail traders compensated for providing liquidity? *Journal of Financial Economics* 120, 146-168.

Bekaert, Geert, and Campbell R. Harvey, 2000, Foreign speculators and emerging equity markets, *Journal of Finance* 55, 565-613.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad, 2005, Does financial liberalization spur growth? *Journal of Financial Economics* 77, 3-55.

Bernile, Gennaro, Jianfeng Hu, and Yuehua Tang, 2016, Can information be locked up? Informed trading ahead of macro-news announcements, *Journal of Financial Economics* 121, 496-520.

Bian, Jiangze, Kalok Chan, Bing Han, and Donghui Shi, 2020, Cross-Border Equity Flows and Information Transmission: Evidence from Chinese Stock Markets, Working Paper, University of International Business and Economics.

Boehmer, Ekkehart, Charles M. Jones, Juan Wu, and Xiaoyan Zhang, 2020, What do short sellers know? *Review of Finance* 24, 1203-1235.

Boehmer, Ekkehart, Charles M Jones, Xiaoyan Zhang, and Xinran Zhang, 2021, Tracking retail investor activity, *Journal of Finance* 76, 2249-2305.

Bradley, Daniel, Jonathan Clarke, Suzanne Lee, and Chayawat Ornthanalai, 2014, Are analysts' recommendations informative? Intraday evidence on the impact of time stamp delays, *Journal of Finance* 69, 645-673.

Brennan, Michael J., and H. Henry Cao, 1997, International portfolio investment flows, *Journal of Finance* 52, 1851-1880.

Campbell, John Y., Tarun Ramadorai, and Allie Schwartz, 2009, Caught on tape: Institutional trading, stock returns, and earnings announcements, *Journal of Financial Economics* 92, 66-91.

Chan, Kalok, Albert J. Menkveld, and Zhishu Yang, 2008, Information asymmetry and asset prices: Evidence from the China foreign share discount, *Journal of Finance* 63, 159-196.

Chen, Deqiu, Yujing Ma, Xiumin Martin, and Roni Michaely, 2022, On the fast track: Information acquisition costs and information production, *Journal of Financial Economics* 143, 794-823.

Chen, Keqi, Yuehan Wang, and Xiaoquan Zhu, 2019, The Value of Information in the China's

Connected Market, Working Paper, Tsinghua University.

Choe, Hyuk, Bong-Chan Kho, and René M. Stulz, 2005, Do domestic investors have an edge? The trading experience of foreign investors in Korea, *Review of Financial Studies* 18, 795-829.

Dong, Rui, Raymond Fisman, Yongxiang Wang, and Nianhang Xu, 2021, Air pollution, affect, and forecasting bias: Evidence from Chinese financial analysts, *Journal of Financial Economics* 139, 971-984.

Dvořák, Tomáš, 2005, Do domestic investors have an information advantage? Evidence from Indonesia, *Journal of Finance* 60, 817-839.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607-636.

Ferreira, Miguel A., Massimo Massa, and Pedro Matos, 2010, Shareholders at the gate? Institutional investors and cross-border mergers and acquisitions, *Review of Financial Studies* 23, 601-644.

Ferreira, Miguel A., and Pedro Matos, 2008, The colors of investors' money: The role of institutional investors around the world, *Journal of Financial Economics* 88, 499-533.

Ferreira, Miguel A., Pedro Matos, João Pedro Pereira, and Pedro Pires, 2017, Do locals know better? A comparison of the performance of local and foreign institutional investors, *Journal of Banking and Finance* 82, 151-164.

Frank, Murray Z., and Ali Sanati, 2018, How does the stock market absorb shocks?, *Journal of Financial Economics* 129, 136-153.

Froot, Kenneth A., Paul G. J. O'Connell, and Mark S. Seasholes, 2001, The portfolio flows of international investors, *Journal of Financial Economics* 59, 151-193.

Froot, Kenneth A., and Tarun Ramadorai, 2008, Institutional portfolio flows and international investments, *Review of Financial Studies* 21, 937-971.

Ge, Huimin, and Xiaoyan Zhang, 2022, News Tone and Stock Return in Chinese Market, Working Paper, Tsinghua University.

Grinblatt, Mark, and Matti Keloharju, 2000, The investment behavior and performance of various investor types: a study of Finland's unique data set, *Journal of Financial Economics* 55, 43-67.

Hendershott, Terrence, Dmitry Livdan, and Norman Schürhoff, 2015, Are institutions informed about news? *Journal of Financial Economics* 117, 249-287.

Huang, Alan Guoming, Hongping Tan, and Russ Wermers, 2020, Institutional trading around corporate news: Evidence from textual analysis, *Review of Financial Studies* 33, 4627-4675.

Huang, Wei, and Tao Zhu, 2015, Foreign institutional investors and corporate governance in emerging markets: Evidence of a split-share structure reform in China, *Journal of Corporate Finance* 32, 312-326.

Irvine, Paul, Marc Lipson, and Andy Puckett, 2007, Tipping, *Review of Financial Studies* 20, 741-768.

Ivković, Zoran, and Narasimhan Jegadeesh, 2004, The timing and value of forecast and recommendation revisions, *Journal of Financial Economics* 73, 433-463.

Jia, Chunxin, Yaping Wang, and Wei Xiong, 2017, Market segmentation and differential reactions of local and foreign investors to analyst recommendations, *Review of Financial Studies* 30, 2972-3008.

Jiang, George J, and Kevin X Zhu, 2017, Information shocks and short-term market underreaction, *Journal of Financial Economics* 124, 43-64.

Jones, Charles M, Donghui Shi, Xiaoyan Zhang, and Xinran Zhang, 2022, Understanding Retail

Investors: Evidence from China, Working Paper, Columbia University.

Kacperczyk, Marcin, Savitar Sundaresan, and Tianyu Wang, 2021, Do foreign institutional investors improve price efficiency? *Review of Financial Studies* 34, 1317-1367.

Kang, Jun-Koo, and René M. Stulz, 1997, Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan, *Journal of Financial Economics* 46, 3-28.

Kelley, Eric K, and Paul C Tetlock, 2013, How wise are crowds? Insights from retail orders and stock returns, *Journal of Finance* 68, 1229-1265.

Li, Zengquan, T. J. Wong, and Gwen Yu, 2020, Information dissemination through embedded financial analysts: Evidence from China, *The Accounting Review* 95, 257-281.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan, 2019, Size and value in China, *Journal of Financial Economics* 134, 48-69.

Ma, Chang, John H. Rogers, and Sili Zhou, 2021, The effect of the China Connect, Working Paper, Fudan University.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

Savor, Pavel, and Mungo Wilson, 2016, Earnings announcements and systematic risk, *Journal of Finance* 71, 83-138.

Seasholes, Mark. 2000, Smart foreign traders in emerging markets, Working Paper, Harvard Business School.

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437-1467.

Yoon, Aaron S., 2021, The role of private disclosures in markets with weak institutions: evidence from market liberalization in China, *The Accounting Review* 96, 433-455.

Table I
Comparison of QFII, RQFII and HKC

This table summarizes the differences in QFII, RQFII and HKC investors in China.

	QFII	RQFII	HKC
Investor	1.Institutional investors such as security companies, commercial bank, asset management company and others 2. Requirements on the scale of assets under management and operational periods.	1.In 2011, only Hong Kong subsidiaries of Chinese financial institution 2.gradually extended to other locations.	Hong Kong and overseas investors, including both retail and institutional investors.
Capital Control	1. 3-month lockdown period for non open-end funds. 2. The monthly remittance of capital and profits could not exceed 20% of the total asset at the end of previous year 3. Restrictions were removed on June 10, 2018	3-month lockdown period for non open-end funds, which was removed on June 11, 2018.	Not required
Investment Quota	1. Basic quota for a single QFII was limited by the scale of assets under management and was no more than $5 billion 2. Aggregated QFII quota was raised to $300 billion on January 14, 2019 3. Restriction cancelled on September 10, 2019.	1. Basic quota for a single RQFII was limited by the scale of assets under management 2. Aggregated RQFII quota varies for different locations. For example, the aggregated quota for Hong Kong was RMB 500 billion on July 4, 2017 3. Restriction cancelled on September 10, 2019.	1. Total investment quota was set at RMB 300 billion. Restriction cancelled on Aug 17, 2016. 2. Initial northbound daily quota was RMB 13 billion, and rose to 52 billion after May 1, 2018.
Funding	1.Remit foreign currency as the principal 2.Both FX and RMB are allowed after May 7, 2020	1.Offshore Chinese Yuan as the principal 2.Both FX and RMB are allowed after May 7, 2020	Not required
Investable Stock	1. All A-share stocks listed on exchanges 2. Fixed income and other financial products	1. All A-share stocks listed on exchanges 2. Fixed income and other financial products	1.Constituent stocks of the SSE 180 Index and SSE 380 Index. 2.A shares with H shares listed in HK
Ownership	1. A single QFII licensee or RQFII licensee or HKC cannot hold more than 10% of a given company's A-shares. 2. Total A shares held by all QFII, RQFII and HKC investors for any given company cannot exceed 30% of its total outstanding shares.		

Table II
Summary statistics of foreign investors and local institutions

This table summarizes trading and holdings by foreign investors and local institutions. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. Foreign investors refer to as Qualified Foreign Institutional Investors (QFII), Renminbi Qualified Foreign Institutional Investor (RQFII) and investors via Shanghai-Hong Kong Connected Scheme (HKC). We refer to local mutual funds, hedge funds, insurance companies, security companies, and other institutional investors as local institutions (Local INST). In Panel A, we report the daily average number of stocks held and traded by investors, the daily average of investors' aggregated trading volume (the mean of buy and sell) in billion RMB, and the daily average of investors' aggregated holdings in billion RMB. At the stock-day level, the investors' order imbalance measure (*Oib*) is defined as buy volume (in shares) minus sell volume (in shares) divided by the sum of buy and sell, as shown in equation (1). Panel B reports the time-series average of the cross-sectional mean and standard deviation. *AR(1)* is the cross-sectional mean of the first-order autocorrelation of the order imbalance measure. We also report the time-series average of the cross-correlations of the order imbalance measure across QFII, RQFII and HKC.

Panel A. Time-series average of investors' aggregate trading and holdings

	QFII	RQFII	HKC	Local INST
Daily trading volume (Bil. RMB)	1.51	0.16	4.33	28.95
Trading volume of total market (%)	0.79%	0.08%	2.24%	14.80%
Number of stocks traded	946	174	561	1,227
Daily Holding (Bil. RMB)	240.23	58.01	311.14	3590.2
Holding shares of total market (%)	0.95%	0.23%	1.20%	14.19%
Number of stocks held	1,261	901	744	1,297

Panel B. Time-series average of cross-sectional statistics of the order imbalance measure

				Correlations		
	Mean	Std	AR(1)	Oib(QFII)	Oib(RQFII)	Oib(HKC)
Oib(QFII)	-0.01	0.86	0.09			
Oib(RQFII)	0.02	0.82	0.44	0.09		
Oib(HKC)	0.02	0.58	0.12	0.14	0.04	
Oib(Local INST)	-0.01	0.47	0.18	0.09	0.06	0.06

Table III
Stock return prediction of foreign investors and local institutions

This table presents estimation results on whether foreign investors and local institutions can predict the cross-sectional stock returns in both short-term and long-term horizons. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate daily Fama-MacBeth (1973) regressions. Panel A presents results on the next day's return prediction, as in equation (2). Panel B presents the coefficients on the order imbalance measure in the *w* weeks cumulative return prediction, as in equation (3). The key independent variable is the order imbalance measure on the previous day (*Oib(d-1)*). *Ret(d-1)* is the previous day's stock return. *Ret(d-6, d-2)* is the cumulative daily return over the [-6, -2] window. *Ret(d-27, d-7)* is the cumulative daily return over the [-27, -7] window. We include control variables: the log of market capitalization *(Lnsize)*, earnings-to-price ratio *(EP)* and monthly turnover *(Turnover)*, all measured at the end of previous month. *Adj-R^2* is the time-series average of adjusted R-squared in the cross-sectional regression. *Interquartile* is the time-series average of the cross-sectional interquartile range of the order imbalance variable. *Interquartile Return* represents the magnitude of investors' return predictability, defined as *Interquartile* multiplied by the estimated coefficient on the order imbalance. To account for potential serial correlation in the coefficients, the standard errors are adjusted using Newey-West (1987) with five lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. One-day return prediction

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Oib(d-1)	0.0649***	0.0247***	0.0783***	0.1330***
	(17.02)	(3.10)	(10.44)	(18.57)
Ret(d-1)	0.7388	-0.3870	0.2152	2.2033***
	(1.54)	(-0.61)	(0.37)	(4.32)
Ret(d-6, d-2)	-0.8924***	-0.5660**	-0.7376***	-1.1077***
	(-4.48)	(-2.06)	(-3.20)	(-5.88)
Ret(d-27, d-7)	-0.2353***	0.0237	-0.2095**	-0.3077***
	(-2.75)	(0.17)	(-2.03)	(-4.68)
Lnsize	-0.0078	0.0034	0.0045	-0.0016
	(-0.78)	(0.32)	(0.44)	(-0.16)
EP	1.3757***	1.2805	1.5416***	1.4607***
	(2.91)	(1.62)	(2.82)	(3.22)
Turnover	-0.0521***	-0.1848***	-0.1121***	-0.0556***
	(-2.66)	(-3.49)	(-4.35)	(-3.25)
Adj-R^2	8.96%	14.75%	10.07%	8.83%
Interquartile	1.8295	1.2342	0.9666	0.7012
Interquartile Return	0.1188%***	0.0305%***	0.0757%***	0.0933%***
	QFII-Local	RQFII-Local	HKC-Local	
Interquartile Return Difference	0.0255%***	-0.0626%***	-0.0184%***	
	(3.29)	(-5.52)	(-2.64)	

Panel B. 12-week cumulative return prediction

Dep: Cumulative Ret(w)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Week number w	QFII		RQFII		HKC		Local INST
1	0.1123***		0.0686***		0.0985***		0.2717***
2	0.1289***		0.1102***		0.1184***		0.3631***
3	0.1524***		0.1380***		0.1271***		0.4159***
4	0.1688***		0.1494***		0.1338***		0.4588***
5	0.1779***		0.2089***		0.1348***		0.4822***
6	0.1834***		0.2065***		0.1594***		0.5250***
7	0.2068***		0.2157***		0.1858***		0.5669***
8	0.2172***		0.2434***		0.1874***		0.6038***
9	0.2119***		0.2356***		0.1598***		0.6010***
10	0.2284***		0.2665***		0.1701***		0.6242***
11	0.2387***		0.3205***		0.1725**		0.6375***
12	0.2507***		0.3240***		0.1677**		0.6510***
Interquartile Cumulative Return	QFII	QFII-Local	RQFII	RQFII-Local	HKC	HKC-Local	Local INST
1	0.2054%***	0.0149%	0.0847%***	-0.1060%***	0.0952%***	-0.0946%***	0.1905%***
2	0.2358%***	-0.0188%	0.1361%***	-0.1191%***	0.1144%***	-0.1436%***	0.2546%***
3	0.2789%***	-0.0128%	0.1704%***	-0.1221%**	0.1228%***	-0.1725%***	0.2916%***
4	0.3088%***	-0.0129%	0.1844%***	-0.1381%**	0.1293%***	-0.1917%***	0.3217%***
5	0.3255%***	-0.0127%	0.2578%***	-0.0815%	0.1303%***	-0.2091%***	0.3381%***
6	0.3356%***	-0.0325%	0.2549%***	-0.1146%	0.1541%***	-0.2141%***	0.3681%***
7	0.3784%***	-0.0191%	0.2662%***	-0.1331%	0.1796%***	-0.2185%***	0.3975%***
8	0.3973%***	-0.0261%	0.3004%***	-0.1243%	0.1811%***	-0.2431%***	0.4234%***
9	0.3877%***	-0.0337%	0.2908%***	-0.1325%	0.1544%***	-0.2701%***	0.4214%***
10	0.4178%***	-0.0199%	0.3289%***	-0.1102%	0.1644%***	-0.2734%***	0.4377%***
11	0.4367%***	-0.0103%	0.3956%***	-0.0527%	0.1667%**	-0.2797%***	0.4470%***
12	0.4586%***	0.0021%	0.3999%***	-0.0576%	0.1621%**	-0.2941%***	0.4565%***

Table IV

Foreign investors' predictive power in different counterparties

This table presents results about foreign investors' return predictive power with different counterparties. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We separate counterparties into three groups: foreign investors, local institutions and retail investors (RT). According to the sign of order imbalances, we separate investors' daily trade directions into buy (B) and sell (S). With the three groups of investors, and two sides of trades, all observations are divided into six bins: BBS, SSB, BSS, SBB, BSB and SBS. The first letter indicates the trade direction of foreign investors, the second indicates trade direction of local institutions, and the third for retail investors. Panel A reports the proportion of trades in the six counterparty groups. We further examine whether foreign investors' return predictive pattern changes with different counterparties with Fama-MacBeth regressions as in equation (3). The indicator variable *I(k,d,G)* is equal to one if trades from foreign investor group *G* for stock *i* on day *d-1* fall in the *k*-th counterparty bin, otherwise it is zero. As estimating regressions for RQFII, we find in 24 cross-sectional regressions, the six interaction variables and the constant are highly colinear, leading to extreme estimations. Therefore, we drop the 24 estimated cross-sectional coefficients when calculating the time-series average. Panel B reports the estimated coefficients and related interquartile returns. The dependent variables are expressed in percentage. All control variables are same as those in equation (2). To account for serial correlation in the coefficients, the standard errors of the time-series are adjusted using Newey-West (1987) with 5 lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. The proportion of trades in different groups

Trade direction	(1) QFII-Local INST-RT	(2) RQFII-Local INST-RT	(3) HKC-Local INST-RT
BBS	27%	24%	26%
SSB	26%	25%	25%
BSS	4%	5%	5%
SBB	4%	4%	4%
BSB	19%	22%	21%
SBS	19%	20%	18%
Observations	755, 991	134,851	430,067

Panel B. Predictive patterns with different counterparties

Dep: Ret(d)	(1)	(2)	(3)
	QFII	RQFII	HKC
Oib(d-1)*BBS	0.0500***	0.0695	0.1890***
	(4.06)	(1.33)	(9.92)
Oib(d-1)*SSB	0.2009***	0.1194**	0.1188***
	(15.63)	(2.34)	(6.80)
Oib(d-1)*BSS	-0.0097	-0.0550	0.0808***
	(-0.56)	(-0.83)	(3.44)
Oib(d-1)*SBB	0.1388***	0.1132*	0.0612**
	(7.72)	(1.83)	(2.23)
Oib(d-1)*BSB	-0.0749***	-0.0405	0.0009
	(-5.78)	(-0.77)	(0.04)
Oib(d-1)*SBS	0.0432***	-0.0743	-0.0729***
	(3.27)	(-1.44)	(-3.91)
Interquartile return			
BBS	0.0101%***	0.0049%	0.0940%***
SSB	0.0333%***	0.0075%**	0.0633%***
BSS	-0.0017%	-0.0048%	0.0342%***
SBB	0.0226%***	0.0089%*	0.0275%**
BSB	-0.0185%***	-0.0032%	0.0004%
SBS	0.0108%***	-0.0059%	-0.0379%***

Table V
Account performance of foreign investors

This table presents the account performance of foreign investors as well as local institutions. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. Given the detailed data on holding and trading, we design a methodology to obtain an estimate of aggregate performances for foreign investors and local institutions. On each trading day, we compute the total gain for each investor group as in equation (4). The total account performance equals the capital gain from holdings plus trading proceeds minus transaction costs. Transaction costs include commission cost (0.05%) imposed on both the buy and sell side, with a minimum of 5 CNY for each trade), the stamp tax (0.10% of the sales amount), and the transfer fee (0.002% imposed on both sides). We add up all of the daily gain from the entire sample and divide the total number by 3.5 to get the annual performance. To have a sense of return percentage of investment, we calculate the account performance in percentage by dividing the aggregate cash flows by the daily average of investors' aggregate holdings as presented in Table II Panel A.

Investor	Total in billion CNY	Cost in billion CNY	Total (%)	Cost (%)
QFII	42.82	49.88	17.83%	-0.31%
RQFII	12.53	13.57	21.61%	-0.14%
HKC	63.67	64.85	20.46%	-0.67%
Local INST	411.12	504.01	11.45%	-0.39%

Table VI

Stock return predictive power of overlapping and specific order imbalances

This table reports results on the predictive power of overlapping and specific order imbalances by foreign investors and local institutions. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate Fama-MacBeth regressions as in equation (2), but replace foreign investors' order imbalance measure by *Oib(overlap with local)* and *Oib(foreign specific)*, which are calculated in equation (5) and (6). We use a similar procedure to decompose local institutions' order imbalances and obtain *Oib(overlap with foreign)* and *Oib(local specific)*. *Interquartile* is the time-series average of the cross-sectional interquartile ranges of investors' overlapping and specific trading. *Interquartile return* is defined as *Interquartile* multiplied by the estimated coefficient on the related order imbalance. *Adj-R^2* is the time-series average of adjusted R-squareds in the cross-sectional regression. Control variables are the same as those in equation (2). To spare the space, we omit the coefficients on control variables. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Oib(d-1, overlap with local)	0.3553	4.0874	1.9399	
	(0.18)	(1.14)	(0.84)	
Oib(d-1, foreign specific)	0.0593***	0.0197**	0.0700***	
	(15.67)	(2.45)	(10.02)	
Oib(d-1, overlap with foreign)				0.7173***
				(4.50)
Oib(d-1, local specific)				0.2355***
				(11.03)
Adj-R^2	9.18%	15.07%	10.32%	16.39%
Interquartile Return				
Oib(d-1, overlap with local)	0.0399%	0.4916%	0.1330%	
Oib(d-1, foreign specific)	0.1034%	0.0240%	0.0670%	
Oib(d-1, overlap with foreign)				0.0769%
Oib(d-1, local specific)				0.1205%

Table VII

Stock return predictive power, earnings, and analyst-related events

This table presents stock return prediction results related to firm earnings announcements and analyst-related events. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. Our sample covers 15,477 earnings announcements and 41,722 analyst-related events, totaling 50,331 event days. We estimate quarterly Fama-MacBeth regressions. Panel A presents results for equation (7) on investors' predictive power with cumulative abnormal returns around event days. The dependent variable *CAR* (*AR*) is the cumulative stock return minus the cumulative market return over the event period *[d, d+k]* (on day d). Panel B presents results for equation (8) on whether investors have stronger return predictive power on firm-level news days. The indicator variable *Bignews(i, d)* is equal to one if stock *i*'s return on event day *d* is outside the 5th and 95th percentiles of all event day returns, and otherwise it is zero. *NBignews(i, d)* is equal to one if stock *i*'s return on event day *d* is inside the 5th and 95th percentiles of event day returns and otherwise it is zero. All dependent variables are expressed as percentages. Control variables are the same as those in equation (2). The standard errors are adjusted using Newey-West (1987) with five lags. We omit coefficients on control variables and *t*-statistics. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Cumulative abnormal return prediction around firm event days

	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Coefficients on Oib(-1)				
AR[0]	0.0878***	0.0580	0.1018*	0.3610***
CAR[0,1]	0.1464***	0.0914***	0.2064**	0.5515***
CAR[0,61]	0.2712**	0.2933	0.3680***	1.9058***
CAR[0, 251]	0.6318***	0.6809***	1.5919*	3.3235***
Interquartile return				
AR[0]	0.1713%	0.1161%	0.0891%	0.2270%
CAR[0,1]	0.2856%	0.1829%	0.1806%	0.3468%
CAR[0,61]	0.5292%	0.5867%	0.3221%	1.1985%
CAR[0, 251]	1.2327%	1.3617%	1.3931%	2.0901%

Panel B. Stock return prediction with firm-level event dummy variables

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
$\widehat{e1}$: Oib(d-1)	0.0977***	0.0433***	0.0954***	0.2120***
$\widehat{e2}$: Oib(d-1)×Bignews(d)	0.5177***	0.6788***	0.3035	2.4531***
$\widehat{e3}$: Oib(d-1)×NBignews(d)	-0.0342***	0.0292	0.0824**	-0.0042
Interquartile (Oib)× $\widehat{e1}$: $\widehat{Ret}1$(Non-event)	0.1787%	0.0535%	0.0922%	0.1487%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e2})$: $\widehat{Ret}2$(Bignews)	1.1259%	0.8913%	0.3856%	1.8688%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e3})$: $\widehat{Ret}3$(NBignews)	0.1161%	0.0896%	0.1718%	0.1457%
Contribution of Bignews days (0.49%)	3.06%	7.38%	1.94%	5.83%
Contribution of NBignews days (4.45%)	2.86%	6.73%	7.87%	4.13%

Table VIII
Stock return predictive power and media news

This table presents stock return prediction results related to local firm-level financial media news. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We obtain news information from CNRDS. Our sample includes 353,551 news days accounting for 34.69% of all observations. We estimate similar regressions as in equation (7) and (8) by focusing on media news days. Panel A presents results on cumulative abnormal return predictions. The dependent variable *CAR* (*AR*) is the cumulative stock return minus the cumulative market return over the event period *[d, d+k]* (on day d). Panel B presents results on whether investors have stronger return predictive power on media news days. *Bignews(i, d)* is equal to one if stock *i*'s return on financial media news day *d* is outside the 5th and 95th percentiles of all news day returns and otherwise it is zero. *NBignews(i, d)* is equal to one if stock *i*'s return on news day *d* is inside the 5th and 95th percentiles of all news day returns and otherwise it is zero. All dependent variables are expressed in a percentage. Control variables are same as those in equation (2). The standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. To spare the space, we omit coefficients on control variables and *t*-statistics in the table. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Cumulative abnormal return prediction around media news days

	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Coefficients on Oib(-1)				
AR[0]	0.0798***	0.0140***	0.0775***	0.1839***
CAR[0,1]	0.1300***	0.0390***	0.1307***	0.3061***
CAR[0,61]	0.3653***	0.2363**	0.2746***	1.1970***
CAR[0, 251]	0.5475***	0.6655***	0.9926*	1.9756***
Interquartile return				
AR[0]	0.1570%	0.0279%	0.0732%	0.1239%
CAR[0,1]	0.2559%	0.0779%	0.1235%	0.2063%
CAR[0,61]	0.7189%	0.4716%	0.2593%	0.8069%
CAR[0, 251]	1.0774%	1.3281%	0.9374%	1.3317%

Panel B. Stock return prediction with media news dummy variables

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
$\widehat{e1}$: Oib(d-1)	0.0906***	0.0323*	0.0910***	0.1847***
$\widehat{e2}$: Oib(d-1)×Bignews (d)	0.3550***	0.2713	0.1473	1.4984***
$\widehat{e3}$: Oib(d-1)×NBignews (d)	-0.0085	0.0162	0.0156	-0.0553*
Interquartile (Oib)× $\widehat{e1}$: $\widehat{Ret}1$(Non-news)	0.1657%	0.0398%	0.0880%	0.1295%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e2})$: $\widehat{Ret}2$(Bignews)	0.8153%	0.3747%	0.2304%	1.1802%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e3})$: $\widehat{Ret}3$(NBignews)	0.1502%	0.0599%	0.1031%	0.0908%
Contribution of Bignews days (3.47%)	15.42%	22.53%	8.18%	26.61%
Contribution of NBignews days (31.23%)	25.57%	32.39%	32.97%	18.42%

Table IX

Stock return predictive power and firms with cross-border business

This table presents estimation results on whether foreign investors better predict stock returns on firms with cross-border business. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We investigate foreign investors' return predictive power on stocks with cross-border business activities with daily Fama-MacBeth regressions, where we modify the benchmark regression in equation (2) by interacting the order flow variables with $|Overseas(i, d-1)|$. *|Overseas(i, d-1)|* is the absolute value of the ratio of overseas revenue to total revenue, and it is equal to zero if there is no overseas revenue. We report interquartile returns at the bottom. All dependent variables are expressed as percentages. Control variables are the same as those in equation (2). To spare the space, we omit coefficients on control variables. The standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
$\widehat{a1}$: Oib(d-1)	0.0625***	0.0225**	0.0741***	0.1292***
	(15.48)	(2.55)	(9.44)	(17.63)
$\widehat{a2}$: Oib(d-1)×|Overseas(d-1)|	0.0427**	0.0524	0.0720**	0.0601**
	(2.33)	(0.48)	(2.04)	(2.01)
Interquartile Returns				
Interquartile Oib(d-1)× $\widehat{a1}$	0.1144%	0.0278%	0.0716%	0.0906%
Interquartile Oib(d-1)× $(\widehat{a1} + \widehat{a2})$	0.1925%	0.0925%	0.1412%	0.1327%

Table X

Investors' predictive power and market-level news

This table presents results on investors' predictive power with respect to market-level news. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We use a value-weighted portfolio of A-share stocks and the MSCI World Index as proxies for the Chinese and global stock markets, respectively. Panel A presents estimation results of equation (9) and (9'). In equation (9), the dependent variables *MktRet* are the Chinese stock market returns on day *d* and over the period *[d, d+4]*. The independent variable *AggOib(d-1,G)* is the aggregate buy volume in RMB across all stocks, minus the aggregate sell volume, divided by the sum of aggregate buys and sells for investor group *G* on day *d-1*. In equation (10'), we investigate whether investors have stronger market return predictive power on market-level news days. The dependent variable is the Chinese stock market return on day *d*. We interact *AggOib(d-1,G)* with the indicator variable *BignewsMkt(d)*, which is equal to one if, on day *d,* the market return is outside the 5^{th} and 95^{th} percentiles of all market returns, and zero otherwise. We consider both Chinese and global stock market news days and they account for 9.89% of all sample days, separately. Panel B presents results about the cross-sectional return predictive power with respect to market-level news for equation (10). The independent variable $\widehat{a1}(d)$ is the time-series coefficients on *Oib(d-1)* for equation (2). Market returns are expressed as percentages. The contribution of market-level news days is calculated following the methodology in Table V. The standard errors are adjusted using Newey-West (1987) with five lags. To spare the space, we omit *t*-statistics; ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Aggregate order flows predict market returns over time

	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Equation (9): Market return prediction				
MktRet[0]	0.6602***	0.1884	0.7181**	0.7735*
MktRet[0,4]	1.4568**	0.5493	0.8818	0.0285
Interquartile MktRet[0]	0.1736%	0.1010%	0.1628%	0.0940%
Interquartile MktRet[0,4]	0.3830%	0.2943%	0.1999%	0.0035%
Equation (9'): Chinese stock market news				
$\widehat{f1}$: AggOib(d-1)	0.1306	0.0742	0.1308	0.5017*
$\widehat{f2}$: AggOib(d-1)×BignewsMkt(d)	3.9059***	1.1607	5.3511**	2.7660
Interquartile(AggOib)× $\widehat{f1}$	0.0343%	0.0398%	0.0297%	0.0609%
Interquartile(AggOib)×($\widehat{f1}$+$\widehat{f2}$)	1.0613%	0.6617%	1.2427%	0.3969%
Contribution of BignewsMkt days (9.89%)	77.23%	64.61%	82.14%	41.69%
Equation (9'): Global stock market news				
$\widehat{f1}$: AggOib(d-1)	0.3856*	0.2209*	0.4533*	0.6514
$\widehat{f2}$: AggOib(d-1)×BignewsMkt(d)	1.6800**	-0.2380	2.0441	2.3672
Interquartile(AggOib)× $\widehat{f1}$	0.1014%	0.1184%	0.1028%	0.0791%
Interquartile(AggOib)×($\widehat{f1}$+$\widehat{f2}$)	0.5431%	-0.0091%	0.5662%	0.3667%
Contribution of BignewsMkt days (9.89%)	37.02%	-0.85%	37.68%	33.71%

Panel B. Order flows predict cross-sectional stock returns with respect to market-level news

Dep: $\widehat{a1}(d)$	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Chinese stock market news				
$\widehat{h0}$: Intercept	0.0673***	0.0276***	0.0754***	0.1378***
$\widehat{h1}$: BignewsMkt(d)	-0.0241	0.0173	0.0285	-0.0487**
Interquartile (Oib)$\times\ \widehat{h0}$	0.1231%	0.0341%	0.0729%	0.0966%
Interquartile (Oib)$\times\ (\widehat{h0}+\widehat{h1})$	0.0790%	0.0554%	0.1004%	0.0625%
Contribution of BignewsMkt days (9.89%)	6.58%	15.15%	13.13%	6.63%
Global stock market news days				
$\widehat{h0}$: Intercept	0.0627***	0.0247***	0.0750***	0.1300***
$\widehat{h1}$: BignewsMkt(d)	0.0229*	0.0450	0.0376	0.0316
Interquartile (Oib)$\times\ \widehat{h0}$	0.1147%	0.0304%	0.0725%	0.0911%
Interquartile (Oib)$\times\ (\widehat{h0}+\widehat{h1})$	0.1566%	0.0860%	0.1088%	0.1133%
Contribution of BignewsMkt days (9.89%)	13.03%	23.66%	14.15%	12.00%

Table XI

Stock return predictive power and market liberalization

This table presents estimation results on how foreign investors' return predictive power changes after the relaxation on regulations. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We apply a two-step regression procedure, as in equation (11). In the first step, we perform an OLS regression on each day and obtain time-series coefficients on *Oib(d-1)*, $\widehat{a1}(d)$. In the second step, we regress the estimated coefficient on a series of indicator variables related to regulations. There are three major policy changes for QFII. First, on February 3, 2016, the State Administration of Foreign Exchange (SAFE) announced an increase in the maximum basic investment quota for a single QFII from $1 billion to $5 billion. Second, on June 10, 2018, SAFE announced the removal of the 3-month lock-up period and the maximum 20% capital repatriation limitation for QFII. Third, on January 14, 2019, SAFE announced an increase in QFII's total investment quota from $150 billion to $300 billion. There are two major policy changes for the RQFII program. RQFII originally were not allowed to invest in stocks or stock investment funds at levels that exceeded 20% of its raised capital. CSRC verbally announced the lifting of that restriction at a press conference on September 30, 2016. Then, on June 11, 2018, SAFE announced the removal of the 3-month lock-up period for RQFII. Finally, there are two regulatory changes for the HKC program. First, on August 16, 2016, the RMB 300 billion aggregated quota was removed. Second, on May 1, 2018, the daily quota increased from RMB 13 billion to RMB 52 billion. Based on these regulations, we define seven regulation dummy variables, *Quota2016QFII, FX2018QFII, Quota2019QFII, Invest2016RQFII, FX2018RQFII, Quota2016HKC* and *Quota2018HKC*. Each dummy variable is equal to zero before the related event occurs and one afterward. Panels A, B and C show the results on the second step regressions for QFII, RQFII and HKC, respectively. *Adj-*R^2 is the adjusted R-squared in the second-step regression. To account for serial correlation in the coefficients, the standard errors are adjusted using Newey-West (1987) with five lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Regulation changes on QFII

Dep: $\widehat{a1}(d)$	(1) QFII
Intercept	0.0095
	(0.61)
Quota2016QFII(d-1)	0.0585***
	(3.66)
FX2018QFII(d-1)	0.0028
	(0.27)
Quota2019QFII(d-1)	-0.0187
	(-1.18)
Adj-R^2	0.59%

Panel B. Regulation changes on RQFII

Dep: $\widehat{a1}(d)$	(1) RQFII
Intercept	0.0260*
	(1.81)
Invest2016RQFII(d-1)	-0.0099
	(-0.54)
FX2018RQFII(d-1)	0.0215
	(1.09)
Adj-R^2	-0.09%

Panel C. Regulations changes on HKC

Dep: $\widehat{a1}(d)$	(1) HKC
Intercept	0.0289***
	(3.43)
Quota2016HKC(d-1)	0.0234**
	(2.05)
Quota2018HKC(d-1)	0.0913***
	(5.25)
Adj-R^2	6.58%

Table XII
Further Discussions and Robustness checks

This table presents results for several robustness checks. Our sample period is January 1, 2016, to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. In Panel A, we decompose the order flow measures into two parts, a 20-day moving average measure, *MAOib*, to capture the persistent component of the order imbalance variable, and the difference between the *Oib* and the *MAOib*, *ShockOib*, to capture the day-to-day fluctuations in order flows. We investigate the return predictive power of *MAOib* and *ShockOib* with daily Fama-MacBeth regressions as in equation (2). Panel B presents the results on the next-day's stock return prediction for investors' buy and sell. For stock *i* on day *d* from investor group *G*, *Buy(i,d,G)* and *Sell(i,d,G)* are defined as the buy volume in shares and sell volume divided by outstanding A-shares, respectively. If there is no buy or sell, two variables are equal to zero. Panel C and D present results on the next day's return prediction and *w* weeks cumulative return prediction in a subsample with only eligible stocks in HKC program, as specified in equations (2), respectively. In Panel F and G, we investigate foreign investors' return predictive power on index constituents' stocks as shown in equation (2). The indicator variable *CSI (i, d-1)* is equal to 1 if stock *i* on day *d-1* is included in the CSI 300 Index, and zero otherwise. *MSCI (i, d-1)* is equal to 1 if stock on day *d-1* is announced to be included in the MSCI Emerging Market Index and zero otherwise. *Adj-R^2* is the time-series average of adjusted R-squared in the cross-sectional regression. All dependent variables are in percentage except for the test in Panel E. To spare the space, we omit coefficients on control variables and *t*-statistics. The standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Moving average vs. shock component of order flows

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
MAOib(d-1)	0.1170***	0.0375**	0.0869***	0.3079***
ShockOib(d-1)	0.0601***	0.0215**	0.0797***	0.1136***
Adj-R^2	9.01%	14.98%	10.13%	8.93%
Interquartile Return of MAOib	0.0382%***	0.0225%**	0.0179%***	0.0603%***
Interquartile Return of ShockOib	0.0924%***	0.0222%**	0.0669%***	0.0732%***

Panel B. Stock return predictive power of buy and sell

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Buy(d-1)	1.7667***	1.3619	1.2712***	0.2919***
Sell(d-1)	-0.2000	-0.2824	-0.5577***	-0.2111***
Interquartile Ret (Buy)	0.0112%	0.0000%	0.0295%	0.0525%
Interquartile Ret (Sell)	-0.0012%	0.0000%	-0.0121%	-0.0388%

Panel C. One-day return prediction with eligible stocks in HKC program

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Oib(d-1)	0.0627***	0.0367***	0.0904***	0.1754***
Interquartile	1.7954	1.1716	0.9437	0.6195
Interquartile Return	0.1125%***	0.0430%***	0.0853%***	0.1087%***

Panel D. 12-week cumulative return prediction with eligible stocks in HKC program

Interquartile Cumulative Return	(1)	(2)	(3)	(4)
Week number *w*	QFII	RQFII	HKC	Local INST
1	0.1620%***	0.0866%***	0.1079%***	0.1923%***
2	0.1741%***	0.1158%***	0.1393%***	0.2670%***
3	0.2179%***	0.1511%***	0.1523%***	0.3125%***
4	0.2496%***	0.1647%***	0.1559%***	0.3763%***
5	0.2540%***	0.1891%***	0.1593%***	0.3967%***
6	0.2786%***	0.1646%**	0.1906%***	0.4418%***
7	0.3170%***	0.1750%**	0.2276%***	0.4775%***
8	0.3409%***	0.1947%**	0.2343%***	0.5056%***
9	0.3274%***	0.1829%**	0.2166%***	0.5037%***
10	0.3667%***	0.2376%***	0.2395%***	0.5354%***
11	0.3623%***	0.3140%***	0.2447%***	0.5419%***
12	0.3713%***	0.3490%***	0.2511%***	0.5528%***

Panel E. Stocks as constituents of the CSI300 Index

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Oib(d-1)	0.0677***	0.0364**	0.0781***	0.1267***
Oib(d-1)×CSI(d-1)	-0.0064	-0.0147	0.0012	0.0644***
Adj-R^2	9.10%	15.33%	10.23%	8.96%

Panel F. Stocks as constituents of the MSCI Emerging Market Index

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Oib(d-1)	0.0637***	0.0404	0.1414***	0.1853 ***
Oib(d-1)×MSCI(d-1)	0.0001	0.0015	0.0181	0.0975***
Adj-R^2	7.63%	15.14%	9.39%	7.67%

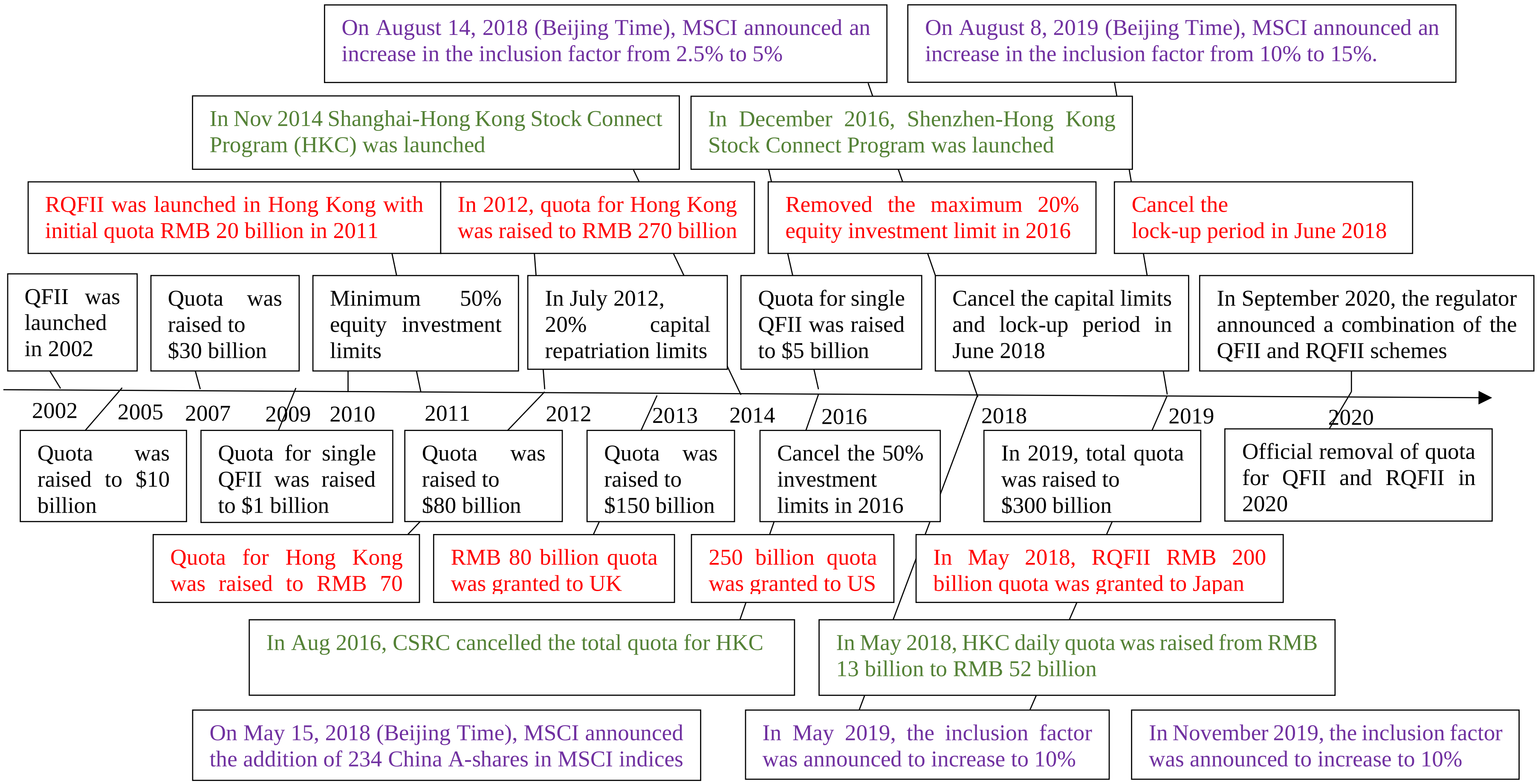
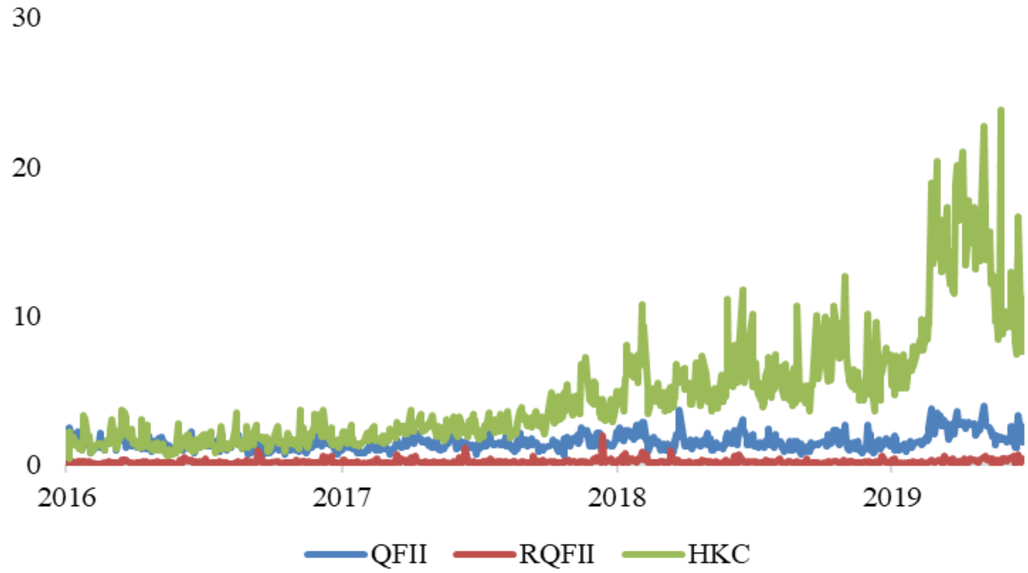


Figure 1. The timeline of QFII, RQFII and HKC in China. This figure presents the key events during the development of QFII, RQFII and HKC in the Chinese stock market. QFII is with black font, RQFII is with green font, HKC is with green font. We also introduce events that MSCI announced to include A-shares into its indices in purple font.

Panel A. Aggregate trading volume in billion RMB



Panel B. Aggregate holding in billion RMB

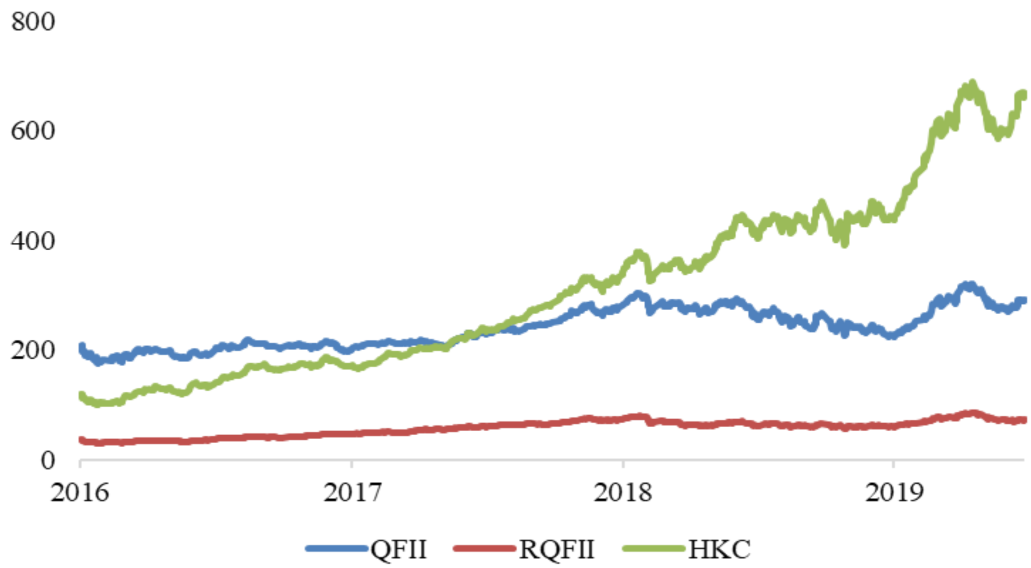


Figure 2. Aggregate trading and holding for QFII, RQFII and HKC. The figure shows the time-series aggregate trading volume and holdings by QFII, RQFII and HKC from January 1, 2016 to June 30, 2019. Panel A shows the time-series aggregate trading volume in billion RMB. Panel B shows the time-series aggregate holdings in billion RMB

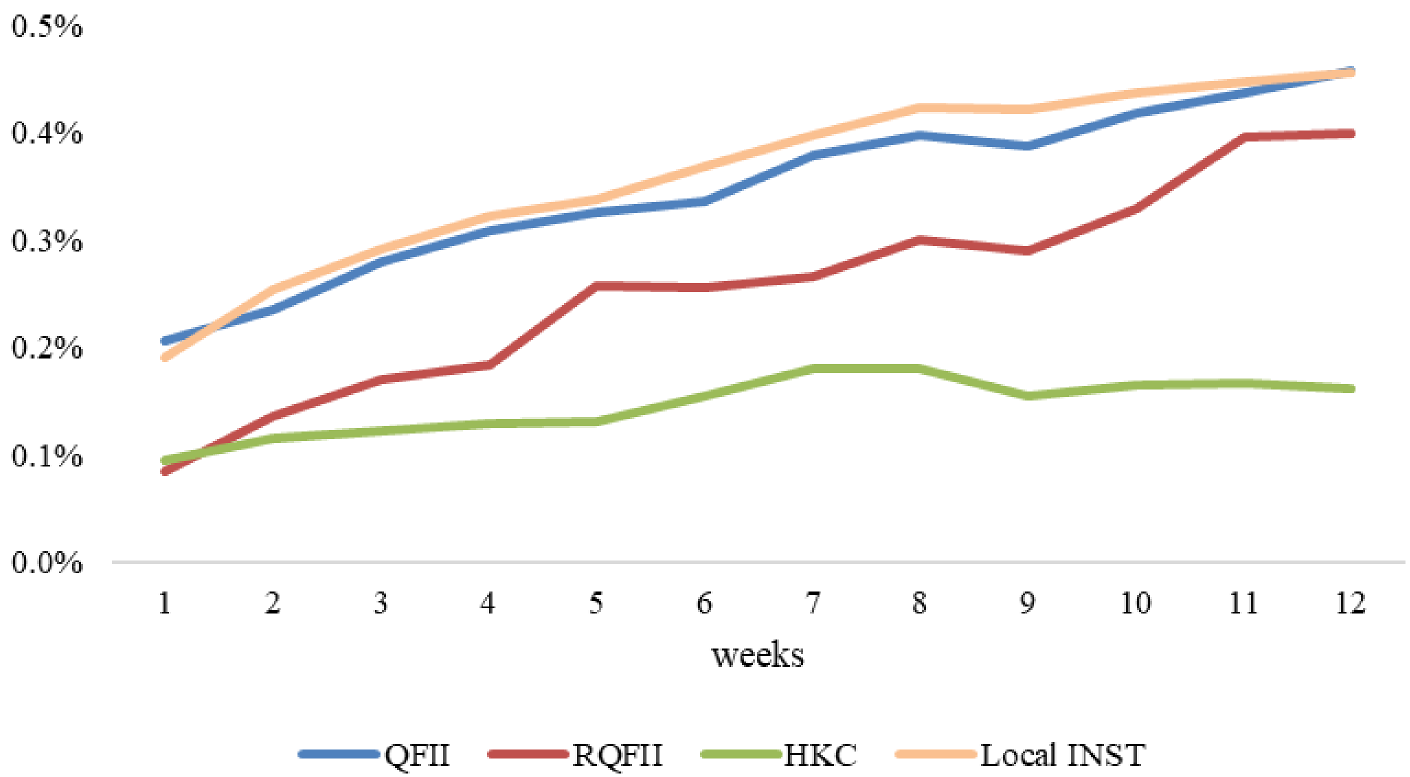


Figure 3. Investors' return predictive power in longer horizons. In this figure, we present the cumulative interquartile returns for different investors over the future 12 weeks, which are calculated as the interquartile of order imbalance multiplied by the coefficients of order imbalance in Table III Panel B.

Internet Appendix for:

"Are Foreign Investors Informed? Trading Experiences of Foreign Investors in China"

Christian T. Lundblad, Donghui Shi, Xiaoyan Zhang, and Zijian Zhang*

A. Construct Analyst Dataset

This section describes how we construct the analyst dataset. To build a comprehensive analyst dataset, we obtain analyst forecasts and recommendations data from four leading data vendors in China: CSMAR, WIND, RESSET and SUNTIME. Following Li, Wong and Yu (2020), we start with the CSMAR analyst database, then add new observations from the other three. To ensure accuracy, we require that the observation in final dataset be recorded in at least two of the four databases with same analyst forecast.

We only include firm-level annual EPS earnings forecasts made for the current fiscal year before the earnings announcements. The stocks' consensus forecast is the arithmetic average of all outstanding EPS forecasts made since the last earnings announcement date (Ivković and Jegadeesh, 2004). We calculate the forecast revision as the current consensus forecast minus the previous consensus forecast. In terms of recommendations, these databases usually divide them into five categories: strong buy, buy, hold, sell, and strong sell. We keep the original rankings in the databases and assign numerical values of 2, 1, 0, -1, and -2 to strong buy, buy, hold, sell, and strong sell, respectively. The analyst's recommendation change is the current numeric recommendation minus the previous recommendation made by the same analyst within one year (Jia, Wang and Xiong, 2017). If no previous recommendation matches, the change is the difference between the current recommendation and zero. Finally, we compute the mean of analyst recommendation at stock-day level.

* Christian T. Lundblad is with Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Donghui Shi is with Fanhai International School of Finance, Fudan University, Xiaoyan Zhang is with PBC School of Finance, Tsinghua University, and Zijian Zhang is with PBC School of Finance, Tsinghua University.

Table IA.I

Stocks characteristics, sectors and investors' trading and holding behaviors

This table presents summary statistics on stock characteristics and sectors conditional on investors' trading and holding behaviors. Our sample period is January 1, 2016, to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. In Panel A, on each day, we sort stocks into two groups based on investors' daily holding shares in percentage of stocks' A-share outstanding. Then we calculate the time-series average of cross-sectional mean of stocks' size (in billion RMB), earnings-to-price ratio and monthly turnover. We also present sectors classified by CSRC with lowest and highest investors' holdings at industry level. In Panel B, we sort stocks into two groups based on investors daily trading volumes in percentage of total volume and report the similar statistics.

Panel A. Size, earnings-to-price ratio, turnover and sectors conditional on investors' holdings

	QFII		RQFII		HKC		Local INST	
	Low	High	Low	High	Low	High	Low	High
Size	10.36	37.19	11.91	35.65	8.27	39.28	19.27	28.29
EP	0.0036	0.0096	0.0045	0.0087	0.0035	0.0097	0.0045	0.0088
Turnover	61.94%	36.24%	67.93%	30.26%	68.99%	29.20%	60.68%	37.50%
Sector	Education	Manufacturing	Education	Manufacturing	Education	Manufacturing	Education	Finance

Panel B. Size, earnings-to-price ratio, turnover and sectors conditional on investors' trading

	QFII		RQFII		HKC		Local INST	
	Low	High	Low	High	Low	High	Low	High
Size	14.31	33.28	12.79	100.02	7.02	43.03	11.07	36.48
EP	0.0055	0.0077	0.0056	0.0133	0.0036	0.0101	0.0038	0.0094
Turnover	60.81%	37.37%	53.94%	22.87%	65.11%	31.10%	68.82%	29.37%
Sector	Education	Manufacturing	Education	Manufacturing	Education	Manufacturing	Education	Manufacturing

Table IA.II
Predicting long-term risk-adjusted stock return

This table presents results about whether foreign investors and local institutions can predict long-term risk-adjusted stock returns. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We adopt the three-factor model in Liu, Stambaugh and Yuan (2019) and estimate cumulative risk-adjusted stock returns over future w weeks following the procedure in Boehmer et al. (2022). The daily Fama-MacBeth (1973) regressions is specified in equation (IA1),

$$\hat{\alpha}(i,w) = a0(d,G) + a1(d,G)Oib(i,d-1,G) + a2(d,G)'Controls(i,d-1) + \epsilon(i,d,G) \quad \text{(IA1)}$$

where the main independent variable is the investor G's order imbalance on previous day. All control variables are same as those in equation (2). To spare the space, we only report the interquartile cumulative risk-adjusted returns over future w weeks, with w from 1 to 12. To account for potential serial correlation in the coefficients, the standard errors are adjusted using Newey-West (1987) with five lags. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Interquartile Cumulative Return	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Week number w	QFII	QFII-Local	RQFII	RQFII-Local	HKC	HKC-Local	Local INST
1	0.1898%***	0.0184%	0.0771%***	-0.0944%***	0.0886%***	-0.0828%***	0.1714%***
2	0.2102%***	-0.0144%	0.1216%***	-0.1033%**	0.1055%***	-0.1239%***	0.2245%***
3	0.2435%***	-0.0059%	0.1557%***	-0.0942%*	0.1176%***	-0.1389%***	0.2493%***
4	0.2740%***	0.0121%	0.1551%***	-0.1075%*	0.1186%***	-0.1482%***	0.2619%***
5	0.2675%***	0.0015%	0.2130%***	-0.0544%	0.1057%***	-0.1679%***	0.2660%***
6	0.2844%***	-0.0075%	0.2101%***	-0.0834%	0.1110%***	-0.1884%***	0.2919%***
7	0.3255%***	0.0092%	0.2123%***	-0.1060%	0.1271%***	-0.1968%***	0.3163%***
8	0.3463%***	0.0016%	0.2302%***	-0.1157%	0.1306%***	-0.2192%***	0.3447%***
9	0.3345%***	-0.0151%	0.2081%**	-0.1434%	0.1031%*	-0.2530%***	0.3496%***
10	0.3619%***	0.0028%	0.2418%***	-0.1190%	0.1105%*	-0.2548%***	0.3591%***
11	0.3860%***	0.0190%	0.3125%***	-0.0563%	0.1154%*	-0.2569%***	0.3670%***
12	0.4119%***	0.0504%	0.3121%***	-0.0507%	0.0995%	-0.2675%***	0.3615%***

Table IA.III
Horse race test

This table presents results on horse race test of foreign investors return predictive power. The sample period is from January 2016 to June 2019. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of the Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. As shown in equation (IA2), we put order imbalances from QFII, RQFII, HKC and local institutions together and estimate Fama-MacBeth regressions.

$$Ret(i,d) = a0(d,G) + a1(d,G)'Oib(i,d-1,Foreign) + a2(d,G)Oib(i,d,Local\ INST) + a3(d)'Controls(i,d-1) + \epsilon(i,d,G). \quad \text{(IA2)}$$

We present the estimated coefficients on order imbalances and related interquartile returns in the horse race test. The dependent variable is expressed as a percentage. All control variables are same as those in equation (2). To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Dep: Ret(d)	(1)
QFII(d-1)	0.0600***
RQFII(d-1)	0.0248***
HKC(d-1)	0.0842***
Local INST(d-1)	0.2410***
Adj-R^2	17.13%
Interquartile Return	
QFII	0.0933%
RQFII	0.0291%
HKC	0.0664%
Local INST	0.1273%

Table IA.IV

Robustness checks for stock return prediction and firm-level news

This table presents robustness results on investors' stock return predictive power. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate quarterly Fama-MacBeth regressions as in equation (IA3),

$$\begin{aligned} Ret(i,d) = {} & c0(q,G) \\ & + [c1(q,G) + c2(q,G)Bignews(i,d) + c3(q,G)NBignews(i,d)] \\ & \times Oib(i,d-1,G) + c4(q,G)Bignews(i,d) + c5(q,G)NBignews(i,d) \\ & + c6(q,G)'Controls(i,d) + \epsilon(i,d,G). \end{aligned} \tag{IA3}$$

The indicator variable *Bignews(i, d)* is equal to one if stock i's return on event day d is outside the 5th and 95th percentiles of all firm-level news day returns, otherwise zero. *NBignews(i, d)* is equal to one if stock i's return on event day d is inside the 5th and 95th percentiles of firm-level news day returns, otherwise it is zero. In Panel A, we use earnings announcements and analyst activities as proxies for firm-level news. In Panel B, we use financial media news as the proxy for firm-level information. Control variables are same as those in equation (2). To spare the space, we omit coefficients of control variables and *t*-statistics in the table. *Adj-R^2* is the time-series average of adjusted R-squared in the cross-sectional regression. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Earnings announcement and analyst activities

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Oib(d-1)	0.0978***	0.0440***	0.0970***	0.2120***
Oib(d-1)×Bignews(d)	0.6296***	0.8099***	0.2045	2.3867***
Oib(d-1)×NBignews (d)	-0.0339**	0.0020	0.0528	-0.0077
Bignews(d)	1.3757***	1.5974***	1.4899***	1.0471***
NBignews(d)	0.2353***	0.2423***	0.2622***	0.2279***
Adj-R^2	1.34%	1.90%	1.51%	1.19%

Panel B. Media news

Dep: Ret(d)	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
Oib(d-1)	0.0910***	0.0366*	0.0928***	0.1838***
Oib(d-1)×Bignews(d)	0.4312***	0.2861**	0.2333	1.5658***
Oib(d-1)×NBignews (d)	-0.0131	0.0085	0.0112	-0.0429
Bignews(d)	1.2390***	1.2591**	1.1337**	1.1179***
NBignews(d)	0.1648***	0.1085***	0.1452***	0.1789***
Adj-R^2	3.26%	3.67%	3.29%	2.91%

Table IA.V

Stock return prediction on firm-level news days with positive and negative returns

This table presents results about investors' return predictive power on positive and negative firm-event days. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate quarterly Fama-MacBeth (1973) regressions as in equation (IA4),

$$\begin{aligned} Ret(i,d) = e0(q,G) + [&e1(q,G) + e2(q,G)PosBignews(i,d) + e3(q,G)NegBignews(i,d) \\ &+ e4(q,G)NBignews(i,d)] \times Oib(i,d-1,G) + e5(q,G)'Controls(i,d-1) \\ &+ \epsilon(i,d,G) \end{aligned} \quad \text{(IA4)}$$

The indicator variable *PosBignews(i, d)* (*NegBignews(i, d)*) is equal to one if stock *i*'s return on firm news day *d* is above (below) the 95th (5th) percentile of all firm news day returns, and otherwise it is zero. *NBignews(i, d)* is equal to one if stock *i*'s return on firm news day *d* is inside the 5th and 95th percentiles of all firm news day returns, and otherwise it is zero. Panel A presents the results for earnings announcements and analyst activities. Panel B presents the results for financial media news. Control variables are same as those in equation (2). To spare the space, we omit coefficients of control variables and *t*-statistics in the table. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. The coefficients are multiplied by 100 for readability. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. Earnings announcements and analyst activities

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
$\widehat{e1}$: Oib(d-1)	0.0977***	0.0437***	0.0954***	0.2121***
$\widehat{e2}$: Oib(d-1)×PosBignews (d)	0.2342	0.7363*	0.5087	2.0315***
$\widehat{e3}$: Oib(d-1)×NegBignews(d)	0.9860***	0.9531**	0.1349	2.8057***
$\widehat{e4}$: Oib(d-1)×NBignews(d)	-0.0342***	0.0292	0.0824**	-0.0042
Interquartile (Oib)× $\widehat{e1}$: $\widehat{Ret}1$(Non-event)	0.1787%	0.0539%	0.0922%	0.1487%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e2})$: $\widehat{Ret}2$(PosBignews)	0.6073%	0.9627%	0.5839%	1.5732%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e3})$:$\widehat{Ret}3$(NegBignews)	1.9827%	1.2302%	0.2226%	2.1161%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e4})$:$\widehat{Ret}4$(NBignews)	0.1161%	0.0899%	0.1719%	0.1457%

Panel B. Media news

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
$\widehat{e1}$: Oib(d-1)	0.0907***	0.0327*	0.0913***	0.1849***
$\widehat{e2}$: Oib(d-1)×PosBignews (d)	0.1687	0.2213	0.2382	0.3799*
$\widehat{e3}$: Oib(d-1)×NegBignews(d)	0.7248***	0.4156	0.4341	2.6589***
$\widehat{e4}$: Oib(d-1)×NBignews(d)	-0.0085	0.0163	0.0158	-0.0551
Interquartile (Oib)× $\widehat{e1}$: $\widehat{Ret}1$(Non-event)	0.0082%	0.0054%	0.0055%	0.0069%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e2})$: $\widehat{Ret}2$(PosBignews)	0.0258%	0.0096%	0.0088%	0.0345%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e3})$:$\widehat{Ret}3$(NegBignews)	0.0470%	0.0189%	0.0323%	0.0284%
Interquartile (Oib)× $(\widehat{e1}+\widehat{e4})$:$\widehat{Ret}4$(NBignews)	0.1141%	0.0277%	0.0607%	0.0892%

Table IA.VI

Robustness checks for stock return prediction and firm-level news

This table presents results about investors' predictive power between earnings announcements and analyst-related events. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate quarterly Fama-MacBeth regressions, as in equation (IA5),

$$\begin{aligned} Ret(i,d) = {} & h0(q,G) \\ & + [h1(q,G) + h2(q,G)BignewsEarn(i,d) + h3(q,G)NBignewsEarn(i,d) \\ & + h4(q,G)BignewsAnalyst(i,d) + h5(q,G)NBignewsAnalyst(i,d)] \\ & \times Oib(i,d-1,G) + h6(q,G)'Controls(i,d) + \epsilon(i,d,G), \end{aligned} \tag{IA5}$$

where *BignewsEarn(i, d)* is equal to one if stock *i*'s return on earnings day *d* is outside the 5th and 95th percentiles of all earnings day returns, otherwise it is zero. *NBignewsEarn(i, d)* is equal to one if stock *i*'s return on earnings day *d* is within the 5th and 95th percentiles of all earnings day returns, otherwise it is zero. *BignewsAnalyst(i, d)* is equal to one if stock *i's* return on analyst activity day *d* is outside the 5th and 95th percentiles of all analyst-related day returns. *NBignewsAnalyst(i, d)* is equal to one if stock *i*'s return on analyst activity day *d* is inside the 5th and 95th percentiles of all analyst-related day returns. Control variables are same as those in equation (2). To spare the space, we omit coefficients of control variables and *t*-statistics in the table. *Adj-*R^2 is the time-series average of adjusted R-squared in the cross-sectional regression. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
$\widehat{h1}$: Oib(d-1)	0.0978***	0.0436***	0.0959***	0.2134***
$\widehat{h2}$: Oib(d-1)×BignewsEarn(d)	0.4794	0.2702	-0.0564	0.8032*
$\widehat{h3}$: Oib(d-1)×NBignewsEarn (d)	0.0235	-0.0908	0.0651	-0.1393***
$\widehat{h4}$: Oib(d-1)×BignewsAnalyst (d)	0.3681***	0.4750*	0.1332	2.6030***
$\widehat{h5}$: Oib(d-1)×NBignewsAnalyst (d)	-0.0370***	0.0418	0.0878**	0.0264
Interquartile (Oib)× $\widehat{h1}$:	0.1789%	0.0538%	0.0927%	0.1496%
Interquartile (Oib)× $(\widehat{h1}+\widehat{h2})$	1.0560%	0.3873%	0.0382%	0.7128%
Interquartile (Oib)× $(\widehat{h1}+\widehat{h3})$	0.2219%	-0.0582%	0.1556%	0.0519%
Interquartile (Oib)× $(\widehat{h1}+\widehat{h4})$	0.8523%	0.6401%	0.2214%	1.9749%
Interquartile (Oib)× $(\widehat{h1}+\widehat{h5})$	0.1111%	0.1055%	0.1776%	0.1681%
Contribution of BignewsEarn days (0.15%)	0.87%	1.01%	0.06%	0.68%
Contribution of NBignewsEarn days (1.37%)	1.67%	-1.39%	2.18%	0.45%
Contribution of BignewsAnalyst days (0.41%)	1.92%	4.57%	0.93%	5.11%
Contribution of NBignewsAnalyst days (3.68%)	2.25%	6.77%	6.70%	3.92%

Table IA.VII
Stock return predictive power and large stock price changes

This table presents results about investors' return predictive power on stocks with large price changes. Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate quarterly Fama-MacBeth regressions as in equation (IA6),

$$\begin{aligned} Ret(i,d) = l0(q,G) + [l1(q,G) + l2(q,G)Bigday(i,d)] \times Oib(i,d-1,G) \\ + l3(q,G)'Controls(i,d-1) + \epsilon(i,d,G) \end{aligned} \tag{IA6}$$

where the indicator variable *Bigday(i, d)* is equal to one if the return for stock *i* on day *d* is outside the 5th and 95th percentile of all sample returns, otherwise it is zero. All control variables are same as those in equation (2). We omit coefficients of control variables and *t*-statistics in the table. The standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
$\widehat{l1}$: Oib(d-1)	0.0683***	0.0346**	0.0742***	0.0868***
$\widehat{l2}$: Oib(d-1)×Bigday(d)	0.3512***	0.1827	0.3636**	1.2889***
Interquartile (Oib)× $\widehat{l1}$	0.1249%	0.0427%	0.0717%	0.0609%
Interquartile (Oib)× ($\widehat{l1} + \widehat{l2}$)	0.7673%	0.2683%	0.4232%	0.9647%
Contribution of big days (10%)	40.57%	41.08%	39.60%	63.78%

Table IA.VIII

Stock return predictive power and Citigroup Economic Surprise Index

This table presents estimation results on whether the return predictive power of foreign investors and local institutions is related to macroeconomic surprise. Our sample period is January 1, 2016 to June 30, 2019. We use Citigroup Economic Surprise Indices (CESI) of China (CNY) and G10 countries (G10) as proxies for macroeconomic surprise. The indicator variable *CESI(d)* is equal to one if the index on day d is outside the 5th and 95th percentile of the index values in the sample period, otherwise it is zero. We apply a two-step regression procedure as shown in (IA7),

$$\widehat{a1}(d,G) = p0(G) + p1(G)CESI(d) + \epsilon(d,G) \tag{IA7}$$

In the first step, we perform OLS regression on each day and obtain the time-series coefficients $\widehat{a1}(d)$ on *Oib(d-1)*. In the second step, we regress the estimated coefficient on the CESI indicator variable. Panel A and Panel B present the second step regression results for CESI CNY Index and CESI G10 Index, respectively. The standard errors are adjusted using Newey-West (1987) with five lags. To spare the space, we omit *t*-statistics in the table. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. CESI CNY Index

Dep: $\widehat{a1}(d)$	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
$\widehat{p0}$: Intercept	0.0663***	0.0286***	0.0763***	0.1329***
$\widehat{p1}$: CESI (d, CNY)	-0.0125	0.0121	0.0200	-0.0078
Interquartile (Oib)× $\widehat{p0}$:$\widehat{Ret}1$	0.1212%	0.0354%	0.0738%	0.0932%
Interquartile (Oib)× $(\widehat{p0}+\widehat{p1})$:$\widehat{Ret}2$(Large surprise)	0.0983%	0.0503%	0.0931%	0.0877%
Contribution of large surprise days (10%)	8.26%	13.65%	12.30%	9.47%

Panel B. CESI G10 Index

Dep: $\widehat{a1}(d)$	(1)	(2)	(3)	(4)
	QFII	RQFII	HKC	Local INST
$\widehat{p0}$: Intercept	0.0632***	0.0297***	0.0794***	0.1328***
$\widehat{p1}$: CESI (d, G10)	0.0177	0.0014	-0.0121	0.0023
Interquartile (Oib)× $\widehat{p0}$:$\widehat{Ret}1$	0.1157%	0.0367%	0.0767%	0.0931%
Interquartile (Oib)× $(\widehat{p0}+\widehat{p1})$:$\widehat{Ret}2$(Large surprise)	0.1481%	0.0384%	0.0651%	0.0947%
Contribution of large surprise days (10%)	12.46%	10.42%	8.61%	10.16%

Table IA.IX
Stock return prediction, dual-listed stocks, and SOEs

This table presents the estimation results on how investors' return predictive power is related to dual-listed stocks and State-Owned Enterprises (SOEs). Our sample period is January 1, 2016 to June 30, 2019, and our sample includes common stocks listed on the main board of Shanghai Stock Exchange (SSE) with at least fifteen non-zero volume trading days in the previous month. We estimate daily Fama-MacBeth regressions as shown in equation (IA8),

$$\begin{aligned} Ret(i,d) = r0(d,G) + [r1(d,G) + r2(d,G)Duallist(i,d-1)] \times Oib(i,d-1,G) \\ +r3(d,G)Duallist(i,d-1) + r4(d,G)'Controls(i,d-1) + \epsilon(i,d,G) \end{aligned} \tag{IA8}$$

Panel A presents results for dual-listed stocks. The dummy variable *Duallist(i, d-1)* is equal to 1 if stock *i* on day *d-1* has a dual-listed H share, otherwise it equals zero. Panel B presents results for SOEs. We replace *Dualist(i, d-1)* with an SOE dummy, *SOE(i, d-1)*, which equals one if the controlling shareholders for stock *i* are state-owned enterprises, and otherwise zero. To spare the space, we omit coefficients of control variables. *Adj-R^2* is the time-series average of adjusted R-squared in the cross-sectional regression. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags. We report *t*-statistics in the parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level.

Panel A. AH dual-listed stocks

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Oib(d-1)	0.0655***	0.0283***	0.0793***	0.1294***
	(16.86)	(3.14)	(10.25)	(17.87)
Oib(d-1)×Duallist(d-1)	-0.0099	0.0078	-0.0199	0.0712***
	(-1.02)	(0.36)	(-1.27)	(3.46)
Adj-R^2	9.01%	14.82%	10.19%	8.90%

Panel B. SOEs

Dep: Ret(d)	(1) QFII	(2) RQFII	(3) HKC	(4) Local INST
Oib(d-1)	0.0683***	0.0203*	0.0879***	0.1371***
	(13.07)	(1.69)	(9.06)	(15.45)
Oib(d-1)×SOE(d-1)	-0.0071	0.0080	-0.0204*	-0.0100
	(-1.16)	(0.56)	(-1.81)	(-1.00)
Adj-R^2	9.30%	15.11%	10.43%	9.14%

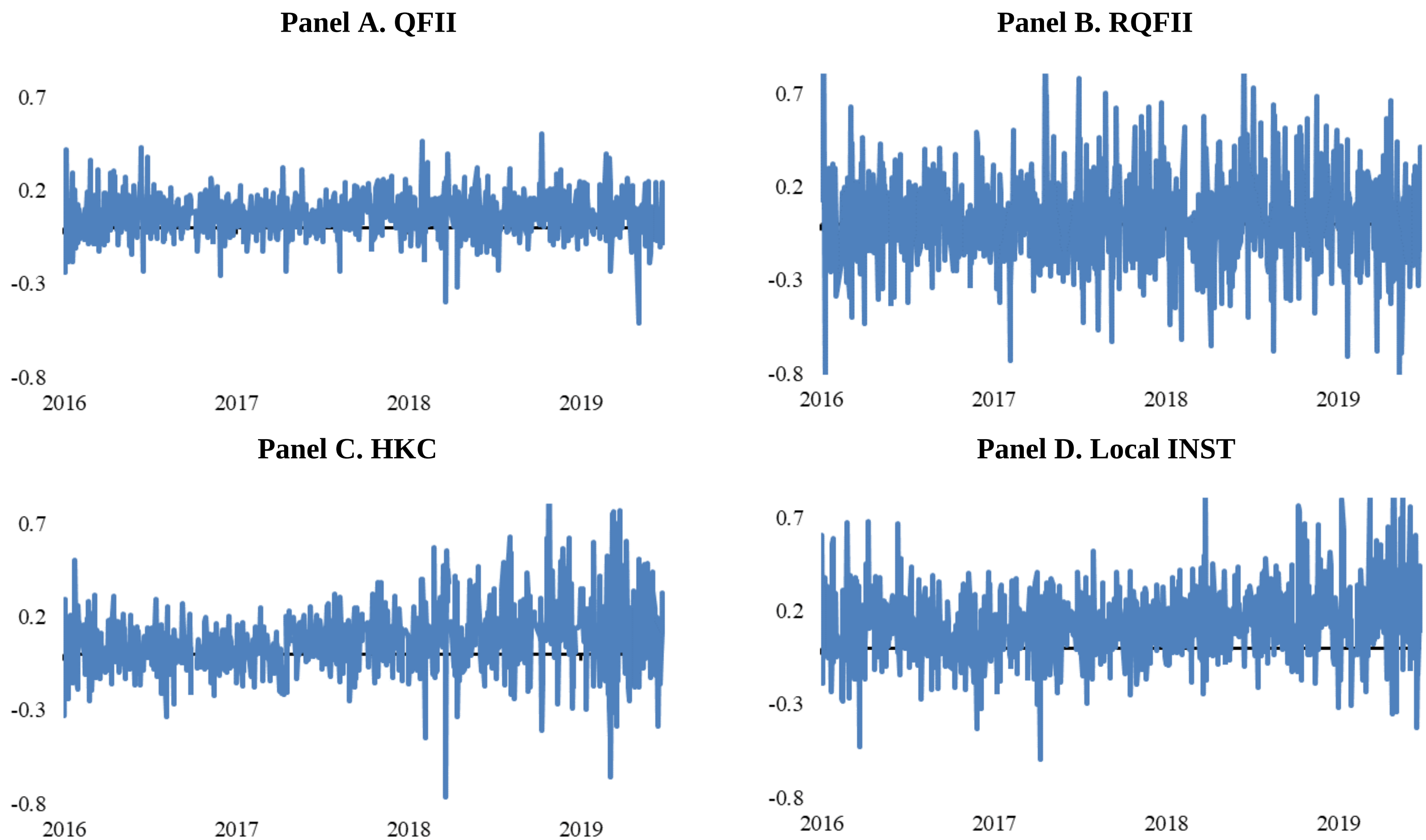


Figure IA.1. The time-series coefficients of the order imbalance in the next day's return prediction. In equation (2), we use the Fama-MacBeth regression to examine investors' predictive power on the next day's stock return. We plot the time-series coefficients on the previous day's order imbalance in the first-stage regression.