The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. , NO. 0 • XXXX 2021

Tracking Retail Investor Activity

EKKEHART BOEHMER, CHARLES M. JONES, XIAOYAN ZHANG,
and XINRAN ZHANG

ABSTRACT

We provide an easy method to identify marketable retail purchases and sales using recent, publicly available U.S. equity transactions data. Individual stocks with net buying by retail investors outperform stocks with negative imbalances by approximately 10 bps over the following week. Less than half of the predictive power of marketable retail order imbalance is attributable to order flow persistence, while the rest cannot be explained by contrarian trading (proxy for liquidity provision) or public news sentiment. There is suggestive, but only suggestive, evidence that retail marketable orders might contain firm-level information that is not yet incorporated into prices.

CAN RETAIL EQUITY INVESTORS PREDICT future stock returns, or do they make systematic, costly mistakes in their trading decisions? The answers to these questions are important for other market participants looking for useful signals about future price moves, for behavioral finance researchers, and for policymakers deciding whether these investors should be protected from themselves.

Ekkehart Boehmer is with Singapore Management University, LKC School of Business. Charles M. Jones is with Columbia Business School. Xiaoyan Zhang is with Tsinghua University, PBC School of Finance. Xinran Zhang is with Central University of Finance and Economics, School of Finance. None of the authors has conflicts of interest to disclose. Jones serves on two related advisory boards: the Nasdaq Quality of Market Committee, which meets quarterly to discuss the state of the U.S. equity market, Nasdaq's innovations in equity trading, and the regulation of U.S. equity trading, and FINRA's Economic Advisory Committee, which meets twice a year to discuss issues related to its supervision of most on-exchange and off-exchange trading of equities in the United States, issues related to its supervision of broker-dealers in the United States, and related macroeconomic issues. Xiaoyan Zhang acknowledges financial support from the National Natural Science Foundation of China [Grant 71790605]. We thank Kevin Crotty, Larry Glosten, Frank Hatheway, Eric Kelley, Jamil Nazarali, Jeff Pontiff, Meijun Qian, Tarun Ramadorai, Paul Tetlock, and seminar participants at Brigham Young University, University of California at Riverside, the Frankfurt School, Humboldt University Berlin, INSEAD, Purdue University, Rice University, Singapore Management University, Southern Methodist University, Tsinghua University, University of Utah, National Singapore University, and Nanyang Technology University, and conference audiences at the 2017 SFS Cavalcade, 2017 German Finance Association, 2017 ABFER, 2017 CICF, and 2018 AFA for helpful comments.

Correspondence: Xiaoyan Zhang, PBC School of Finance, Tsinghua University in Beijing, China 100083; e-mail: zhangxiaoyan@pbcsf.tsinghua.edu.cn.

DOI: 10.1111/jofi.13033

Many researchers have concluded that retail equity investors are generally uninformed and make systematic mistakes when selecting equity investments (see, for example, Barber and Odean (2000, 2008)). More recent studies, however, suggest otherwise (Kaniel, Saar, and Titman (2008), Barber, Odean, and Zhu (2009), Kaniel et al. (2012), Kelley and Tetlock (2013), Fong, Gallagher, and Lee (2014), Barrot, Kaniel, and Sraer (2016)). These studies show that retail investors' trading can predict future stock returns. Unfortunately, most existing studies of retail order flow are based on proprietary data sets with relatively small subsets of overall retail order flow. For example, Barber and Odean (2000) use data from a single U.S. retail brokerage firm, while Barber and Odean (2008) examine individual investor trading data from a total of three different retail or discount brokerage firms. Kelley and Tetlock (2013) use data from a single U.S. wholesaler, Fong, Gallagher, and Lee (2014) analyze data from the Australian Securities Exchange (ASX), and Barrot, Kaniel, and Sraer (2016) use data from a single French brokerage firm. Kaniel, Saar, and Titman ((2008), Kaniel et al. (2012), and Boehmer, Jones, and Zhang (2008) use proprietary account-type data from the NYSE during the early 2000s. During that period, only a small number of brokerages sent their retail order flow to the NYSE, and thus, the NYSE's market share of overall retail order activity was (and has remained) quite small.

In existing work, many researchers use trade size as a proxy for retail order flow. Before the spread of computer algorithms that "slice and dice" large institutional parent orders into a sequence of small child orders, small trades were much more likely to come from retail customers, while institutions were likely behind the larger reported trades. For example, Lee and Radhakrishna (2000) use a \$20,000 cutoff to separate smaller individual trades from larger institutional trades. More recently, Campbell, Ramadorai, and Schwartz (2009) allow the cutoff to vary using through a regression approach that is calibrated to observed quarterly changes in institutional ownership, but they maintain the same basic assumption that small trades are more likely to arise from individual trading. However, with algorithms becoming an important feature of institutional order executions in the early 2000s, a trade-size partition has become far less useful as a proxy for retail order flow. Indeed, the tendency of algorithms to slice orders into smaller and smaller pieces has progressed so far that during our recent sample period, the retail order flow that we identify actually has a slightly larger average trade size than other flow.

Given the current automated and algorithm-driven market structure, researchers need an alternative measure to isolate retail order flow. We introduce such a measure in this paper. As one of our main contributions, we show that our measure can identify a broad swath of marketable retail order flow. Our measure builds on the fact that, due to regulatory restrictions in the United States and the resulting institutional arrangements, retail order flow—but not institutional order flow—can receive price improvement, measured in small fractions of a cent per share. We use this observation to identify marketable retail price-improved orders from the Trade and Quote (TAQ) data, a publicly available data set that contains all transactions for stocks listed on a national exchange in the United States. Specifically, we identify trades that execute at share prices with fractional pennies. Most such price-improved transactions take place off-exchange and are reported to a Trade Reporting Facility (TRF). Using these TRF data, we identify transactions as retail buys if the transaction price is slightly below the round penny and as retail sells if the transaction price is slightly above the round penny. This approach separates retail investors' marketable orders from those of institutions because institutional trades generally cannot receive this type of fractional penny price improvement.[1] We discuss our approach in greater detail in Section I.B. Notice that our retail order flow measure only includes marketable orders, that is, it does not include limit orders. Overall, we believe that our method of retail trade identification is conservative, and we cross-validate the accuracy of our approach using a small sample of NASDAQ TRF audit trail data.

We analyze retail marketable order flow from the U.S. equity market for the six years between January 2010 and December 2015. We find that retail investors are slightly contrarian at a weekly horizon, and that the cross-section of weekly marketable retail order imbalances predicts the cross-section of returns over the next several weeks, consistent with the findings of Kaniel, Saar, and Titman (2008), Kaniel et al. (2012), Kelley and Tetlock (2013), Fong, Gallagher, and Lee (2014), and Barrot, Kaniel, and Sraer (2016) but inconsistent with the findings of many others.

The predictability of marketable retail order flow for future returns is consistent with persistence in retail order flow, liquidity provision, and informed trading. We conduct a decomposition exercise and separate the marketable retail order imbalance into proxies for these three components. The empirical findings show that persistence in order flow and order flow driven by return reversals (our proxy for liquidity provision) account for about half of the predictive power of the marketable retail order imbalance for future returns. We attribute the other half of this predictive power to potential informed trading.

We also investigate the nature of the information embedded in retail trading. Our results show that the marketable retail order imbalance is positively correlated with some firm-level surprises in public news, and that marketable retail order flow has predictive power beyond public news, which suggests (but only suggests) that retail investors may possess firm-level information that is not yet incorporated into prices.

Finally, we conduct a battery of robustness checks and provide further discussion. We find that our results are robust, and we provide additional evidence that, despite the predictive power of marketable retail order flow in the cross-section, aggregate marketable retail flows cannot predict future market returns.

[1] In contrast, institutional trades often occur at the midpoint of the prevailing bid and ask prices. If the bid-ask spread is an odd number of cents, the resulting midpoint trade price ends in a half-penny. Many of these midpoint trades take place on crossing networks and are reported to the TRF. Thus, trades at or near a half-penny are likely to be from institutions and are not assigned to the retail category.

Given the nature of our data, our work is also related to recent studies of off-exchange trading in the United States. For instance, Kwan, Masulis, and McInish (2015) study the competition between traditional stock exchanges and new dark-trading venues and find that the minimum pricing increment regulation (typically one penny) drives orders to dark pools and limits the competitiveness of the exchanges. Battalio, Corwin, and Jennings (2016) examine make-take fees and how brokers route order flow, and suggest that current order-routing practices may not maximize the quality of limit order execution. Menkveld, Yueshen, and Zhu (2017) directly investigate the pecking order of trading venues in dark pools and document that investors strategically put low-cost, low-immediacy orders in front of high-cost, high-immediacy orders.

Compared to the earlier literature on retail orders and studies of off-exchange trades, we make three main contributions. First and most importantly, we propose a novel methodology for identifying and signing marketable retail trades using publicly available data with substantial coverage. Second, we show that the marketable retail trades that we identify can predict the cross-section of future stock returns. Third, we analyze the nature of the predictive power of marketable retail order flow and show that half of its predictability is likely driven by order imbalance persistence and liquidity provision, while the other half is consistent with informed trading. We also track potential informed trading to different types of news and provide suggestive evidence on the nature of the information possessed by these retail investors.

Two studies, Kaniel, Saar, and Titman (2008) and Kelley and Tetlock (2013), study similar questions and are closely related to our research, but they employ different data and draw different interpretations. For instance, using proprietary data from the NYSE between January 2000 and December 2003, Kaniel, Saar, and Titman (2008) document that retail order flows can predict stock returns. Kaniel, Saar, and Titman (2008) examine the contemporaneous relation between their retail order flows and stock returns. They find that the contemporaneous return is significantly positive for stocks that retail investors sell and negative for stocks that they buy, which is consistent with a liquidity provision interpretation and inconsistent with the information story. We follow their approach using our new marketable retail order flow variables. We are able to replicate the predictive relation between retail order flow and future stock returns, but our results for the contemporaneous relation are different: the contemporaneous return is significantly negative for stocks that retail investors sell and positive for stocks that they buy when they use marketable orders. Our findings are more in line with an information interpretation than a liquidity provision interpretation.

Kelley and Tetlock (2013) obtain data from a major retail wholesaler between February 2003 and December 2007. Their data allow them to separate retail orders into market orders and limit orders. They find that both retail market orders and limit orders can predict future stock returns, but for different reasons. The aggressive market orders can correctly predict future news, suggesting that these trades are informed, while the passive limit orders are contrarian, consistent with liquidity provision. Our marketable retail order flow measure only identifies market orders, and for these marketable orders, we follow their tests and replicate their results. In addition, we decompose our marketable retail order imbalance into components related to order flow persistence, contrarian trading, public news, and a residual, which potentially contains nonpublic information. The decomposition exercise shows that public news contributes little to the predictive power of marketable retail trades, whereas the residual part is more important. With more recent data and wider coverage, our study provides interesting new findings that complement the studies by Kaniel, Saar, and Titman (2008) and Kelley and Tetlock (2013).

The remainder of this paper is organized as follows. We describe the data and our identification method in Section I. Section II presents our main empirical results. We provide further discussion of the results and perform robustness and plausibility checks in Section III. Finally, Section IV concludes the paper.

I. Identifying Retail Order Flows

As we note in the introduction, our most important contribution is to provide a simple, new method to identify a wide swath of marketable retail order flow using publicly available equity transaction data. We introduce our data sources in Section I.A. In Section I.B, we provide the institutional background. Sections I.C and I.D report summary statistics and results of cross-validation tests, respectively.

A. Data Sources

From TAQ trade data, we start with only trades that occur off-exchange, designated with exchange code "D." We merge these TAQ data with stock returns and accounting data from CRSP and Compustat, respectively. We include only the common stocks with share code 10 or 11 (which mainly excludes ETFs, ADRs, and REITs) listed on the NYSE, NYSE MKT (formerly Amex), and NASDAQ. We remove low-priced stocks by requiring that the minimum stock price be $1 at the previous month-end.

Our sample spans the period January 3, 2010 to December 31, 2015. Data on subpenny price improvement actually extend back to 2005. In Internet Appendix Figure IA.1[2], we plot the time series from January 2005 (the start of Regulation National Market System, or Reg NMS, which established the current regulatory framework for subpenny price improvement in the United States) to December 2017. We choose to study the 2010 to 2015 period for two reasons. First, during the first few years under Reg NMS, there is a strong upward trend in the number of subpenny trades, possibly because an increasing number of brokerage firms were adopting the practice of providing fractional cents of price improvement to retail investors via internalization or wholesalers. The upward trend disappears and stabilizes after 2009. Second, from

[2] The Internet Appendix is available in the online version of this article on *The Journal of Finance* website.

2016 to September 2018, the SEC adopted a tick size pilot program (TSPP) that affected tick size and brokers' ability to provide price improvement for many stocks, which likely affected the prevalence of subpenny price improvements unevenly in the cross-section. Our main analysis therefore focuses on the middle part of these data, from 2010 to 2015. For each day, we have an average of around 3,000 firms in the sample.

B. Institutional Background and Methodology

In the United States, most marketable equity orders initiated by retail investors do not take place on one of the dozen or so registered exchanges. Instead, these retail orders are typically executed by wholesalers or via internalization, meaning that orders are filled from a broker's own inventory. Orders executed by wholesalers or through internalization must be publicly reported; they are usually reported to a Financial Industry Regulatory Authority (FINRA) TRF, which provides broker-dealers with a mechanism through which to report transactions that occur off-exchange. These TRF executions are then included in the TAQ "consolidated tape" of all reported transactions with exchange code "D." Many orders that are internalized or executed by wholesalers are given a small-price improvement relative to the national best bid or offer (NBBO).[3] For instance, wholesalers are willing to provide a small price improvement to induce the retail trader's broker to route the order to the wholesaler. Internalizers, who are subject to Regulation 606T, need to show that they execute their clients' orders optimally and thus also have incentives to provide price improvement to their clients. This price improvement is typically only a small fraction of a cent. Common price improvement amounts are 0.01, 0.1, and 0.2 cent.

Brokerage firms in the United States are required to provide regular summary statistics in SEC Rule 606 filings about their order-routing practices for nondirected orders. A directed order instructs the broker to execute an order on a given exchange or trading venue, while a nondirected order gives the broker discretion regarding the execution venue. The vast majority of retail orders are nondirected. For example, Charles Schwab reports that 98.6% of their security orders during the second quarter of 2016 were nondirected orders. The corresponding figure for TD Ameritrade is 99%. According to the Rule 606 filings by these two retail brokerage firms, more than 90% of these orders receive price improvement.

Our communications with a major retail wholesaler and a major exchange suggest that these types of price improvement are not a feature of institutional order executions, as institutional orders are almost never internalized or sold

[3] As a rough estimate of the frequency of subpenny price improvement, we find in an NASDAQ subsample used for robustness tests (introduced in Section I.D) that 60% of trades on "retail" venues receive subpenny price improvements, with 14% reported at the half-penny and 46% taking place at a different subpenny. For subpenny trades that do not execute at half-pennies and constitute the focus of our study, more than 99% are reported to a TRF with exchange code "D."

to wholesalers. Instead, their orders are sent to exchanges and dark pools, and Reg NMS prohibits these orders from having subpenny limit prices. Thus, institutional transaction prices are usually in round pennies. The only exception applies to midpoint trades. Reg NMS has been interpreted to allow executions at the midpoint between the best bid and best offer. As a result, institutions are heavy users of crossing networks and midpoint peg orders that generate transactions at this midpoint price. Since the quoted spread is now typically 1 cent per share, this means that many institutional transactions are reported at a half-penny price. In the early part of our sample, a small number of dark pools allowed some subpenny orders and provided nonmidpoint subpenny execution prices, but our results hold when we exclude this subperiod.[4]

Based on these institutional arrangements, identifying transactions initiated by retail customers is fairly straightforward. Transactions with a retail seller tend to be reported on a TRF at prices that are just above a round penny due to the small price improvement, while transactions with a retail buyer tend to be reported on a TRF at prices just below a round penny. More precisely, for all trades reported to an FINRA TRF (exchange code "D" in TAQ), let P_{it} be the transaction price in stock i at time t, and let $Z_{it} \equiv 100 * \text{mod}(P_{it}, 0.01)$, where $Z_{it} \in [0, 1)$ be the fraction of a penny associated with that transaction price. If Z_{it} is in the interval (0,0.4), we identify it as a retail sell transaction. If Z_{it} is in the interval (0.6,1), the transaction is coded as a retail buy transaction. To be conservative, transactions at a round penny ($Z_{it} = 0$) or near the half-penny ($0.4 \leq Z_{it} \leq 0.6$) are not assigned to the retail category.

As discussed above, Reg NMS requires that limit orders be priced at round pennies, so by definition, our approach will only identify marketable retail orders.[5] The 606 filings by brokerage firms are also partitioned into market and limit orders, which allows us to gauge the relative prevalence of these two types of orders. For example, the Charles Schwab brokerage firm reports that in the second quarter of 2016, market orders account for 50.0% of its customers' nondirected orders in NYSE-listed securities, while limit orders account for 45.1% and other orders account for the remainder. For securities listed on

[4] According to SEC litigation releases (see, for example, https://www.sec.gov/litigation/admin/2015/33-9697.pdf and https://www.sec.gov/litigation/admin/2016/33-10013.pdf), at least two dark pool operators (Credit Suisse and UBS) were accused of violations of Regulation NMS in accepting, ranking, and executing orders based on subpenny prices. These alleged violations occurred through mid-2011 and were eventually settled. A back-of-the-envelope calculation suggests that these violations could have accounted for about 0.5% of total share volume during this part of our sample period. Since these dark pools cater to institutions, including high-frequency traders, our identification of retail flows using subpenny trades could be "contaminated" during this period, and we cannot use public TAQ data to identify which trades are from the affected dark pools. Given the potential contamination accounts for a small part of overall subpenny trades in our sample period, our main analysis still focuses on the period 2010 to 2015. For robustness, we conduct subsample analysis for the period 2012 to 2015 in Section III.B and find that the results are similar to those for the full sample. We thank the Associate Editor for pointing this out.

[5] Marketable orders by definition demand immediacy and, according to Kelley and Tetlock (2013), market orders are more informed than limit orders. Thus, any predictive power from retail market orders is likely to be stronger than that of retail limit orders and overall retail orders.

Table I
Summary Statistics

This table reports summary statistics for our measure of marketable retail investor trading activity. Our sample period is January 2010 to December 2015, and our sample firms are common stocks listed on all U.S. stock exchanges with a share price of at least \$1. Across all stocks and all days, we report the pooled sample mean for the daily number of shares traded (*vol*), marketable retail buy volume (*mrbvol*), marketable retail sell volume (*mrsvol*), number of trades (*trd*), marketable retail buy trades (*mrbtrd*), marketable retail sell trades (*mrstrd*), as well as their odd lot counterparts (prefix *odd*). Odd lot measures are available starting at the end of 2013. We include data related to odd lots starting January 2014. We compute retail order imbalance measures (variables containing *mroib*) as in equations (1) to (4).

	N	Mean	Std	Median	Q1	Q3
Round Lots and Odd Lots						
Vol	4,628,957	1,229,004	6,849,849	221,234	51,768	819,615
Trd	4,628,957	5,917	13,909	1,505	312	5,502
Mrbvol	4,628,957	42,481	280,474	5,165	1,200	20,681
Mrsvol	4,628,957	42,430	264,704	5,635	1,369	21,828
Mrbtrd	4,628,957	110	410	22	5	79
Mrstrd	4,628,957	108	355	24	6	81
Mroibvol	4,628,957	−0.038	0.464	−0.027	−0.301	0.217
Mroibtrd	4,628,957	−0.032	0.437	−0.010	−0.276	0.205
Odd Lots Only						
Oddvol	1,446,749	6,561	20,141	1,811	629	5,250
Oddtrd	1,446,749	222	669	64	21	186
Oddmrbvol	1,446,749	1,108	5,054	211	58	690
Oddmrsvol	1,446,749	968	3,488	210	62	663
Oddmrbtrd	1,446,749	37	171	7	2	23
Oddmrstrd	1,446,749	33	114	7	2	23
Oddmroibvol	1,446,749	−0.004	0.559	0.014	−0.338	0.331
Oddmroibtrd	1,446,749	−0.017	0.506	0.000	−0.290	0.250

NASDAQ, limit orders are slightly more prevalent than market orders at Schwab, with market orders accounting for 44.0% and limit orders 50.7%. Note that nonmarketable limit orders may be canceled without being executed, so most overall retail trading activity is likely to arise from marketable orders. Our approach is therefore likely to pick up a majority of the overall retail trading activity.[6]

C. Summary Statistics

Table I presents summary statistics on the marketable retail orders identified by our method. We pool observations across stocks and days, and compute the mean, standard deviation, median, and 25th and 75th percentiles. Our sample comprises over 4.6 million stock-day observations. For the number of

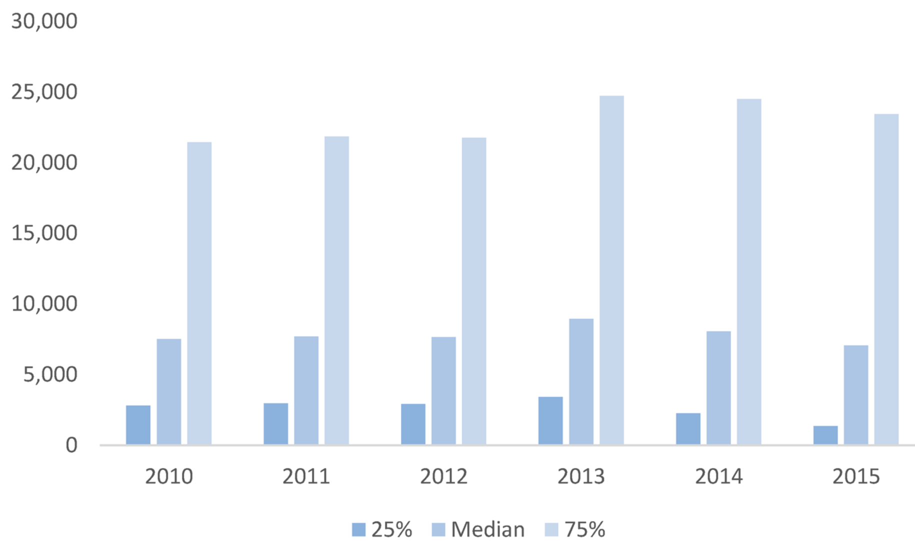
[6] One might wonder whether market or marketable orders can be offset, in aggregate, by limit or nonmarketable orders in the opposite direction. This is possible. Unfortunately, we do not have data to directly check this possibility.

shares traded per day (*vol*), the mean share volume is around 1.23 million, and the standard deviation is about 6.85 million shares. The average stock has 5,917 trades each day (*trd*). These numbers suggest that the average trade size over the sample period is about 200 shares. Our identified marketable retail investor activity represents only a small part of overall trading volume. The identified average daily buy volume from marketable retail orders (*mrbvol*) is 42,481 shares, and the average daily sell volume from marketable retail orders (*mrsvol*) is 42,430 shares. Throughout the paper, we use "*mr*" to represent "marketable retail." We therefore identify an average of 84,911 shares per stock-day traded by marketable retail orders, about 6.91% of the average total shares traded each day. The average number of buy trades from marketable retail orders (*mrbtrd*) each day is 110, and the average number of sell trades from marketable retail orders (*mrstrd*) each day is 108. Thus, the total number of identified trades per stock-day from marketable retail orders is 218, around 3.68% of the total number of trades. Interestingly, the buy volumes closely match the sell volumes, and the number of buy trades matches the number of sell trades, both indicating that many marketable retail trades offset each other. In terms of average share volumes and number of trades, there is slightly more buying than selling by marketable retail trades over our sample period.

Information on odd lot trades (trades of fewer than 100 shares) is reported on the TRF and on the consolidated tape beginning in December 2013 (see O'Hara, Yao, and Ye (2014)). During the December 2013 to December 2015 period, for which odd lot data are available, the daily averages of odd lot marketable retail buy and sell volumes (*oddmrbvol* and *oddmrsvol*, respectively) are 506 and 443 shares, respectively, totaling 949 shares traded by marketable retail investors in odd lots per average stock-day. This is about one-third of the total odd lot share volume at 3,027 shares. The pattern for the number of trades is similar. Prior studies of odd lots generally find that these marketable retail-dominated orders are virtually uninformed, so later in the paper, we study odd lots separately to determine whether the information content of odd lots executed by marketable retail trades differs from that of marketable retail round lots.

Figure 1 provides further statistics on the overall properties of our identified marketable retail trades. Panel A presents trade size in dollars. For each marketable retail trade, we compute its trade size in dollars by multiplying the number of executed shares by the transaction price. For each year in our sample, we compute the 25^{th} percentile, median, and 75^{th} percentile of marketable retail trade size. The median marketable retail trade size is around \$8,000, and the interquartile range is mostly between \$2,000 and \$25,000. Panel B shows the distribution of subpenny prices. We separate all trades into 12 groups or bins. In particular, we separate out trades that take place at a round penny or a half-penny, and the other bins are each 0.1 cent wide. We pool the sample across days and stocks and show the number of shares reported in the different subpenny buckets. Not surprisingly, most of the share volume occurs at round and half-pennies, with average stock-day share volumes of around 27,000 and 7,000, respectively. The next most prevalent occurrence, averaging

Panel A. Marketable retail order trade size in dollars



Panel B. Median share volumes for different subpenny groups

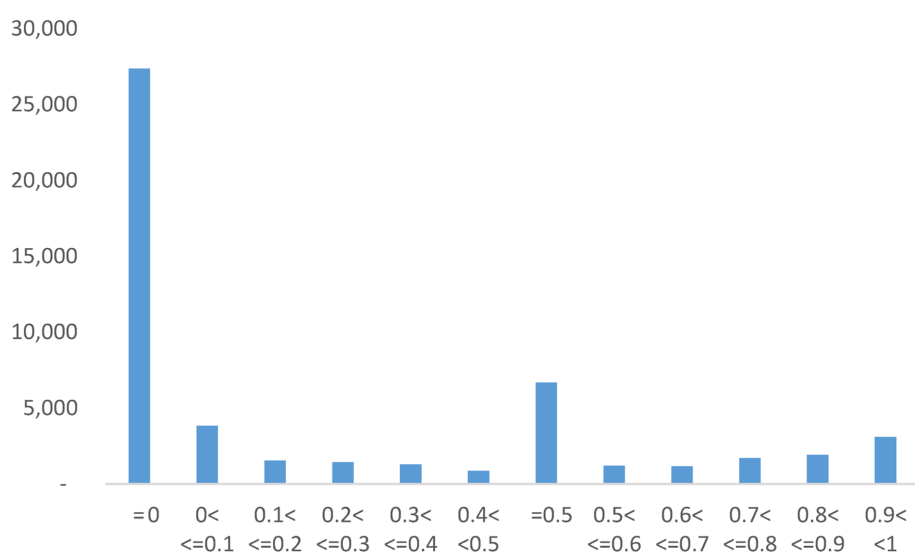


Figure 1. Distribution of trade size and subpenny prices for marketable retail orders. These figures graph summary statistics for the marketable retail investor trading we identify. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. In Panel A, we compute the trade size in dollars as the number of shares multiplied by transaction price. For each year, we report the cross-sectional median, 25th percentile, and 75th percentile. In Panel B, we separate trades into 12 groups based on subpenny increments: trades at the whole penny, at the half penny, and in buckets that are 0.1 cent wide. We show the cross-sectional median of the daily number of shares traded in each group.
(Color figure can be viewed at wileyonlinelibrary.com)

around 3,000 shares per day per stock, is a subpenny price within 0.1 cent of a round penny. Other subpenny bins are less prevalent, with most averaging around 1,000 shares per stock-day.

We measure marketable retail investors' directional trades by computing four order imbalance measures for each stock i on each day t:

$$mroibvol\ (i,t) = \frac{mrbvol\,(i,t) - mrsvol\,(i,t)}{mrbvol\,(i,t) + mrsvol\,(i,t)}\,, \tag{1}$$

$$mroibtrd\ (i,t) = \frac{mrbtrd\,(i,t) - mrstrd\,(i,t)}{mrbtrd\,(i,t) + mrstrd\,(i,t)}\,, \tag{2}$$

$$oddmroibvol\ (i,t) = \frac{oddmrbvol\,(i,t) - oddmrsvol\,(i,t)}{oddmrbvol\,(i,t) + oddmrsvol\,(i,t)}\,, \tag{3}$$

$$oddmroibtrd\ (i,t) = \frac{oddmrbtrd\,(i,t) - oddmrstrd\,(i,t)}{oddmrbtrd\,(i,t) + oddmrstrd\,(i,t)}\,. \tag{4}$$

The first two measures are calculated using marketable retail round lot executions between January 2010 and December 2015 and by aggregating round lot and odd lot executions thereafter, while the last two measures are calculated using marketable retail odd lots and thus begin in December 2013 instead of December 2010.

Summary statistics on the marketable retail order imbalance measures are reported at the bottom of Table I. Across all stock days, the mean order imbalance for share volume, *mroibvol*, is −0.038 with a standard deviation of 0.464, and the mean order imbalance for trade, *mroibtrd*, is −0.032 with a standard deviation of 0.437. The correlation between *mroibtrd* and *mroibvol* is around 85%. Our discussions below mostly focus on *mroibvol*, but the results using these two measures are quite similar given the high correlation between the two. Overall, the order imbalance measured in shares is close to zero on average but sells are slightly more prevalent than buys, consistent with the findings in Kaniel, Saar, and Titman (2008). More importantly, the sizable standard deviation measures indicate that there is substantial cross-sectional variation in the activity levels and trading direction of retail investors. The odd lot order imbalance measures exhibit similar patterns.

In Figure 2, we plot the time series of the cross-sectional mean, median, and 25^{th} and 75^{th} percentiles of the marketable retail order imbalance measures over the six-year sample period. Across all four order imbalance measures, the means and medians are all close to zero, while the 25^{th} percentiles are mostly around −0.3 and the 75^{th} percentiles are mostly around 0.2. There are no obvious time trends or structural breaks in the time-series observations.

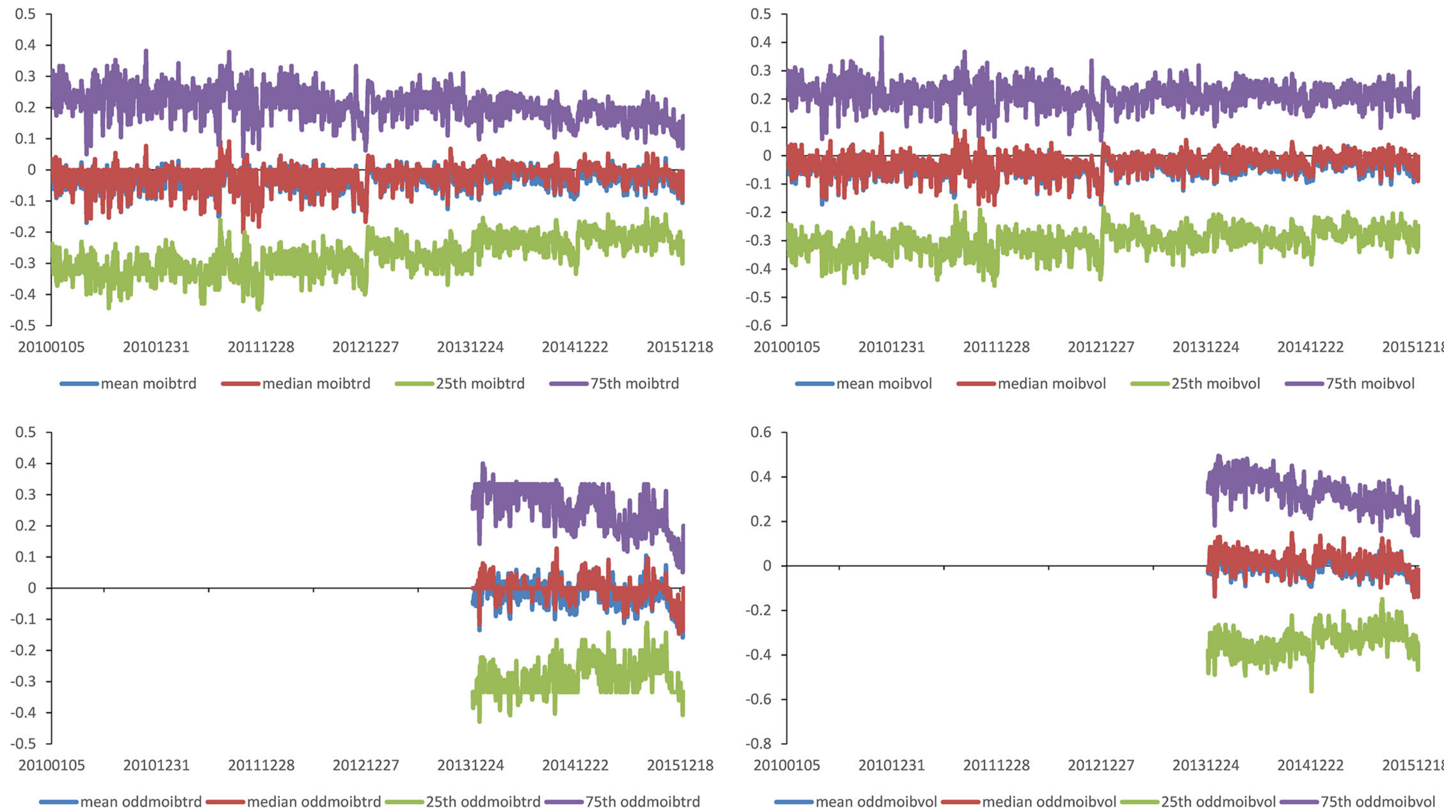


Figure 2. Time series of marketable retail investor order imbalances. These figures plot time-series statistics of identified marketable retail investor trading activity. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We plot the cross-sectional mean, median, 25th percentile, and 75th percentile for each day. (Color figure can be viewed at wileyonlinelibrary.com)

We extensively examine other properties of the marketable retail order imbalance measures. To save space, we report them in the Internet Appendix. The order imbalance measure's daily autocorrelations are reported in Figure IA.2, Panel A. The daily order imbalance measures are mostly significantly positively correlated with their nearby lags, while the cross-firm median correlation is 0.15. This positive autocorrelation is statistically significant over horizons up to a few months. The persistence of marketable retail order flow is slightly higher for larger firms than for smaller firms. Figure IA.2, Panel B, presents the time-series correlation between the marketable retail order imbalance measure and past returns. The results display a V-shape, indicating that the correlation between the current marketable retail imbalance and the previous one-day return is positive on average (consistent with momentum trading), and then becomes negative (consistent with contrarian trading) for the next 30 trading days. Finally, Internet Appendix Table IA.I reports results for the measure's seasonality and its relation with variables reflecting firm fundamentals.

D. Cross-Validation Using NASDAQ TRF Data

Our main data source is TAQ, which does not provide direct information on the direction of trade or the identity of traders. We validate our marketable retail order imbalance algorithm using a small sample of proprietary NASDAQ data.[7] The same data set is used in Menkveld, Yueshen, and Zhu (2017), who provide more details about the data. The NASDAQ sample covers all intraday transactions on its TRF for 117 stocks in October 2010. The 117 stocks are chosen from different size groups, but they are generally larger than a typical firm in TAQ.[8]

For each trade, the NASDAQ TRF data provide a trade direction indicator: "buy," "sell," or "cross." Our algorithm identifies all subpenny trades with subpenny prices between 0.61 and 0.99 cents inclusive as "buy" trades. We separate all subpenny "true buy" trades (as indicated in the NASDAQ TRF data) with a price below \$100 into two categories: "identified buy" and "identified sell (false identification)." We falsely identify 1.37% of all subpenny "true buy" trades as "sell." Similarly, our algorithm identifies all subpenny trades with subpenny prices between 0.01 and 0.39 basis points as "sell" trades. In this case, we falsely identify 2.12% of all subpenny "true sell" trades with a price below \$100 as "buy." If we group identified marketable retail "buys" and "sells" together, for stocks with a share price below \$100, our subpenny approach matches the NASDAQ TRF's correct buy/sell sign 98.2% of the time, while the standard Lee and Ready (1991) trade-signing algorithm gets the trade sign right 96.7% of the time. Overall, we find that our algorithm identifies trade direction accurately.

[7] We thank NASDAQ for generously providing the data.

[8] The smallest market cap of the 117 NASDAQ firms is 257 million dollars, while our sample's 40^{th} percentile market cap is merely 243 million dollars.

Menkveld, Yueshen, and Zhu (2017) explain that when order flows come in, they are routed to different types of off-exchange venues depending on the cost and immediacy of the trade execution. The NASDAQ TRF data identify five types of off-exchange venues: DarkNMid, DarkMid, DarkOther, DarkPrintB, and DarkRetail. Our communication with a major marketable retail wholesaler and the NASDAQ indicates that other than DarkRetail, the venue types are a mix of all kinds of traders. Thus, the venue is not a precise indicator of a trader's identity, and thus even if one had access to the NASDAQ TRF sample for a larger cross-section over a longer period of time, there would still be an important role for our algorithm in identifying marketable retail buys and sells.

Our main measure in this article is order imbalance. The correlation between our order imbalance measure and that calculated using the DarkRetail order imbalance for the 117 stocks is 0.70. This correlation is less than 1 for two main reasons. First, our order imbalance measure includes trades printed on the competing NYSE TRF, while the NASDAQ TRF data set does not. Second, our order imbalance measure includes some subpenny trades from the DarkNMid and DarkMid venues, in addition to those in DarkRetail. Moreover, some marketable retail market orders do not receive price improvement or receive a full half-cent of price improvement. We do not sign these trades or include them in our marketable retail sample because we cannot be sure that we have the correct trade direction. Nevertheless, the high correlation between our marketable retail order imbalance measure and the actual NASDAQ DarkRetail venue data strongly suggests that our order imbalance measures closely reflect the true marketable buy and sell activities of retail investors.[9]

II. Empirical Results

In the previous section, we measure order imbalances at the daily level to minimize the amount of aggregation. For our main empirical analysis, we focus on weekly horizons to reduce the impact of microstructure noise on our results. That is, our main variables of interest are firm-level average marketable retail order imbalances over five-day horizons and five-day firm-level stock returns. Blume and Stambaugh (1983) show that using the end-of-day closing price to compute daily returns can generate an upward bias due to bid-ask bounce.

[9] Kelley and Tetlock (2013) compute retail order imbalance measures using data from one large wholesaler. As part of a conference discussion of our paper, Kelley computed the retail order imbalance measure for 2007 using our algorithm and found that the correlations between our measure and their measure ranged between 0.345 and 0.507 when defining marketable retail flow using different subpenny ranges. For instance, 0.345 is the correlation between our measure and their measure using the number of shares for subpenny prices in the (0, 0.4) and (0.6, 1) cent intervals, while 0.507 is the correlation between our measure and their measure using the number of trades for subpenny prices at 0.99 and 0.01 cents. These correlations should be less than 1 because their flow comes from only one wholesaler while our measure comes from TRF, which covers nearly all retail order executions. We are grateful to Eric Kelley for computing and sharing these calculations with us.

We therefore compute two versions of weekly returns, one by compounding CRSP daily returns based on daily closing prices and one by compounding daily returns using the end-of-day bid-ask average price. We always report results for both types of returns but we focus attention on returns based on closing bid-ask averages.

We start by investigating the properties of the order imbalance measures in Section II.A. In Section II.B, we examine whether past marketable retail order imbalance measures can predict future stock returns using Fama-MacBeth (1973) regressions and long-short portfolios. In Section II.C, we compare alternative hypotheses for the predictive power of marketable retail order imbalances for future stock returns. In Section II.D, we explore the nature of the information contained in marketable retail flow by linking it to Thomson Reuters News Analytics data.

A. What Explains Marketable Retail Investor Order Imbalances?

We start our empirical investigation by examining what drives the trading of retail investors. Specifically, we examine how retail investors' marketable order flow is related to past order flow and past returns. To allow maximal time-series flexibility and focus on cross-sectional patterns, we adopt Fama and MacBeth's (1973) two-stage estimation. In the first stage, for each day, we estimate the predictive regression

$$\begin{aligned} mroib\ (i,w) = b0\,(w) + b1(w)' ret\,(i,w-1) + b2(w)' controls\,(i,w-1) \\ + u1\,(i,w), \end{aligned} \tag{5}$$

where we use various horizons of past weekly returns, $ret(i,w-1)$, and various control variables from the past to explain the order imbalance measure, $mroib(i,w)$, for firm i during week w. The first-stage estimation generates a daily overlapping time series of weekly coefficients, $\{b0(w), b1(w)', b2(w)'\}$. In the second stage, we conduct statistical inference using the time series of the coefficients. Because we use overlapping daily-frequency data for weekly order imbalance and return measures, the standard errors are calculated using Newey-West (1987) with six lags.[10]

To explain the order imbalance over week w, from day 1 to day 5, we first include its own lag, the past week's order imbalance from day -4 to day 0, or $mroib(w-1)$. We also include past returns over three horizons: the previous week ($ret(w-1)$), the previous month ($ret(m-1)$), and the previous six months ($ret(m-7,m-2)$). For control variables, we use log market cap, log book-to-market ratio, turnover (share volume over shares outstanding), and daily return volatility, all computed using the previous month's data.

The results are presented in Table II, with regressions I and II explaining the order imbalance using shares, and regressions III and IV explaining the order

[10] The optimal lag number is chosen using Bayesian Information Criterion.

Table II

Determinants of Marketable Retail Order Imbalances

This table reports determinants of retail investor trading activity. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We estimate Fama-MacBeth (1973) regressions as specified in equation (5). The dependent variables are two scaled marketable retail order imbalance measures: *mroibvol* (based on the number of shares traded) and *mroibtrd* (based on the number of trades). As independent variables, we include the previous-week return, *ret*($w-1$), previous-month return, *ret*($m-1$), and previous six-month return, *ret*($m-7$, $m-2$). We compute weekly returns in two ways: using the end-of-day bid-ask average price or using the CRSP closing price. The control variables are monthly turnover (*lmto*), monthly volatility of daily returns (*lvol*), log market cap (*size*), and log book-to-market ratio (*lbm*), all measured at the end of the previous month. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey-West (1987) with six lags.

Regression Dependent Variables Return	I Mroibvol Bid-Ask Return		II Mroibvol CRSP Return		III Mroibtrd Bid-Ask Return		IV Mroibtrd CRSP Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat
Intercept	−0.4013	−20.03	−0.4065	−20.19	−0.4326	−22.00	−0.4357	−22.01
Mroib($w-1$)	0.2200	92.53	0.2201	92.57	0.2865	150.01	0.2866	150.06
Ret($w-1$)	−0.9481	−40.60	−0.9620	−41.43	−0.9003	−35.92	−0.9156	−36.74
Ret($m-1$)	−0.2778	−19.24	−0.2784	−19.30	−0.2258	−14.84	−0.2262	−14.87
Ret($m-7$,$m-2$)	−0.0586	−11.49	−0.0584	−11.46	−0.0380	−6.50	−0.0378	−6.48
Lmto	0.0003	5.31	0.0003	5.19	0.0002	3.93	0.0002	3.83
Lvol	0.8100	8.37	0.8478	8.79	0.4366	4.24	0.4633	4.51
Size	0.0154	12.06	0.0157	12.31	0.0209	16.37	0.0211	16.48
Lbm	−0.0275	−17.66	−0.0274	−17.61	−0.0274	−18.09	−0.0273	−18.05
Adj. R^2	6.00%		6.01%		9.49%		9.50%	

imbalance using the number of trades. In the first regression, the order imbalance using share volume, *mroibvol*, has a positive correlation with its own lag, with a highly significant coefficient of 0.22, indicating that directional marketable retail trading activity is somewhat persistent over successive weeks, as suggested by Chordia and Subrahmanyam (2004). The coefficients for the past one-week, past one-month, and past six-month returns are −0.9481, −0.2778, and −0.0586, respectively. All three coefficients are negative and highly significant, which suggests that marketable retail order flows are contrarian for horizons ranging between one week and six months. The control variables indicate that investors tend to buy more aggressively in larger firms, growth firms, and firms with higher turnover and higher volatility. All coefficients are highly significant. The average adjusted-R^2 from the first-stage cross-sectional estimation is about 6%.

We use different return and order imbalance measures for regressions II, III, and IV. At the weekly horizon, the results are similar across methods of computing returns and order imbalances. We focus our discussion on bid-ask midpoint returns, which do not have bid-ask bounce and thus exhibit a smaller degree of time-series predictability than returns based on transaction prices. We also include CRSP returns in the results for the sake of completeness and robustness.

The negative coefficients on past returns match some of the findings in the literature. For example, marketable retail order flows are found to be contrarian by Kaniel, Saar, and Titman (2008) over monthly horizons and by Barrot, Kaniel, and Sraer (2016) over daily and weekly horizons. In contrast, Kelley and Tetlock (2013) paint a more complex picture. They find that at weekly horizons, marketable retail order imbalance measures are contrarian and have negative coefficients on past returns, whereas over shorter (daily) horizons, market order imbalances actually have a positive coefficient on the lagged one-day return, which implies momentum rather than contrarian behavior.

Internet Appendix Figure IA.2, Panel B, plots the correlation between daily order imbalance and past returns for the previous 1 to 80 trading days. Similar to Kelley and Tetlock (2013), the correlation between the current marketable retail order imbalance and the previous-day return is positive, indicating a momentum trading pattern on average. However, at lags between 2 and 30 days, our average correlation is slightly negative. Our results are thus consistent with the findings of Kelley and Tetlock (2013) at short horizons and with those of other researchers at longer horizons.[11,12]

[11] Lee et al. (2004) also find a mixed pattern of contrarian and momentum trading, using the overall market order imbalance. They find that overall trades tend to follow a momentum pattern after up-market moves, while overall trades tend to be contrarian after downmarket moves. We provide similar results using daily retail order flows in Internet Appendix Table IA.I, Panel A. When we use weekly retail order flows, both patterns become contrarian, as shown in Table IA.I, Panel B.

[12] In addition, we examine how firm-level order imbalance measures are related to firm fundamentals, as in Chordia, Huh, and Subrahmanyam (2007). The results in Internet Appendix Table

Our results in Table II reveal two important factors affecting the weekly order imbalance. The first is its own lag, which indicates that the marketable retail order imbalance measures are persistent. The second is past returns, for which we find both contrarian and momentum patterns, with the contrarian pattern prevailing at weekly horizons.

B. Predicting Future Stock Returns with Marketable Retail Order Imbalance Measures

B.1. Methodology and Overall Predictive Power

Can marketable retail investors' activity provide useful information for future stock returns? In this section, we examine the predictive power of our order imbalance measures using Fama-MacBeth (1973) regressions. Specifically, we estimate the following specification,

$$ret(i, w) = c0\,(w) + c1\,(w)\,mroib\,(i, w-1) + c2(w)'controls\,(i, w-1) + u2\,(i, w)\,, \tag{6}$$

where we use the marketable retail order imbalance measure from the previous week, $mroib(i, w-1)$, and various control variables to predict the next week's stock return, $ret(i, w)$, for firm i during week w. As in the previous section, because we use overlapping daily-frequency data for weekly order imbalances and return measures, the standard errors of the time series are adjusted using Newey-West (1987) with five lags. If past marketable retail order imbalances predict the cross-section of future returns in the same direction, the coefficient $c1$ should be significantly positive. For example, if retail buys dominate retail sells for a particular stock during a particular week, a positive $c1$ would mean that the stock's future return tends to be above the cross-sectional average. Such a pattern could have several possible explanations, none of which is mutually exclusive: there could be persistence in marketable retail order imbalances, marketable retail orders could be compensated for providing liquidity, or retail traders may have valuable information that is incorporated into stock prices at some point after they trade. We examine these hypotheses in Section II.C. If coefficient $c1$ turns out to be significantly negative, then there could be several possible explanations: these retail investors may be making systematic trading mistakes, or retail investors may be mainly "liquidity" or "noise" traders who end up trading at temporarily disadvantageous prices because rational but risk-averse market makers require compensation for trading with them. Either way, a negative $c1$ would constitute a drag on the overall returns of these retail investors. Finally, if $c1$ is insignificantly different from zero, we cannot reject the null that our measure of marketable retail order flow is uninformative on average about the cross-section of future stock returns.

IA.I, Panel E, indicate that retail order imbalances are positively related to firm size, number of analysts, analyst dispersion, and leverage and negatively related to past return, firm age, and book-to-market ratio.

We again include past returns as control variables, using three different horizons: the previous week, the previous month, and the previous six months (month $m-7$ to month $m-2$). In addition, we include log market cap, log book-to-market ratio, turnover, and daily return volatility, all from the previous month. We report the estimation results in Table III. In regression I, we use the order imbalance based on share volume, *mroibvol*, to predict the next week's return based on bid-ask midpoints. The coefficient on *mroibvol* is 0.0009, with a t-statistic of 15.60. The positive and significant coefficient indicates that if retail investors buy more than they sell in a given week, the return on that stock in the next week is significantly higher. In terms of magnitude, we report at the bottom of the table that the interquartile range for *mroibvol* is 1.1888 per week. Multiplying the interquartile difference by the regression coefficient of 0.0009 generates a weekly return difference of 10.89 bps (or 5.66% per year) when moving from the 25^{th} to the 75^{th} percentile of *mroibvol*. The same pattern is present when we use different order imbalance and return measures, and the weekly interquartile difference in the conditional mean return ranges from 9.31 to 11.44 bps (4.84% to 5.94% per year). Whether economic magnitudes are large or small is open to interpretation, but this strikes us as a nontrivial amount of cross-sectional predictability that lasts for a relatively long time (weeks, not days, as we show later in the paper). Overall, past-week marketable retail order imbalances can significantly predict future returns in the correct direction.

Turing to the control variables, we observe negative coefficients on the previous week's return, which indicates weekly return reversals, and positive coefficients on the other longer horizon returns, which indicates momentum. Size, book-to-market, turnover, and volatility all carry the expected signs, and most are not statistically significant. This result also confirms that the predictability we find is not simply a manifestation of some other size, book-to-market, turnover, or volatility anomaly. The average adjusted R^2s from the first-stage cross-sectional estimation are mostly around 3.85%.

B.2. Subgroups in the Cross-Section

Our sample includes on average more than 3,000 firms each day. Is the predictive power of marketable retail order imbalances restricted to a particular type of firm? Do informed retail investors have preferences for particular types of firms? In this section, we investigate these questions by analyzing various firm subgroups. We first sort all firms into three groups based on a firm or stock characteristic observed at the end of the previous month. We then estimate equation (6) within each characteristic group. That is, we allow all coefficients in equation (6) to be different within each group, which allows substantial flexibility in the possible predictive relationship across these different groups.

To save space, we only include the results on weekly returns that are computed using the end-of-day bid-ask average price. We first sort all stocks into three size groups based on market capitalization: small, medium, and large. The results are reported in Panel A of Table IV. In the left panel, we report

Table III

Predicting Next-Week Returns Using Marketable Retail Order Imbalances

This table reports estimation results on whether retail investors' trading activity can predict the cross-section of one-week-ahead returns. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We estimate Fama-MacBeth (1973) regressions as specified in equation (6). The dependent variable is weekly individual stock returns, computed in two ways: using the end-of-day bid-ask average price or using the CRSP closing price. The independent variables are two scaled marketable retail order imbalance measures: *mroibvol* (based on the number of shares traded) and *mroibtrd* (based on the number of trades). As independent variables, we include the previous-week return, *ret*($w-1$), previous-month return, *ret*($m-1$), and previous six-month return, *ret*($m-7$, $m-2$). The control variables are log book-to-market ratio (*lbm*), log market cap (*size*), monthly turnover (*lmto*), and monthly volatility of daily returns (*lvol*), all measured at the end of the previous month. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags.

Regression Order Imbalance Dependent Variables	I Mroibvol Bid-Ask Return		II Mroibvol CRSP Return		III Mroibtrd Bid-Ask Return		IV Mroibtrd CRSP Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat
Intercept	0.0050	2.58	0.0056	2.85	0.0050	2.58	0.0056	2.85
Mroib($w-1$)	0.0009	15.60	0.0010	16.29	0.0008	12.30	0.0008	13.20
Ret ($w-1$)	−0.0185	−5.83	−0.0220	−6.85	−0.0186	−5.88	−0.0222	−6.91
Ret ($m-1$)	0.0006	0.35	0.0006	0.34	0.0005	0.29	0.0005	0.29
Ret ($m-7$, $m-2$)	0.0008	1.16	0.0008	1.16	0.0008	1.12	0.0008	1.12
Lmto	0.0000	−3.37	0.0000	−3.76	0.0000	−3.36	0.0000	−3.75
Lvol	−0.0223	−1.41	−0.0205	−1.31	−0.0217	−1.37	−0.0198	−1.27
Size	−0.0001	−0.86	−0.0001	−0.92	−0.0001	−0.90	−0.0001	−0.96
Lbm	−0.0001	−0.39	0.0000	−0.07	−0.0001	−0.42	0.0000	−0.10
Adj. R^2	3.85%		3.85%		3.84%		3.84%	
Interquartile	1.1888		1.1888		1.2292		1.2292	
Interquartile weekly return diff	0.1089%		0.1144%		0.0931%		0.0997%	

Table IV

Marketable Retail Return Predictability within Subgroups

This table reports whether marketable retail investor order imbalances can predict the cross-section of returns for subsets of stocks. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We first sort all firms into three groups based on previous month-end characteristics. We then estimate Fama-MacBeth (1973) regressions, specified in equation (6), for each subgroup. The dependent variable is weekly returns on individual stocks computed using the end-of-day bid-ask average price. The independent variables are two scaled marketable retail order imbalance measures: *mroibvol* (based on number of shares traded) and *mroibtrd* (based on number of trades). To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey-West (1987) with five lags. For each regression, we also provide the interquartile range for the relevant explanatory order imbalance along with the difference in predicted week-ahead returns for observations at the two ends of the interquartile range. Control variables are the same as in Table III and are not reported.

Panel A: Market Cap Groups

Mroib Measure	Mroibvol				Mroibtrd			
Mkt Cap	Coef.	t-Stat	Interquartile	Weekly Return Diff	Coef.	t-Stat	Interquartile	Weekly Return Diff
Small	0.0013	13.90	1.662	0.219%	0.0012	11.58	1.736	0.207%
Medium	0.0007	9.18	1.323	0.087%	0.0004	5.63	1.346	0.059%
Big	0.0003	3.68	0.892	0.026%	0.0002	2.52	0.929	0.019%

Panel B: Share Price Groups

Mroib Measure	Mroibvol				Mroibtrd			
Price Groups	Coef.	t-Stat	Interquartile	Weekly Return Diff	Coef.	t-Stat	Interquartile	Weekly Return Diff
Low	0.0014	13.34	1.432	0.205%	0.0012	10.34	1.586	0.185%
Medium	0.0007	10.00	1.289	0.089%	0.0005	7.56	1.309	0.070%
High	0.0002	3.23	0.961	0.020%	0.0002	2.19	0.961	0.015%

Panel C: Turnover Groups

Mroib Measure	Mroibvol				Mroibtrd			
Turnover Groups	Coef.	t-Stat	Interquartile	Weekly Return Diff	Coef.	t-Stat	Interquartile	Weekly Return Diff
Low	0.0011	15.60	1.837	0.205%	0.0011	14.71	1.777	0.195%
Medium	0.0008	10.21	1.219	0.094%	0.0006	7.05	1.228	0.071%
High	0.0007	4.98	0.910	0.065%	0.0004	2.55	1.005	0.037%

coefficients on *mroibvol*, the order imbalance computed from share volume. When we move from the smallest one-third of firms by market cap to the largest tercile, the coefficient on *mroibvol* decreases from 0.0013 to 0.0003, and the t-statistic decreases from 13.90 to 3.68. Clearly, the predictive power of marketable retail order imbalances is much stronger for smaller firms than for larger-cap firms, but the predictability remains reliably present in all three groups. Economically, the interquartile difference in weekly returns is 21.9 bps for the smallest firms (11.39% per year) and 2.6 bps for the largest firms (1.35% per year). The results in the right panel using order imbalance based on the number of trades (*mroibtrd*) are quite similar.

In Panel B of Table IV, we sort all firms into three groups based on the previous month-end share price. In the left panel, moving from the lowest share-price firms to the highest, the coefficient on *mroibvol* decreases from 0.0014 to 0.0002, and the t-statistic decreases from 13.34 to 3.23. In terms of magnitude, the interquartile weekly return difference is 20.5 bps (10.66% per year) for the lowest price firms and only 2.0 bps for the firms with the highest share price (1.04% per year). The results are similar for specifications using *mroibtrd*, reported in the right panel, with slightly lower coefficients and t-statistics. The pattern is clear: the predictive power of marketable retail order imbalances for future returns is stronger for low-price firms.

Next, we sort all firms based on previous-month turnover, which may be a proxy for liquidity. In the left panel, moving from the tercile of low trading activity to firms with more turnover, the coefficient on *mroibvol* decreases from 0.0011 to 0.0007, and the t-statistic decreases from 15.60 to 4.98. In terms of magnitude, the interquartile weekly return difference is 20.5 bps (10.66% per year) for the firms with the lowest turnover and 6.5 bps for the firms with the highest turnover (3.38% per year). For specifications based on *mroibtrd* in the right panel, the results are similar, with slightly lower coefficients and t-statistics. Overall, marketable retail order imbalances better predict returns for firms with lower trading activity.

In this section, we find that the predictive power of the marketable retail order imbalance is significant and positive for all but one subgroup, which shows that the predictive power is not driven by special subgroups. However, a clear cross-sectional pattern is observed for the predictive power. In particular, the predictive power of the marketable retail order imbalance is much stronger for small firms and firms with low share price and low liquidity.

B.3. Longer Horizons

The results in the previous section show that marketable retail order imbalances can predict next week's returns positively and significantly. It is thus natural to ask whether the predictive power is transient or persistent. If the predictive power quickly reverses, the retail investors may be capturing price reversals; if the predictive power continues over time and then vanishes beyond some horizon, the retail investors may be informed about information related to firm fundamentals. To address this question, we extend equation (6) to longer horizons as follows:

$$ret(i, w+k) = c0\,(w) + c1\,(w)\,mroib\,(i, w) + c2(w)'controls\,(i, w) + u3\left(i, w+k\right). \tag{7}$$

That is, we use one week of order imbalance measures to predict k-week ahead returns, $ret(i,w+k)$, where $k = 1$ to 12. To observe the decay of the predictive power of marketable retail order imbalance, the return to be predicted is a weekly return over a one-week period, rather than a cumulative return over n weeks, which is an average over all weeks involved. If marketable retail order imbalances have only short-lived predictive power for future returns, the coefficient $c1$ should decrease to zero within a couple of weeks. Alternatively, if the marketable retail order imbalance has longer predictive power, the coefficient $c1$ should remain statistically significant for a longer period. In our empirical estimation, we choose k ranging from 2 to 12 weeks.

We report the results in Table V, with the results based on bid-ask average returns in Panel A and those based on closing transaction prices in Panel B. In Panel A, when we extend the window from 2 to 12 weeks, the coefficient on *mroibvol* decreases monotonically from 0.00055 to 0.00007, and the coefficient on *mroibtrd* decreases from 0.00048 to 0.00006. The coefficients are statistically significant up to six or eight weeks ahead. The results in Panel B are similar with no evidence of price reversals at any horizon. Thus, our marketable retail order imbalances potentially capture either longer lived information or slow information diffusion.

B.4. Long-Short Portfolios

One might wonder whether we can use marketable retail order imbalances as a signal to form a profitable trading strategy. As discussed earlier, both *mroibvol* and *mroibtrd* are publicly available information. In this section, we form quintile portfolios based on the previous week's average order imbalance and then hold the quintile portfolios for up to 12 weeks. If retail investors on average can select the right stocks to buy and sell, then firms with higher or positive marketable retail order imbalance should outperform firms with lower or negative order imbalance. Notice that this exercise uses marketable retail order imbalance measures merely as a signal to predict future stock returns, and thus, it provides no information on whether retail investors with marketable orders profit from their own trades. We ignore trade frictions and transaction costs here, and thus the results do not have implications for whether outsiders can profit from these signals.

Table VI reports long-short portfolio returns, where we buy the stocks in the highest order imbalance quintile and short the stocks in the lowest order imbalance quintile each day using the previous five-day marketable retail order flow measures and then hold them for the next few weeks. Portfolio returns are value-weighted using the previous month-end market cap. Because the holding period can be as long as 12 weeks, we report both raw and risk-adjusted

Table V

Predicting Returns k Weeks Ahead

This table reports estimation results on whether marketable retail investor trading activity can predict the cross-section of stock returns at more distant horizons. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We estimate Fama-MacBeth (1973) regressions, as specified in equation (7). The dependent variable is the weekly individual stock return k-weeks ahead, computed in two ways: using the end-of-day bid-ask average price (Panel A) or using the CRSP closing price (Panel B). The independent variables are two scaled marketable retail order imbalance measures: *mroibvol* (based on the number of shares traded) or *mroibtrd* (based on the number of trades). To account for serial correlation in the coefficients, the standard deviations of the time series are adjusted using Newey-West (1987) with five lags. Control variables are the same as in Table III and are not reported.

Panel A: Predict Bid-Ask Average Return k Weeks Ahead

	Mroibvol		Mroibtrd	
# of Weeks Ahead	Coef.	t-Stat	Coef.	t-Stat
1 week	0.00092	15.60	0.00076	12.30
2 weeks	0.00055	9.35	0.00048	7.89
4 weeks	0.00031	5.56	0.00026	4.66
6 weeks	0.00022	3.90	0.00015	2.60
8 weeks	0.00021	3.47	0.00011	1.75
10 weeks	0.00010	1.82	0.00002	0.35
12 weeks	0.00007	1.29	0.00009	1.52

Panel B: Predict CRSP Return k Weeks Ahead

	Mroibvol		Mroibtrd	
# of Weeks Ahead	Coef.	t-Stat	Coef.	t-Stat
1 week	0.00096	16.29	0.00081	13.20
2 weeks	0.00058	9.99	0.00052	8.57
4 weeks	0.00032	5.92	0.00028	5.05
6 weeks	0.00024	4.18	0.00017	2.93
8 weeks	0.00021	3.50	0.00011	1.80
10 weeks	0.00011	2.04	0.00005	0.81
12 weeks	0.00008	1.39	0.00010	1.76

returns using the Fama and French (1993) three-factor model. Given the use of overlapping data, we adjust the standard errors of the portfolio return time series using Hansen and Hodrick (1980) standard errors with the corresponding number of lags.[13]

In Panel A, the long-short strategy is based on the previous week's *mroibvol*, and we report bid-ask average returns. Over a one-week horizon, the long-short portfolio return is 0.092%, or 4.78% per year annualized. The t-statistic is 2.66.

[13] For example, for a portfolio with a one-week holding period, we use Hansen and Hodrick (1980) with five lags. For a two-week holding period, we use Hansen and Hodrick (1980) with 10 lags, and so on.

Table VI

Long-Short Strategy Returns Based on Marketable Retail Order Imbalances

This table reports portfolio returns using a long-short strategy where we buy the stocks in the highest quintile of scaled marketable retail order imbalance and short the stocks in the lowest scaled marketable retail order imbalance quintile. The order imbalance is computed during the previous week. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. Portfolio returns are value-weighted, and market cap terciles are based on the previous month-end market cap. Because the holding period can be as long as 12 weeks, we report both raw returns and risk-adjusted returns using the Fama and French (1993) three-factor model. As our data are overlapping, we adjust the standard errors of the portfolio return time series using Hansen-Hodrick (1980) standard errors with the corresponding number of overlapping lags.

Panel A: Form Portfolios on the Previous Week's Marketable Retail Order Imbalance Based on Number of Shares Traded

Holding	Full Sample				Small		Medium		Big	
Period	Mean	*t*-Stat	alpha	*t*-Stat	alpha	*t*-Stat	alpha	*t*-Stat	alpha	*t*-Stat
1 week	0.092%	2.66	0.084%	2.43	0.403%	9.16	0.170%	6.24	0.067%	1.78
2 weeks	0.147%	2.45	0.135%	2.46	0.669%	9.01	0.292%	6.81	0.105%	1.70
4 weeks	0.223%	1.89	0.208%	2.00	1.124%	10.43	0.423%	6.36	0.143%	1.22
6 weeks	0.310%	1.72	0.277%	1.73	1.399%	13.02	0.558%	6.07	0.171%	1.05
8 weeks	0.448%	1.92	0.460%	2.26	1.709%	17.13	0.623%	4.18	0.342%	1.69
10 weeks	0.515%	1.99	0.484%	1.81	1.704%	11.17	0.578%	3.87	0.381%	1.53
12 weeks	0.588%	2.09	0.629%	1.89	1.857%	7.65	0.556%	3.20	0.477%	1.48

Panel B: Form Portfolios on the Previous Week's Marketable Retail Order Imbalance Based on the Number of Trades

Holding	Full Sample				Small		Medium		Big	
Period	Mean	*t*-Stat	alpha	*t*-Stat	alpha	*t*-Stat	alpha	*t*-Stat	alpha	*t*-Stat
1 week	0.056%	1.34	0.061%	1.44	0.343%	7.04	0.104%	3.52	0.055%	1.42
2 weeks	0.137%	1.72	0.143%	1.89	0.557%	6.72	0.194%	4.02	0.119%	1.61
4 weeks	0.238%	1.61	0.251%	1.88	0.880%	6.98	0.277%	3.75	0.214%	1.61
6 weeks	0.311%	1.50	0.350%	1.93	1.145%	6.25	0.313%	2.62	0.304%	1.84
8 weeks	0.427%	1.58	0.523%	2.26	1.468%	6.40	0.353%	1.91	0.449%	2.19
10 weeks	0.454%	1.41	0.539%	1.74	1.442%	5.37	0.292%	1.56	0.483%	1.64
12 weeks	0.529%	1.47	0.667%	1.70	1.672%	5.30	0.228%	1.05	0.567%	1.51

Risk adjustment using the Fama and French (1993) three-factor model does not make much difference: the weekly Fama-French alpha for the long-short portfolio is 0.084%, with a t-statistic of 2.43. When we increase the holding horizon to 12 weeks, the mean return becomes 0.588%, with a t-statistic of 2.09. The general pattern is that holding-period returns (and alphas) continue to grow at a decreasing rate over time. We observe no evidence of a reversal in returns. In terms of statistical significance, the t-statistics are significant or marginally significant up to the 12-week horizon. These results are slightly weaker than those of the Fama-MacBeth (1973) regressions, mainly because in this section, we value-weight the portfolio returns across firms, while the Fama-MacBeth (1973) approach implicitly weights each stock equally.

When we restrict portfolio formation to one of the three market cap groups, the one-week return is 0.403% (or 20.96% per year) with a t-statistic of 9.16 for the smallest firms, while the one-week return is 0.067% (or 3.48% per year) with a t-statistic of 1.78 for the largest firms. When the holding horizon becomes longer, the return on the long-short strategy is still significant and positive for up to 12 weeks for the smallest third of firms, but the results are statistically insignificant for the largest tercile. The results in Panel B, obtained using *mroibtrd*, are qualitatively similar but with smaller magnitudes and lower statistical significance. This result is expected since, as mentioned, the information provided by *mroibvol* is similar but finer than that provided by *mroibtrd*.[14]

To ensure that the statistical significance in return differences is not driven by particular sample periods, in Figure 3, we provide a time-series plot of the return differences between quintiles 1 and 5 for the portfolios sorted on *mroibvol* and a holding period of one week. Over our six-year sample period, we observe both time variation in the return differences and positive and negative spikes. However, most data points are positive, and the positive returns are not driven by particular sample subperiods.

C. Alternative Hypotheses for Marketable Retail Order Imbalance Predictive Power for Future Returns

The predictive power of marketable retail order imbalances for future stock returns is consistent with three hypotheses. First, as in Chordia and Subrahmanyam (2004), order flows are persistent, and, since the retail buying/selling pressure is also persistent, this could lead directly to the predictability of

[14] We also conduct a rough calculation that includes transaction costs. Frazzini, Israel, and Moskowitz (2018) state that a reasonable estimate of the one-way transaction cost on value-weighted U.S. stocks is about 12 bps for the period January 2006 to June 2016. To be conservative, we assume that for each rebalance, we change 100% of the positions. That is, each rebalance incurs a 2*12 bps = 24 bps rebalance cost. For instance, for a weekly rebalance or one week holding period, each year's transaction cost would be 52 rebalances * 2 * 12bps = 1,248 bps. After this drastic transaction cost adjustment, the mean returns and alphas remain positive and significant for small firms over all holding horizons. For medium and big firms, the mean returns and alphas stay positive for longer holding periods but are mostly insignificant.

Panel A. Weekly portfolio return difference using the end-of-day bid-ask average prices

Panel B. Weekly portfolio return difference using the CRSP closing prices

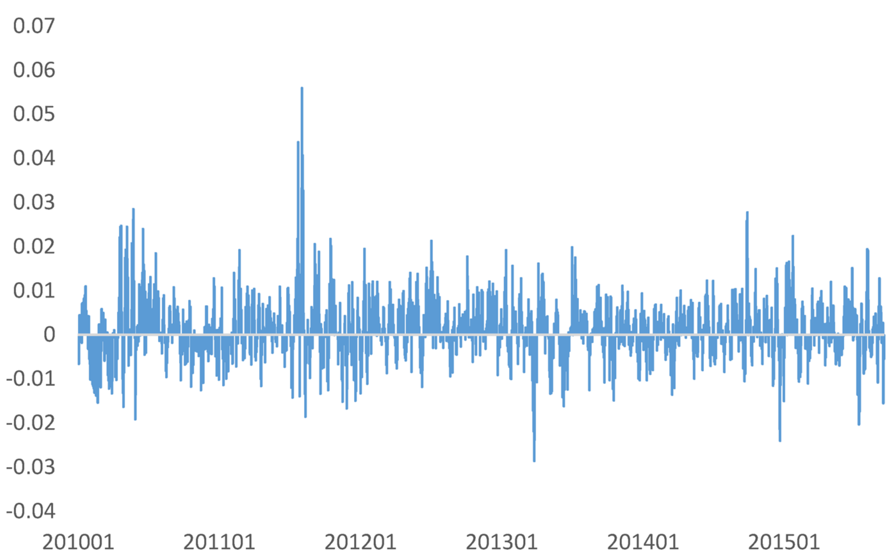


Figure 3. Portfolio return difference using previous week's marketable retail order imbalance. These figures plot weekly value-weighted portfolio return differences between quintile 5 and quintile 1, where stocks are sorted on the previous week's marketable retail order imbalance calculated using the number of shares traded (*mroibvol*). Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. Portfolio returns are computed using the end-of-day bid-ask average price (*bidaskret*) in Panel A and the CRSP closing price (*crspret*) in Panel B.
(Color figure can be viewed at wileyonlinelibrary.com)

future returns. Second, as in Kaniel, Saar, and Titman (2008), these retail traders are contrarian at weekly horizons, and since their contrarian trading provides liquidity to the market, their trades might positively predict future returns. Third, as in Kelley and Tetlock (2013), retail investors, especially aggressive investors using market orders, may have valuable information about the firm, and thus, their trading could correctly predict the direction of future returns. The above three hypotheses are not mutually exclusive. In Section II.C.1, we conduct a simple decomposition to separate alternative hypotheses. In Section II.C.2, we provide more evidence on the liquidity provision hypothesis.

C.1. Two-Stage Decomposition

To distinguish among the alternative hypotheses for the predictive relation between previous-period marketable retail order imbalance and next-period stock returns as in equation (6), we adopt a two-step decomposition. In the first step, we decompose the independent variable in equation (6), the previous week's marketable retail order imbalance *mroib(w−1)*, into three components using cross-sectional regressions for each week $w-1$:

$$\begin{aligned} mroib\ (i, w-1) = {} & d0\,(w-1) + d1\,(w-1)\, mroib\,(i, w-2) \\ & + d2(w-1)' ret\,(i, w-2) + u4\,(i, w-1)\,. \end{aligned} \tag{8}$$

For each week $w-1$, we obtain the time series of coefficients, $\{\widehat{d0}(w-1), \widehat{d1}(w-1), \widehat{d2}(w-1)'\}$. Next, we calculate the three components of *mroib(i,w-1)* as follows:

$$\begin{aligned} \widehat{mroib}_{i,w-1}^{persistence} &= \widehat{d1}\ (w-1)\, mroib\,(i, w-2)\,, \\ \widehat{mroib}_{i,w-1}^{contrarian} &= \widehat{d2}\ (w-1)' ret\ (i, w-2)\,, \\ \widehat{mroib}_{i,w-1}^{other} &= \widehat{u4}\ (i, w-1) + \widehat{d0}\,(w-1)\,. \end{aligned} \tag{9}$$

From equations (8) and (9), we know that

$$mroib\ (i, w-1) = \widehat{mroib}_{i,w-1}^{persistence} + \widehat{mroib}_{i,w-1}^{contrarian} + \widehat{mroib}_{i,w-1}^{other}. \tag{10}$$

We denote the part of $mroib(i, w-1)$ related to the past order imbalance as "persistence," which is related to the price pressure hypothesis. The part related to past returns over different horizons is labeled "contrarian," which relates to the liquidity provision hypothesis. After accounting for predictability due to "persistence" and "contrarian" trading, we are left with "other," which potentially contains other relevant information about future returns. Note that this empirical decomposition is an identity: when we add up these three components, by definition, we obtain the explanatory variable $mroib(i, w-1)$ in our basic predictive regression in equation (6).

In the second stage, we replace $mroib(i, w-1)$ in equation (6) by its three components and we estimate the following regression using the Fama-MaBeth (1973) methodology:

$$ret(i,w) = e0\,(w) + e1\,(w)\,\widehat{mroib}^{\,persistence}_{i,w-1} + e2\,(w)\,\widehat{mroib}^{\,contrarian}_{i,w-1} + e3\,(w)\,\widehat{mroib}^{\,other}_{i,w-1} + e4(w)'controls\,(i, w-1) + u5\,(i,w). \quad (11)$$

Since we decompose the original order imbalance measure $mroib(i, w-1)$ into three parts, related to order flow persistence, a contrarian trading pattern, and the residual, the coefficient estimates in equation (11) reveal the contribution of each component of $mroib(i, w-1)$ to future stock returns. The advantage of the two-stage decomposition is that it includes components of $mroib(i, w-1)$ from alternative hypotheses in a unified and internally consistent empirical framework. The disadvantage of this approach is that without a structural model, interpreting the results may be more difficult. In particular, we must make empirical assumptions on proxies for the persistence and contrarian components. These empirical assumptions appear to be reasonable, but we nevertheless caution that interpretation of the results depends on the validity of our empirical assumptions.

We report the decomposition results in Table VII. Panel A presents the first-stage estimation as in equation (8), which is quite similar to those reported in Table II. Take the first regression as an example. The order imbalance measure, *mroibvol*, has a highly significant and positive coefficient on its own lag at 0.22, which indicates order persistence. In terms of past returns, the coefficients for the past week, past month, and past six-month returns are −0.9286, −0.2029, and −0.0267, respectively, all implying contrarian trading patterns.

After we decompose the previous week's order imbalance into "persistence," "contrarian," and "other," we include them together to predict future stock returns, as in equation (11). In the first regression, we use the past week's *mroibvol* to predict future bid-ask returns. The coefficient estimate on *mroib(persistence)* is 0.0027, with a t-statistic of 8.75, which implies that price pressure significantly and positively contributes to the predictive power of marketable retail flow. The coefficient estimate on *mroib(contrarian)* is −0.0044 and insignificantly different from zero, implying that we cannot reject the null hypothesis that the contrarian component does not contribute significantly to the predictive power of marketable retail order imbalances. Finally, for the *mroib(other)* component, the coefficient is 0.0008, with a strongly significant t-statistic of 14.47.[15]

[15] We also try to include the past order imbalance as a control variable for the second-stage estimation. We cannot directly include $mroib(w-1)$ or $mroib(w-2)$, because doing so would lead to collinearity issues. We therefore control past *mroib* using either $mroib(w-3)$ or $mroib(m-1)$.

Table VII

Predictability Decomposition

This table reports estimation results on a decomposition of the predictive power of marketable retail order flow for the cross-section of future stock returns. Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We estimate two-stage Fama-MacBeth (1973) regressions. Panel A reports the first-stage estimation results, where the order imbalance measures are decomposed into three components as specified in equation (8). Panel B reports the second-stage decomposition of order imbalance's predictive power, as specified in equations (9) to (11). The weekly returns are computed in two ways: using the end-of-day bid-ask average price or using the CRSP closing price. The scaled marketable retail order imbalance measures are *mroibvol* (based on the number of shares traded) and *mroibtrd* (based on the number of trades). The variable *mroib*($w-1$, *persistence*) is estimated in the first stage using past order imbalance and reflects price pressure. The variable *mroib*($w-1$, *contrarian*) is estimated in the first stage using past returns over different horizons and is connected to the liquidity provision hypothesis. The residual part of the previous week's order imbalance from the first-stage estimation is denoted as "other," which can be attributed to private information about future returns on the part of these marketable retail investors. As additional control variables, we include previous-week return, *ret*($w-1$), previous-month return, *ret*($m-1$), and previous six-month return, *ret*($m-7$, $m-2$). The control variables are log book-to-market ratio (*lbm*), log market cap (*size*), monthly turnover (*lmto*), and monthly volatility of daily returns (*lvol*), all measured at the end of the previous month. To account for serial correlation in the coefficients, the standard errors of the time series are adjusted using Newey-West (1987) with five lags.

Panel A: First Stage of Projecting Order Imbalance on Persistence and Past Returns

Regression	I		II		III		IV	
Dependent Variables	Mroibvol(w-1)		Mroibvol(w-1)		Mroibtrd(w-1)		Mroibtrd(w-1)	
Return	Bid-Ask Return		CRSP Return		Bid-Ask Return		CRSP Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat
Intercept	−0.1413	−24.66	−0.1408	−24.61	−0.1054	−17.23	−0.1049	−17.19
Mroib($w-2$)	0.2227	96.20	0.2228	96.20	0.2906	149.82	0.2907	149.85
Ret($w-2$)	−0.9286	−38.93	−0.9422	−39.80	−0.8926	−34.92	−0.9076	−35.81
Ret($m-1$)	−0.2029	−13.93	−0.2025	−13.90	−0.1591	−10.72	−0.1588	−10.70
Ret($m-7$,$m-2$)	−0.0267	−4.98	−0.0268	−4.99	−0.0054	−0.86	−0.0055	−0.88
Adj. R^2	5.62%		5.63%		8.99%		9.00%	

(*Continued*)

Table VII—*Continued*

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

Regression Order Imbalance Dependent Variables	I Mroibvol(w) Bid-Ask Return		II Mroibvol(w) CRSP Return		III Mroibtrd(w) Bid-Ask Return		IV Mroibtrd(w) CRSP Return	
	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat
Intercept	0.0046	2.25	0.0052	2.54	0.0046	2.23	0.0052	2.52
Mroib($w-1$,persistence)	0.0027	8.75	0.0029	9.41	0.0018	7.80	0.0019	8.56
Mroib($w-1$,contrarian)	−0.0044	−0.42	−0.1310	−1.46	−0.0073	−0.73	0.0328	1.62
Mroib($w-1$,other)	0.0008	14.47	0.0009	15.48	0.0006	10.51	0.0007	11.64
Ret($w-1$)	−0.0176	−5.41	−0.0206	−6.27	−0.0177	−5.45	−0.0207	−6.30
Ret($m-1$)	−0.0060	−0.67	0.0002	0.03	0.0017	0.56	0.0093	1.13
Ret($m-7$,$m-2$)	−0.0009	−0.65	−0.0127	−1.12	0.0017	0.95	−0.0008	−0.34
Lmto	0.0000	−3.49	0.0000	−3.80	0.0000	−3.48	0.0000	−3.78
Lvol	−0.0230	−1.48	−0.0231	−1.50	−0.0224	−1.44	−0.0225	−1.46
Size	−0.0001	−0.61	−0.0001	−0.67	−0.0001	−0.65	−0.0001	−0.72
Lbm	−0.0001	−0.46	0.0000	−0.14	−0.0001	−0.56	−0.0001	−0.23
Adj. R^2	4.26%		4.27%		4.25%		4.26%	
	Int'quartile range	Return diff	Int'quartile range	Return diff	Int'quartile range	Return diff	Int'quartile range	Return diff
Mroib($w-1$,persistence)	0.2591	0.0688%	0.2593	0.0739%	0.3498	0.0620%	0.3500	0.0679%
Mroib($w-1$,contrarian)	0.0627	−0.0277%	0.0631	−0.8265%	0.0614	−0.0445%	0.0619	0.2031%
Mroib($w-1$,other)	1.1141	0.0915%	1.1141	0.0977%	1.1326	0.0718%	1.1327	0.0792%

In terms of economic magnitude, we compute the interquartile range of all three components of the order imbalance measure. For *mroib(persistence)*, if we move from the 25th percentile firm to the 75th percentile firm, the difference in future one-week stock return is 0.0688% (3.58% per year). For *mroib(other)*, if we move from the 25th percentile firm to the 75th percentile firm, the difference in future one-week stock return is 0.0915% (4.76% per year). For *mroib(contrarian)*, the sign is the opposite and has no statistical significance. The results in the other specifications are quite similar.[16]

The results are presented in Internet Appendix Table IA.II. We also use the contemporaneous return rather than lag returns in equation (8), and the results are reported in Internet Appendix Table IA.III. No matter which specification we use, the main results are quite similar to those in Table VII.

[16] All three components of the marketable order imbalance measure in equation (11) are measured week $w-1$, so we are using this identity to decompose the *predictive* relation between *mroib(w−1)* and *ret(w)* in equation (6). If we are willing to depart from this predictive decomposition framework, we can examine other relationships. For example, we might want to examine the *contemporaneous* relation between *mroib(w)* and *ret(w)*, and use the three fitted values from week w rather than from week $w-1$. In this case, the first-stage estimation for week w becomes

$$mroib(i,w) = d0(w) + d1(w)mroib(i,w-1) + d2(w)'ret(i,w-1) + u4(i,w). \qquad (8')$$

These regressions are period-by-period cross-sectional regressions, which means that when we estimate coefficients $\{\widehat{d0}(w), \widehat{d1}(w), \widehat{d2}(w)'\}$ in equation (8′), they are estimated using information from both week w and week $w-1$. From this first-stage estimation, we can define the relevant persistence estimate as

$$\begin{aligned} \widehat{mroib}_{i,w}^{persistence} &= \widehat{d1}(w)mroib(i,w-1), \\ \widehat{mroib}_{i,w}^{contrarian} &= \widehat{d2}(w)'ret(i,w-1), \qquad (9') \\ \widehat{mroib}_{i,w}^{other} &= \widehat{u4}(i,w) + \widehat{d0}(w). \\ mroib(i,w) &= \widehat{mroib}_{i,w}^{persistence} + \widehat{mroib}_{i,w}^{contrarian} + \widehat{mroib}_{i,w}^{other}. \qquad (10') \end{aligned}$$

Notice that coefficients $\{\widehat{d0}(w), \widehat{d1}(w), \widehat{d2}(w)'\}$ are estimated using information from week w, and thus $\widehat{mroib}_{i,w}^{persistence}$, $\widehat{mroib}_{i,w}^{persistence}$, and $\widehat{mroib}_{i,w}^{persistence}$ all use information from week w. The second-stage estimation for the contemporaneous relation between returns and the marketable retail order imbalance then becomes

$$\begin{aligned} ret(i,w) = e0(w) &+ e1(w)\widehat{mroib}_{i,w}^{persistence} \\ &+ e2(w)\widehat{mroib}_{i,w}^{contrarian} + e3(w)\widehat{mroib}_{i,w}^{other} \qquad (11') \\ &+ e4(w)'controls(i,w-1) + u5(i,w). \end{aligned}$$

Compared with equation (11), equation (11′) gives us something closer to an estimate of the contemporaneous relation between the components of $mroib(i,w)$ and $ret(i,w)$, rather than a *predictive* relation between $mroib(i,w-1)$ and $ret(i,w)$. We present the estimation results in Internet Appendix Table IA.IV. Table IA.IV, Panel A, reports the estimation results of equation (8′), and the results are quite similar to those in Table II and Panel A of Table VII. Table IA.IV, Panel B, shows the estimation results for equation (11′). Table IA.IV, Panel B, shows the estimation results for equation (11′). For instance, in regression I, the coefficient on *mroib(w, persistence)* is 0.0045 with a t-statistic of 14.26, while the coefficient on *mroib(w−1, persistence)* in Table VII, Panel B, is 0.0027 with a t-statistic of 8.75. That is, the coefficient on past order persistence becomes larger and more significant in equation (11′) than in equation (11), indicating that contemporaneous price

Our decomposition exercise shows that close to half of the predictive power of the marketable retail order, imbalance comes from the persistence of the order imbalance measures,[17] with most of the rest coming from the residual component, after we take out order persistence and the contrarian trading pattern. Since this residual component significantly predicts future stock returns, it is consistent with the hypothesis that marketable retail investor trading contains valuable information about future stock price movements.

C.2. A Closer Look at the Liquidity Provision Hypothesis

The liquidity provision hypothesis receives substantial attention in existing literature, so here we take a closer look at this hypothesis. Kaniel, Saar, and Titman (2008) argue that retail investors' contrarian trading provides liquidity to the market and leads to the positive predictive power of past marketable retail order imbalances for future stock returns. Therefore, in equation (8), we use the part of the marketable retail order imbalance related to past returns, *mroib(contrarian)*, as a proxy for liquidity provision. Recall that our results in Table VII show that the contrarian component of marketable retail order flow cannot significantly predict future stock returns. Does this finding completely rule out the liquidity provision hypothesis for the predictive power of marketable retail order flow? We are afraid not. As mentioned in our earlier discussion of our approach's caveat, we can only rule out the liquidity provision hypothesis under the assumption that the contrarian trading pattern captured by *mroib(contrarian)* is a perfect proxy for liquidity provision. This seems to us to be a reasonable assumption, but as far as we can tell, it cannot be directly confirmed by any data that we observe.[18] In this subsection, we provide more results on the liquidity provision hypothesis using approaches other than the predictive regression.

An important piece of evidence in support of the liquidity provision hypothesis in Kaniel, Saar, and Titman (2008) is the contemporaneous relation between the marketable retail order imbalance and stock returns.[19] To be more specific, Kaniel, Saar, and Titman (2008) examine the past, contemporaneous, and future returns of intense buy and sell portfolios of retail investors. In their paper, buy and sell order flows of retail investors are measured using the "net

pressure (proxied by lag order imbalance) is more important than lagged price pressure in equation (11). The coefficient on *mroib(w, contrarian)* stays insignificant. The coefficient on *mroib(w, other)* is 0.0006 with a t-statistic of 5.07, while the coefficient on *mroib*($w-1$, *other)* is 0.0008 with a t-statistic of 14.47 in Table VII, Panel B. The residual component remains significant but becomes slightly smaller in this case.

[17] To be more specific, the retail order imbalance has a low autocorrelation coefficient between 10% and 20%, but the positive autocorrelation lasts for a long period. Here the persistence refers to the long horizon rather than the magnitude.

[18] For example, recent studies such as Arif, Ben-Rephael and Lee (2016) and Chakrabarty, Moulton, and Trzcinka (2017) show that directional trading by active funds is highly persistent and price destabilizing. If the retail trades provide liquidity to these active funds, then liquidity provision can also go through the persistence channel, rather than the contrarian channel.

[19] We thank an anonymous referee for this suggestion.

individual trading" (NIT) measure. Each week, they first sort all firms into decile groups using the previous week's NIT, they then track the excess returns to these different groups over the four weeks before and after the portfolio construction. The excess return of each portfolio is computed by subtracting the return on a market proxy (the equal-weighted portfolio of all stocks in the sample). Here, we follow their approach, but use our marketable retail order flow measures *mroibvol* and *mroibtrd*. Results using *mroibvol* are reported in Table VIII; results using *mroibtrd* are reported in Internet Appendix Table IA.V.

The main results of Kaniel, Saar, and Titman (2008) are reported in their Table III, which contains three main findings. First, the stocks that retail investors sell during the portfolio construction week (week 0)—the intense selling group—experience significantly positive excess returns before week 0, while the stocks that retail investors buy during week 0—the intense buying group—experience negative excess returns. This is a typical contrarian trading pattern of selling winners and buying losers. In Panel A of our Table VIII, the first row corresponds to the firms intensely sold by retail investors using marketable orders. The mean excess return over the 20 days prior to the selling week is 0.67%. The bottom row corresponds to the firms intensely bought by retail investors using marketable orders. The mean excess return on these stocks over the 20 days prior to the selling week is −1.29%. Both numbers are highly significant and confirm Kaniel, Saar, and Titman's (2008) first finding.

The second finding of Kaniel, Saar, and Titman (2008) is that after retail investors buy or sell, the stocks that retail investors sell during week 0 (the intense selling group) experience negative excess returns, while the stocks that retail investors buy during week 0 (the intense buying group) experience positive excess returns. These results show that retail trading can predict returns in the correct direction. In Panel A of Table VIII, we find that firms intensely sold by retail investors using marketable orders (first row) experience a mean excess return over the 20 days after the selling week of −0.30%, while the firms intensely bought by retail investors using marketable orders (bottom row) experience a mean excess return of 0.57%. Both numbers are again highly significant and confirm Kaniel, Saar, and Titman's (2008) second finding.

Finally, for the contemporaneous relation over week 0, Kaniel, Saar, and Titman (2008) find that the contemporaneous excess return is significantly *positive* for stocks that retail investors *sell* and *negative* for stocks that they *buy*. Since the return signs are opposite of the retail trading direction, they interpret this finding as in favor of the liquidity provision hypothesis. From the column for $k = 0$ in our Table VIII, Panel A, however, we find that for firms intensely sold by retail investors using marketable orders, the contemporaneous return is significantly negative at −0.24% with a t-statistic of −5.30, whereas, for firms intensely bought by retail investors, the contemporaneous return is significantly positive at 0.11% with a t-statistic of 2.69. Our findings show consistent, rather than opposite, signs between contemporaneous marketable retail trading and return direction, which does not line up

Table VIII

Marketable Retail Order Imbalance and Contemporaneous Returns, Replicating Kaniel, Saar, and Titman's (2008) table III Using mroibvol

This table presents analysis of market-adjusted returns around net buying and selling activity as given by our scaled marketable retail order imbalance measure *mroibvol* (based on the number of shares traded). For each nonoverlapping week in the sample period (January 2010 to December 2015), we aggregate the daily order imbalance measures weekly to form weekly *Mroib* deciles. Each stock is put into decile according to its *Mroib* in the current week. Decile 1 contains stocks with the most net selling (negative *Mroib*), while decile 10 contains stocks with the most net buying (positive *Mroib*). We present results for four portfolios: (i) decile 1, (ii) deciles 1 and 2, (iii) deciles 9 and 10, and (iv) decile 10. Let k be the number of days prior to or following portfolio formation each week. In Panel A, we calculate eight cumulative return numbers for each of the stocks in a portfolio: $CR(t-k, t-1)$, where $k \in \{20, 15, 10, 5\}$ days and t is the first day of the formation week, and $CR(t+1, t+k)$, where $k \in \{5, 10, 15, 20\}$ days and t is the last day of the formation week. The return on each portfolio is then adjusted by subtracting the return on a market proxy (the equal-weighted portfolio of all stocks in the sample). We present the time-series mean and t-statistic for each market-adjusted cumulative return measure and for the market-adjusted return during the intense trading week ($k = 0$). In Panel B, we present the time-series mean and t-statistic for weekly market-adjusted returns in the four weeks around the formation week (i.e., $CR(t-k, t-k+4)$, where $k \in \{20, 15, 10, 5\}$ days and t is the first day of the formation week, and $CR(t+k-4, t+k)$, where $k \in \{5, 10, 15, 20\}$ and t is the last day of the formation week). ** and * indicate significance at 1% and 5% level, respectively (both against a two-sided alternative). The t-statistic is computed using Newey-West (1987) correction

Panel A: Cumulative Market-Adjusted Return

Mroibvol Bid-Ask Return	Intense Selling (Decile 1)		Selling (Deciles 1 & 2)		Buying (Deciles 9 & 10)		Intense Buying (Decile10)	
	Mean	t-Stat	Mean	t-Stat	Mean	t-Stat	Mean	t-Stat
$k = -20$	0.0067**	5.48	0.0063**	8.10	−0.0111**	−19.00	−0.0129**	−12.49
$k = -15$	0.0056**	5.62	0.0055**	8.71	−0.0096**	−20.87	−0.0109**	−12.91
$k = -10$	0.0041**	5.40	0.0042**	9.04	−0.0074**	−20.83	−0.0084**	−12.93
$k = -5$	0.0027**	6.06	0.0028**	10.07	−0.0047**	−24.09	−0.0053**	−15.62
$k = 0$	−0.0024**	−5.30	−0.0019**	−5.65	0.0011**	4.02	0.0011**	2.69
$k = 5$	−0.0016**	−3.89	−0.0012**	−5.07	0.0018**	9.58	0.0024**	6.99
$k = 10$	−0.0023**	−3.16	−0.0018**	−4.35	0.0028**	8.38	0.0036**	5.89
$k = 15$	−0.0025*	−2.45	−0.0022**	−3.97	0.0036**	8.39	0.0046**	5.51
$k = 20$	−0.0030*	−2.36	−0.0025**	−3.63	0.0043**	8.89	0.0057**	5.74

(*Continued*)

Table VIII—*Continued*

Panel B: Weekly Market-Adjusted Return								
Mroibvol Bid-Ask Return	Intense Selling (Decile 1)		Selling (Deciles 1 & 2)		Buying (Deciles 9 & 10)		Intense Buying (Decile 10)	
	Mean	t-Stat	Mean	t-Stat	Mean	t-Stat	Mean	t-Stat
$k = -20$	0.0010**	2.75	0.0008**	3.36	−0.0016**	−8.15	−0.0019**	−5.90
$k = -15$	0.0016**	4.12	0.0014**	5.66	−0.0022**	−12.70	−0.0026**	−7.97
$k = -10$	0.0014**	3.59	0.0015**	5.88	−0.0028**	−12.66	−0.0030**	−7.62
$k = -5$	0.0027**	6.06	0.0028**	10.07	−0.0047**	−24.09	−0.0053**	−15.62
$k = 0$	−0.0024**	−5.30	−0.0019**	−5.65	0.0011**	4.02	0.0011**	2.69
$k = 5$	−0.0016**	−3.89	−0.0012**	−5.07	0.0018**	9.58	0.0024**	6.99
$k = 10$	−0.0006	−1.51	−0.0006*	−2.42	0.0010**	5.38	0.0013**	3.68
$k = 15$	−0.0001	−0.21	−0.0004	−1.74	0.0009**	4.61	0.0010**	2.77
$k = 20$	−0.0005	−1.34	−0.0003	−1.29	0.0006**	3.84	0.0010**	2.80

with the liquidity provision hypothesis proposed by Kaniel, Saar, and Titman (2008).

What might explain the difference between our results? It might come from differences in the retail order imbalance variable, the sample period, or coverage. Our main variable comes from marketable retail order flows, while Kaniel, Saar, and Titman (2008) use retail order imbalance from both marketable and nonmarketable order flows. Between marketable and nonmarketable orders, marketable orders are more likely to be aggressive. In terms of the sample period, Kaniel, Saar, and Titman (2008) examine the January 2000 through December 2003 period, while our sample is January 2010 through December 2015, making the two samples about 10 years apart. In terms of coverage, Kaniel, Saar, and Titman's (2008) sample covers from NYSE's Consolidated Equity Audit Data (CAUD), which only contains retail trades that are executed on that exchange. During the Kaniel, Saar, and Titman (2008) sample period, only a small number of brokerages sent their retail order flow to NYSE. As a result, NYSE's market share of overall retail order activity was (and remains) quite small. In comparison, our sample covers a broad swath of retail order flow from TAQ, which contains all off-exchange and nearly all retail marketable orders. To summarize, the liquidity provision hypothesis receives at most mixed support in our sample.

D. Public News and Marketable Retail Order Imbalance

Our earlier results indicate that marketable retail order flows may contain valuable information about future stock price movements, which might be a surprise to many. As Kaniel, Saar, and Titman (2008) note, "… it is unclear how individuals, who have far fewer resources than institutions, could gain the upper hand in discovering private information and trading on it profitably in such a widespread fashion."

To better understand whether marketable retail investors can be informed traders and the nature of the information they might possess, in this section, we examine the relation between marketable retail order flow and public news. We introduce the public news data in Section II D.1, and investigate the extent to which the information in marketable retail flow is related to public news in Section II D.2.

D.1. Marketable Retail Order Imbalance and Future Returns across News Topics

We obtain news data from Thomson Reuters News Analytics (TRNA), which contains prominent public news articles for a broad set of firms starting in 2003. TRNA provides key information about each news item, such as the ticker, the time stamp of the news story, the news topics covered in the story, and sentiment scores for each article. News topics are grouped into five categories: cross market, general news, economy, equities, and money/debt. Each category contains several news subtopics; we identify 58 such subtopics in our sample.

The sentiment score measures the probabilities of the article being positive, negative, or neutral and is computed using Thomson Reuters' proprietary algorithm. We compute a net sentiment score as the difference between the positive and negative sentiment scores for each stock day. The news data are available from January 2010 to December 2014, which covers most of our main sample. We use tickers to match the news data with our marketable retail order imbalance data and obtain a merged sample of 3,854,813 stock-day observations.

We first provide simple statistics for the relations among news, returns, and marketable retail order flow. To examine whether our measure of public news can predict future stock returns, we estimate the following Fama-MacBeth (1973) regression:

$$ret\ (i,w) = f0\,(w) + f1\,(w)\ \times sent\,(i,w-1) + f2\,(w)'\,controls\,(i,w-1) + u6\,(i,w). \tag{12}$$

Here, $sent(i, w-1)$ is the average TRNA net sentiment score for firm i during week $w{-}1$, calculated by averaging nonmissing news sentiment for firm i within week $w{-}1$. The results are reported in regressions I and II of Table IX, Panel A. In regression I, the coefficient on the past week's public news sentiment is 0.0008 with a t-statistic of 3.31. The positive and significant coefficient indicates that net sentiment in public news can predict the cross-section of next week's stock returns. When we include the past marketable retail order imbalance in regressions III and IV, the predictive power of public news sentiment stays about the same. Interestingly, in the presence of contemporaneous public news sentiment in regressions III and IV, the coefficients on past marketable retail order imbalances are also positive and significant, with similar magnitudes as in Table III, indicating that public news sentiment does not reduce the predictive power of marketable retail order flow for future stock returns.

To better understand how marketable retail order imbalances are related to public news, we next estimate the contemporaneous relation between the two using the following Fama-MacBeth (1973) specification:

$$sent(i,w) = g0\,(w) + g1\,(w)\ \times mroib\,(i,w) + g2\,(w)'\,controls\,(i,w-1) + u7\,(i,w). \tag{13}$$

We find that the current week's marketable retail order imbalance is significantly and positively related to same-week public news sentiment in 10 out of the 58 subtopics. We present the coefficient estimates for these 10 cases in Panel B of Table IX. These 10 subtopics represent about 38% of total news days, and they mostly contain firm-level news. For instance, for the subtopic "RESF" (results forecast) in the news type "equities," the coefficient $g1$ is 0.0054 with a significant t-statistic of 3.90, indicating that the marketable retail order imbalance has a positive and significant contemporaneous relation with news related to forecasts of company results. Of the 10 subtopics, four belong to the

Table IX

Relation between Public News and Marketable Retail Order Flow

This table reports analysis of the relation between public news and our marketable retail investor order imbalance. Our sample period is January 2010 to December 2014, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. In Panel A, we examine whether public news can predict the cross-section of future stock returns. We estimate the Fama-MacBeth (1973) regression specified in equation (12). The dependent variable is weekly returns, computed in two ways: using the end-of-day bid-ask average price or using the CRSP closing price. The independent variable is *sent*($w-1$), which is the average TRNA net sentiment score for firm i during week w by averaging nonmissing news sentiment for firm i within week $w-1$. We also include past marketable retail order imbalance *mroibvol* (based on the number of shares traded by marketable retail) in regressions III and IV. As control variables, we include the previous-week return, *ret*($w-1$), previous-month return, *ret*($m-1$), and previous six-month return, *ret*($m-7$, $m-2$), log book-to-market ratio (*lbm*), log market cap (*size*), monthly turnover (*lmto*), and monthly volatility of daily returns (*lvol*), measured at the end of the previous month. In Panel B, we examine the relation between contemporaneous public news sentiment and marketable retail order imbalance across subtopics. We estimate the Fama-MacBeth (1973) regression specified in equation (13). The dependent variable is the weekly net sentiment score, *sent(i,w)*. The independent variable is the marketable retail order imbalance measure *mroibvol*. As control variables, we include the previous-week return, *ret*($w-1$), previous-month return, *ret*($m-1$), previous six-month return, *ret*($m-7$,$m-2$), log book-to-market ratio (*lbm*), log market cap (*size*) and monthly turnover (*lmto*), and monthly volatility of daily returns (*lvol*). Coefficients on controls are not reported for brevity. To account for serial correlation in the coefficients, the standard errors of the estimated coefficients are adjusted using Newey-West (1987) with five lags.

Panel A: Predicting Returns Using Public News and Marketable Retail Order Flow

Regression Dependent Variables Order Imbalance	I Bid-Ask Return		II CRSP Return		III Bid-Ask Return Mroibvol		IV CRSP Return Mroibvol	
	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat
Intercept	0.0057	2.48	0.0066	2.83	0.0061	2.66	0.0070	3.02
Sent($w-1$)	0.0008	3.31	0.0009	3.64	0.0008	3.33	0.0009	3.66
Mroib($w-1$)					0.0009	10.61	0.0010	11.53
Ret($w-1$)	−0.0088	−2.65	−0.0105	−3.10	−0.0090	−2.70	−0.0107	−3.16
Ret($m-1$)	0.0008	0.38	0.0009	0.44	0.0013	0.64	0.0015	0.72
Ret($m-7$,$m-2$)	0.0001	0.15	0.0001	0.11	0.0002	0.21	0.0001	0.18
Lmto	0.0000	−1.03	0.0000	−1.29	0.0000	−1.14	0.0000	−1.41
Lvol	−0.0435	−2.15	−0.0465	−2.34	−0.0444	−2.20	−0.0477	−2.40
Size	−0.0001	−0.90	−0.0001	−1.04	−0.0001	−0.99	−0.0002	−1.14
Lbm	0.0001	0.29	0.0001	0.61	0.0001	0.44	0.0002	0.79
Adj. R^2	5.01%		5.01%		5.06%		5.08%	

(*Continued*)

Table IX—*Continued*

Panel B: Contemporaneous Relation between Sentiment and Order Imbalance					
Topic	Type	Description	N	$g1$	t-Stat
RESF	Equities	Results forecast	102,515	0.0054	3.90
AAA	Money/debt	Debt rating news	23,405	0.0131	3.66
DIP	General news	Diplomacy	6,057	0.0362	3.34
DIV	Equities	Dividend	24,282	0.0093	2.77
IGD	Money/debt	Investment grade debt	6,760	0.0261	2.76
DRV	Cross market	Derivatives	18,061	0.0238	2.58
DBT	Money/debt	Debt markets	73,600	0.0060	2.55
MTG	Money/debt	Mortgage-backed debt	7,764	0.0264	2.52
RES	Equities	Corporate results	176,699	0.0031	2.36
JUDIC	General news	Judicial	28,280	0.0096	2.31

category "money and debt" and three to the category "equity," with the highest significant t-statistics obtaining for the subtopics "results forecast" and "debt rating news." Interestingly, the marketable retail order imbalances are never statistically significantly correlated with the news related to the economy. This finding implies that marketable retail investors may have valuable information at the firm level rather than at the market level. We find further support for this view below in Section III.A, where we find that retail investors cannot reliably predict future market-wide returns.

In Internet Appendix Table IA.VI, we also examine whether retail order imbalances can directly predict our measure of public news. We find that the signs of the coefficient are mixed. Indeed, in several cases, marketable retail order flows significantly predict future public news sentiment with negative coefficients. This does not contradict our earlier results of a positive and significant contemporaneous relation between marketable retail order flow and our measure of public news. Rather, these results simply show that marketable retail flows cannot predict our measure of future public news with the expected signs for these subtopics.[20] Using a joint test, we fail to reject the null that marketable retail order flows cannot jointly predict future public news.

D.2. Public Information and Other Information

The results above show that marketable retail order imbalances are associated with some types of contemporaneous public news, particularly firm-level news. In this subsection, we probe deeper into marketable retail order flows' predictive power that is associated with these public news releases because it is also possible that marketable retail traders possess and trade on nonpublic information that eventually makes its way into prices, but not via an identifiable public news release.

We investigate this question empirically using a two-step decomposition procedure similar to that in Section II.C. In the first step, we estimate a Fama-MacBeth (1973) regression and decompose the weekly order imbalance into four components for week $w-1$, as follows:

$$\begin{aligned} mroib(i, w-1) = {} & h0\,(w-1) + h1\,(w)\, mroib\,(i, w-2) + h2(w-1)' ret\,(i, w-2) \\ & + h3\,(w-1)\, sent\,(i, w-1) + u8\,(i, w-1)\,. \end{aligned} \tag{14}$$

Here, $sent(i, w-1)$ is the average TRNA net sentiment score for firm i during week $w-1$, which we use to capture information in contemporaneous public news releases. After we obtain the time series of coefficients,

[20] To give some context for why retail order flow does not show significant predictive power for future public sentiment in our sample, it could be the case that the public news we observe is noisy, that the information retail investors have does not warrant a specific news story, that the public news is published further into the future than the horizons we examine here, or that retail investors incorporate other useful public news into their trading, such as the SeekingAlpha posts studied by Farrell et al. (2020).

$\{\widehat{h0}(w-1), \widehat{h1}(w-1), \widehat{h2}(w-1)', \widehat{h3}(w-1)\}$, we define the following terms:

$$\begin{aligned}
\widehat{mroib}_{i,w-1}^{persistence} &= \widehat{h1}\ (w-1)\, mroib\,(i, w-2)\,,\\
\widehat{mroib}_{i,w-1}^{contrarian} &= \widehat{h2}\ (w-1)' ret\,(i, w-2)\,,\\
\widehat{mroib}_{i,w-1}^{publicnews} &= \widehat{h3}\ (w-1)\, sent\,(i, w-1)\,,\\
\widehat{mroib}_{i,w-1}^{other} &= \widehat{u8}\ (i, w-1) + \widehat{h0}\,(w-1)\,.
\end{aligned} \tag{15}$$

The sum of the four components above is exactly $mroib(i, w-1)$. As before, we denote the part related to past order imbalance as the "persistence" component, which is related to the price pressure hypothesis, the part related to past returns as the "contrarian" component, which is connected to the liquidity provision hypothesis, and the part related to contemporaneous public news sentiment as the "public news" component. Finally, we denote the residual part as the "other" component, which we attribute to marketable retail investors' nonpublic information that is not incorporated into prices via an identifiable news release.

In the second stage, we use the Fama-Macbeth (1973) methodology to estimate the following regression, which is the analog of equation (6):

$$\begin{aligned}
ret\ \ (i, w) = j0\,(w) + j1\,(w)\,\widehat{mroib}_{i,w-1}^{persistence} &+ j2\,(w)\,\widehat{mroib}_{i,w-1}^{contrarian}\\
+ j3\,(w)\,\widehat{mroib}_{i,w-1}^{publicnews} &+ j4\,(w)\,\widehat{mroib}_{i,w-1}^{other}\\
+ j5\,(w)'\, controls\,(i, w-1) &+ u9\,(i, w)\,.
\end{aligned} \tag{16}$$

Since we decompose the original order imbalance measure $mroib(i, w-1)$ into four parts, related to persistence, contrarian trading, public information, and a residual, the coefficients in equation (16) reveal how each component helps predict future stock returns.

Notice that in the first-stage estimation, the public news component is derived from a contemporaneous relation between current news, rather than past news, and current marketable retail order flow. From the perspective of the empirical design, we can link the marketable retail order imbalance with past, contemporaneous, or future public news, but the interpretations would be different. If we use future public news, the interpretation would be that marketable retail order flow "anticipates" future public news. If we use past public news, the interpretation would be that previously "incorporated" public news can be a component of the marketable retail order imbalance. When we use contemporaneous public news sentiment, we interpret the related part of marketable retail order flow as contemporaneously "processed" public news. Here, we choose not to use future public news because if we project *mroib(w*-1) on *sent(w)* in the first stage, then $\widehat{mroib}_{i,w}^{publicnews}$ would capture news from week w

and would have a mechanical correlation with *ret(w)*, which is the dependent variable in the second-stage estimation, and hence, the regression would no longer be predictive. We also choose not to use past public news because we would like to maximize the explanatory power of public news for marketable retail order flow, while contemporaneous public news sentiment likely nests the information in past public news.

Table X, Panel A, reports results for the first-stage decomposition. The patterns with respect to how past marketable retail order imbalances and past returns affect the current order imbalance are similar to those in Table II. The coefficient on contemporaneous public news sentiment ranges between 0.0249 and 0.0305, with t-statistics higher than 10. This clearly indicates that more positive news is associated with more contemporaneous purchases by marketable retail investors. The average adjusted-R^2s for the first-stage estimation are mostly between 5.49% and 8.59%.

Panel B of Table X reports results for the second-stage decomposition. From the top half of the panel, the coefficients on $\widehat{mroib}_{i,w-1}^{persistence}$ are positive and highly significant, and the coefficients on $\widehat{mroib}_{i,w-1}^{contrarian}$ are mostly insignificant, similar to the findings in Table VII. The coefficients on the public news components of order imbalance, $\widehat{mroib}_{i,w-1}^{publicnews}$, are also all insignificant, indicating that the contemporaneous public news component of marketable retail order imbalances does not help predict future returns significantly. In contrast, the "other" component of the marketable retail order imbalance measure is always positive and significant in the regressions. For instance, in the first regression, it has a coefficient of 0.0008 with a highly significant t-statistic of 13.98. This result is consistent with the view that marketable retail investors trade on information that is not incorporated into prices via the public news releases we measure. The bottom half panel shows that when we move from the 25$^{\text{th}}$ percentile firm to the 75$^{\text{th}}$ percentile firm, the "other" component of the marketable retail order flow accounts for 0.07% to 0.10% of weekly return differences, which is more than half of the return difference that the marketable retail order imbalance can predict overall. These results suggest that public news is noisy, the predictive power of marketable retail investors' order imbalance is not related to an identifiable public news release, or the retail order imbalance is only related to public news releases in the more distant future.

Returning to the question raised at the beginning of this subsection, how can retail investors using marketable orders, with far fewer resources than institutions, get the upper hand in discovering nonpublic information? Here we offer two possible explanations.

First, as an investor group, retail investors can be heterogeneous. For instance, it is possible that some individual investors might simply be endowed with nonpublic, valuable firm-specific information. These individuals might naturally obtain value-relevant information by working in the same industry or for a customer or supplier. Alternatively, they may expend effort to acquire information, for example, by studying the parking lots of retailers to assess demand growth. Farrell et al. (2020) provide evidence that some retail

Table X

Predictability Decomposition Using Public News Releases

This table reports estimation results on a decomposition of our marketable retail order flow measure's predictive power for future returns. Our sample period is January 2010 to December 2014, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. We estimate two-stage Fama-MacBeth (1973) regressions specified in equations (14) to (16). The dependent variable is weekly returns, computed in two ways: using the end-of-day bid-ask average price or using the CRSP closing price. The independent variables are scaled marketable retail order imbalance measures: *mroibvol* (based on the number of shares traded) and *mroibtrd* (based on the number of trades). In the first-stage estimation, the order imbalance measures are decomposed into four components. The variable *mroib*($w-1$, *persistence*) is estimated in the first stage using past order imbalance and reflects price pressure. The variable *mroib*($w-1$, *contrarian*) is estimated in the first stage using past returns over different horizons and is connected to the liquidity provision hypothesis. The variable *mroib*($w-1$, *public*) is estimated in the first stage using week $w-1$ news sentiment, which proxies for marketable retail order imbalances that predict returns associated with future news releases. The residual part of the previous-week order imbalance from first-stage estimation is denoted as "other," which we attribute to retail investors' valuable private information about future returns that is incorporated into prices but is not associated with an identifiable public news release. As additional control variables, we include previous-week return, *ret*($w-1$), previous-month return, *ret*($m-1$), and previous six-month return, *ret*($m-7$, $m-2$). Other control variables are log book-to-market ratio (*lbm*), log market cap (*size*), monthly turnover (*lmto*), and monthly volatility of daily returns (*lvol*), all measured at the end of the previous month. To account for serial correlation in the coefficients, the standard deviations of the time series are adjusted using Newey-West (1987) with five lags.

Panel A: First Stage of Projecting Order Imbalance on Persistence, Past Return, and Public News

Regression	I		II		III		IV	
Dependent Variables	Mroibvol($w-1$)		Mroibvol($w-1$)		Mroibtrd($w-1$)		Mroibtrd($w-1$)	
Return	Bid-Ask Return		CRSP Return		Bid-Ask Return		CRSP Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat
Intercept	−0.1510	−24.08	−0.1505	−24.03	−0.1132	−16.89	−0.1127	−16.85
Mroib($w-2$)	0.2208	112.64	0.2209	112.67	0.2827	125.44	0.2829	125.48
Ret($w-2$)	−0.8918	−36.22	−0.9051	−37.00	−0.8940	−34.04	−0.9098	−35.01
Ret($m-1$)	−0.2169	−13.68	−0.2156	−13.60	−0.1702	−10.56	−0.1687	−10.47
Ret($m-7$,$m-2$)	−0.0264	−4.54	−0.0264	−4.55	−0.0080	−1.17	−0.0081	−1.19
Sent($w-1$)	0.0249	11.60	0.0249	11.60	0.0305	13.81	0.0305	13.81
Adj. R^2	5.49%		5.50%		8.58%		8.59%	

(*Continued*)

Table X—*Continued*

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

Regression Order Imbalance Dependent Variables	I Mroibvol(w) Bid-Ask Return		II Mroibvol(w) CRSP Return		III Mroibtrd(w) Bid-Ask Return		IV Mroibtrd(w) CRSP Return	
	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat
Intercept	0.0055	2.40	0.0060	2.63	0.0055	2.39	0.0060	2.62
Mroib($w-1$,persistence)	0.0027	8.50	0.0029	9.16	0.0018	7.28	0.0020	8.02
Mroib($w-1$,contrarian)	0.0088	0.55	0.6947	1.03	−0.0182	−1.00	0.0419	1.18
Mroib($w-1$,public news)	0.1150	1.17	−0.0134	−0.35	−0.0386	−0.84	0.0020	0.04
Mroib(w-1,other)	0.0008	13.98	0.0009	15.16	0.0006	10.15	0.0007	11.35
Ret($w-1$)	−0.0217	−6.25	−0.0250	−7.12	−0.0218	−6.28	−0.0251	−7.16
Ret($m-1$)	0.0059	0.54	0.6585	1.03	0.0106	1.05	0.0119	1.41
Ret($m-7$,$m-2$)	0.0014	1.01	0.0511	1.02	0.0059	0.99	−0.0004	−0.22
Lmto	0.0000	−2.40	0.0000	−2.75	0.0000	−2.35	0.0000	−2.71
Lvol	−0.0273	−1.63	−0.0252	−1.52	−0.0266	−1.59	−0.0244	−1.47
Size	−0.0001	−0.68	−0.0001	−0.73	−0.0001	−0.73	−0.0001	−0.79
Lbm	0.0001	0.31	0.0001	0.66	0.0001	0.25	0.0001	0.59
Adj. R^2	4.22%		4.23%		4.21%		4.22%	
	Interquartile	Return Diff	Interquartile	Return Diff	Interquartile	Return Diff	Interquartile	Return Diff
Mroib($w-1$,persistence + contrarian+public news)	0.2760	0.0707%	0.2763	0.0761%	0.3609	0.0614%	0.3611	0.0678%
Mroib($w-1$,other)	1.1202	0.0932%	1.1203	0.1006%	1.1654	0.0745%	1.1654	0.0829%

investors make use of valuable firm-level analysis contained in SeekingAlpha posts.

Second, our data only contain marketable retail market orders, that is, they exclude retail limit orders. Retail limit orders could have opposite information, which would at least partially offset our findings for marketable retail market orders. In addition, it is not clear that the counterparties to marketable retail orders are necessarily "better informed" institutional investors. For example, Chakrabarty, Moulton, and Trzcinka (2017) document that uninformed "short-term" institutional investors constitute a nontrivial part of the market.

III. Further Discussion

Marketable retail order imbalances can predict future stock returns. This predictive ability lasts up to eight weeks and is stronger for smaller and lower priced firms. In this section, we discuss several related issues to put the predictive power of marketable retail order imbalance into perspective. In Section III.A, we discuss whether marketable retail investors' trading can predict the market's overall movement. In Section III.B, we look into potential contamination of the retail subpenny trades by dark pools using subsample analysis. We examine whether the predictive power is related to overall market conditions in Section III.C, and we investigate the predictive power of odd lot marketable retail orders in Section III.D. Because marketable retail trades occur with different sizes, we examine the predictive power of large versus small trade sizes in Section III.E. In Section III.F, we attempt to shed light on the role of wholesalers in one setup by examining the magnitude of price improvement and the profitability of interacting with marketable retail order flows. We identify the nature of the information captured by marketable retail order flows by linking marketable retail order imbalances to earnings news in Section III.G. In Section III.H, we examine whether marketable retail order imbalances can continue to predict future returns if we control for overall market order imbalances. Finally, we examine the implicit assumption of price improvements in Section III.I. To save space, all returns in this section are bid-ask returns.

A. Aggregate Marketable Retail Order Imbalance

If marketable retail order imbalances can predict future stock returns in the cross-section, retail investors using marketable orders may also be able to predict aggregate market moves. To investigate this possibility, we aggregate marketable retail order imbalances across all firms to predict aggregate stock market returns. We estimate the following equation:

$$mkt\ \left(w+1, w+k\right) = \ m0 + m1\ \times aggmroib\left(w\right) + u10\left(w+1, w+k\right), \quad (17)$$

where $mkt(w{+}1,w{+}k)$ is the future k-week cumulative market return from week $w{+}1$ to week $w{+}k$, and $aggmroib(w)$ is the current aggregated marketable retail order imbalance measure for week w. We compute *aggmroib* using either value-weighted or equal-weighted *mroibvol* or *mroibtrd*. The results are reported in Table XI, Panel A. As can be seen, they are the same regardless of the weighting scheme or order imbalance measure used; thus, there is no evidence that marketable retail order flows can reliably predict future market returns.

Our approach can also be used to identify the marketable retail order flow in exchange-traded funds (ETFs). In Table XI, Panel B, we examine marketable retail order flow in a large cross-section of ETFs over the same period. In cross-sectional predictive regressions of the form in equation (6), the coefficient is mostly around or below one bp, which is much smaller than the comparable coefficients shown in Table III, and the t-statistics are mostly insignificant. These results suggest that marketable retail order flows cannot predict sector returns or overall equity market returns. To separate sector-oriented information from broader market-wide information, we select the six largest ETFs that focus on the overall U.S. equity market by tracking comprehensive U.S. equity indexes: SPY, IVV, VTI, VOO, IWM, and IWB. The results are reported in the last row of Panel B. Consistent with the market timing results in Panel A, we find little evidence that marketable retail order flow can predict future returns on broad equity market ETFs.

B. Subsample Analysis

As mentioned in Section I.B, from 2008 into 2011, a few dark pool operators were accused of violating Regulation NMS in accepting, ranking, and executing subpenny trades and were eventually fined by the SEC. These questionable dark pool trades account for about 0.5% of total share volume during this period. Since these dark pools mainly cater to institutions, our identification of retail flows using subpenny trades could be "contaminated," and we cannot identify which trades are from the affected dark pools. In this subsection, we examine whether our main results hold for both a "contaminated" subsample and a later subsample that is not contaminated by these dark pool subpenny practices.

To be more specific, we reestimate the key results in Table III for the 2010 to 2012 and 2013 to 2015 subsamples. For comparison, in Table III, we find that marketable retail order flow can predict future stock returns, with a regression coefficient of 0.0009, a t-statistic of 15.60, and an interquartile weekly return difference of 10.89 bps. In Table XI, Panel C, for the 2010 to 2012 subsample, the coefficient on marketable retail order flow is 0.0010, with a t-statistic of 11.52 and an interquartile weekly return difference of 12.13 bps. For the 2013 to 2015 subsample, the coefficient on retail orders becomes 0.0009, with a t-statistic of 10.57 and an interquartile weekly return difference of 9.74 bps. We also test whether the coefficients from the two subperiods are significantly different from each other. For the $mroib(w-1)$ coefficients in regressions I and II, the difference is 0.0001 with a t-statistic of 1.85; for the coefficients in regressions III and IV, the difference is 0.0001 with a t-statistic of 0.67. Therefore, we cannot reject the hypothesis that the coefficients from these two subperiods are the same.

Table XI

Additional Analysis

Our sample period is January 2010 to December 2015, and our sample firms are all common stocks listed on U.S. stock exchanges with a share price of at least $1. Standard errors are calculated using Newey-West (1987). In Panel A, we estimate equation (17). The dependent variable is the n-week-ahead weekly value-weighted market return. The independent variables are two scaled marketable retail order imbalance measures: *mroibvol* (based on the number of marketable retail shares traded) and *mroibtrd* (based on the number of marketable retail trades). In all other panels, the regression is specified in equation (6) and estimated using Fama-MacBeth (1973) regressions. In Panel B, the dependent variable is weekly returns on approximately 1,000 ETFs. In Panel C, we estimate the coefficients for different subsamples. In Panel D, we estimate the coefficients for different VIX regimes. In Panel E, the independent variables are two scaled odd lot marketable retail order imbalance measures, *oddmroibvol* (based on the number of odd lot shares traded) and *oddmroibtrd* (based on the number of odd lot trades). In Panel F, we estimate the coefficients for different trade size. In Panel G, we estimate the coefficients for different amounts of price improvement. In Panel H, we estimate a variant of equation (6) that allows the predictive relationship to differ based on the variable event day, an indicator that takes a value of 1 if day t is an announcement day and 0 otherwise. In Panel I, we estimate the coefficients controlling for overall order imbalance computed by Lee-Ready (1991) algorithm. In Panel J, we estimate the coefficient when the effective spread is less than 1 cent. The dependent variable is weekly returns, computed in two ways: using the end-of-day bid-ask average price or using the CRSP closing price. The independent variable is one of the two scaled marketable retail order imbalance measures *mroibvol* or *mroibtrd*. Control variables for the cross-sectional regressions are the same as in Table III, except that we do not include a book-to-market variable in the ETF regression; coefficients on these variables are not reported.

Panel A: Predicting Future n-Week Market Return

	Mroibvol Value Weight		Mroibvol Equal Weight		Mroibtrd Value Weight		Mroibtrd Equal Weight	
Horizon	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat	Coef.	t-Stat
1 week	0.0037	0.50	−0.0053	−0.57	0.0054	0.92	−0.0038	−0.46
2 weeks	0.0101	0.79	−0.0030	−0.20	0.0120	1.21	0.0007	0.06
4 weeks	0.0044	0.20	−0.0236	−1.04	0.0073	0.43	−0.0136	−0.63
6 weeks	−0.0061	−0.22	−0.0356	−1.25	0.0022	0.10	−0.0216	−0.80
8 weeks	0.0075	0.20	−0.0046	−0.10	0.0118	0.41	0.0044	0.11
10 weeks	0.0051	0.11	−0.0114	−0.23	0.0101	0.28	−0.0038	−0.08
12 weeks	−0.0059	−0.10	−0.0315	−0.58	0.0000	0.00	−0.0227	−0.46

Panel B: Using Marketable Retail Mroib to Predict ETF Returns

Order Imbalance Dependent Variables	Mroibvol Bid-Ask Return		Mroibtrd Bid-Ask Return	
	Coef.	t-Stat	Coef.	t-Stat
All ETFs	0.0001	2.04	0.0001	1.68
Interquartile	1.4726		1.4737	
Return diff	0.0153%		0.0118%	
Broad market ETFs	−0.0004	−0.81	0.0005	1.52

(*Continued*)

Table XI—*Continued*

Panel C: Subsample Analysis

Regression	I		II		III		IV	
Period	2010 to 2012		2013 to 2015		2010 to 2012		2013 to 2015	
Order Imbalance	Mroibvol		Mroibvol		Mroibtrd		Mroibtrd	
Dependent Variables	Bid-Ask Return		Bid-Ask Return		Bid-Ask Return		Bid-Ask Return	
	Coef.	*t*-Sat	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat
Mroib	0.0010	11.52	0.0009	10.57	0.0008	9.39	0.0007	8.10
Interquartile	1.2357		1.1424		1.3328		1.1266	
Return diff	0.1213%		0.0974%		0.1041%		0.0826%	

Panel D: Different Market Conditions

	Vix <= 18%		Vix>18%	
Dependent Variables	Bid-Ask Return		Bid-Ask Return	
Independent Variables	Coef.	*t*-Stat	Coef.	*t*-Stat
Mroibvol	0.0009	13.49	0.0010	9.36
Mroibtrd	0.0007	10.32	0.0008	7.60

Panel E: Predicting Stock Returns Using Odd-Lot Order Imbalances

Order Imbalance	Mroibvol		Mroibtrd	
Dependent Variables	Bid-Ask Return		Bid-Ask Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat
Odd lot	0.0001	1.41	0.0001	0.77
Interquartile	1.2734		1.1314	
Return diff	0.0154%		0.0086%	

Panel F: Different Marketable Retail Trade Sizes

Order Imbalance	Mroibvol		Mroibtrd	
Dependent Variables	Bid-Ask Return		Bid-Ask Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat
Small trades (<400 shares)	0.0004	5.77	0.0004	4.48
Large trades (≥400 shares)	0.0009	7.25	0.0008	5.85

Panel G: Different Price Improvement Amounts

Order Imbalance	Mroibvol		Mroibtrd	
Dependent Variables	Bid-Ask Return		Bid-Ask Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat
Less price improvement	0.00071	9.30	0.00042	5.57
More price improvement	0.00021	3.04	0.00018	2.43

(*Continued*)

Table XI—*Continued*

Panel H: Earnings Surprises				
Order Imbalance Dependent Variables	Mroibvol Bid-Ask Return		Mroibtrd Bid-Ask Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat
Mroib	0.0003	8.16	0.0004	11.98
Mroib * eventday	0.0003	1.47	0.0002	1.31

Panel I: Marketable Retail versus Overall Order Imbalance						
Order Imbalance Dependent Variables	Overall Mroib Bid-Ask Return		Mroibvol Bid-Ask Return		Mroibtrd Bid-Ask Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat	Coef.	*t*-Stat
Retail Mroib			0.0011	6.14	0.0006	3.33
Overall Mroib	0.0004	3.32	0.0000	0.10	0.0001	0.51

Panel J: When Effective Spread Is Less than 1 Cent				
Order Imbalance Dependent Variables	Mroibvol Bid-Ask Return		Mroibtrd Bid-Ask Return	
	Coef.	*t*-Stat	Coef.	*t*-Stat
Mroib	0.0008	4.48	0.0004	2.45

We draw two observations from this exercise. First, the predictive power of retail order flow for the “cleaner” second subperiod is significant, which shows that our results are robust. Second, we cannot reject the hypothesis that the relevant coefficients from these two subperiods are the same. In addition, even during the earlier subperiod, we believe that fewer than 10% of all subpenny trades are misclassified as retail. For these reasons, we choose to use the entire 2010 to 2015 sample period to conduct our main analysis.

C. Market Conditions

Barrot, Kaniel, and Srear (2016) find that marketable retail trades contain more information when markets are volatile, in particular, when the VIX option-implied volatility index is high. Their sample spans the period 2002 to 2010, during which the VIX experienced dramatic changes. In contrast, our sample period is 2010 to 2015, when the VIX was far less volatile. Nevertheless, we divide our sample in half based on whether the VIX is higher or lower than its historical median of 18%.

We reestimate equation (6) separately for the high- and low-VIX subsamples. The results are presented in Panel D of Table XI. Comparing the low- and high-VIX regimes, we find that the coefficient on *mroibvol* is quite similar, but the *t*-statistic is higher when VIX is low than when it is high. This result might not be surprising, given that the volatility of all variables increases when VIX is high. Overall, the predictive power in both high- and low-VIX regimes is positive and significant.

D. Odd Lots

In this section, we investigate the behavior of odd lot marketable retail trades over the post-December 2013 period, when odd lot transactions are reported to the consolidated tape. Can odd lot marketable retail order flow predict future firm-level returns? To address this question, in Panel XI, Panel E, we estimate regression (6) using odd lot marketable retail order imbalances. Both coefficients are positive but not statistically significant. We conclude that the predictive power of the odd lot marketable retail order imbalance is much weaker than that of the overall marketable retail order imbalance.

E. Order Sizes

As shown in Figure 1, Panel A, a median market order submitted by a retail investor is around $7,000. The median marketable retail trade is about 400 shares. The "stealth trading" literature argues that medium-size orders are more likely to be informed and that large orders are usually broken into smaller orders.

To examine whether if information content differs according to order size, we partition the orders into large versus small groups using 400 shares as the cutoff and we estimate the predictive regression for each group separately. The results are reported in Table XI, Panel F. We find that both large and small orders predict future stock returns, but the larger orders' predictive power is stronger. Our results suggest that more informed marketable retail investors may demand immediacy by using larger market orders and that stealth trading does not seem to characterize the trading of retail investors who use marketable orders.

F. Wholesaler/Internalizer's Perspective: Profitability of Marketable Retail Order Flow

If marketable retail order flow is sufficiently informed, trading with these orders would be unprofitable. This might raise the question of whether our results are consistent with the apparently profitable business model of internalizers and wholesalers. Ultimately, as long as the information content of marketable retail order flow is less than the bid-ask spread being charged, internalizers and wholesalers can still earn positive revenues by trading with these orders. For example, if a marketable retail buy and a marketable retail sell order arrive at the same time, they offset each other, and a wholesaler earns the full bid-ask spread charged (the quoted spread less the price improvement given). Ultimately, internalizers and wholesalers are only exposed to adverse selection on marketable retail order imbalances. The summary statistics in Table I show that there is a substantial amount of offsetting marketable retail order flow. The interquartile range for the volume-based daily order imbalance measure is from −0.301 to 0.217, indicating that even at the ends of these ranges, more than two-thirds of the marketable retail order flow in such a stock on a given day constitutes offsetting buys and sells.

To get a better sense of the profitability of interacting with marketable retail order flow, we compute standard microstructure information-content measures for the marketable retail trades in our sample. Specifically, we calculate proportional effective spreads, one-minute price impacts, and one-minute realized spreads for all marketable retail buys and sells during 2015. Realized spreads are a standard proxy for trading revenue earned by a liquidity provider such as a wholesaler. We apply standard data filters, eliminating all trades where effective spreads exceed $1, and we calculate dollar-volume-weighted averages across all stocks. We find that the mean effective half-spread is 16 bps. The one-minute price impact is four bps, leaving a realized half-spread of 12 bps. In other words, interacting with our identified marketable retail order flow appears to be profitable (at least before other costs) for wholesalers/internalizers because the bid-ask spreads are sufficiently large. The liquidity provider (in this case, the wholesaler or internalizer) loses about four bps (the price impact) of the bid-ask spread to short-term price moves, but this leaves about 12 bps (the realized spread) of the bid-ask spread as average trading revenue to the liquidity provider. Note that the realized spread is a very crude measure of trading revenue. Furthermore, we cannot measure payments made by wholesalers to introducing brokers, nor can we measure the other costs associated with a wholesaling or internalization operation. However, these realized spreads are considerably higher than the realized spreads associated with on-exchange transactions, so we conclude that the price-improvement business model is potentially quite profitable for wholesalers and internalizers that can successfully segment order flow.

We can also examine some of the segmentation that is performed by these liquidity providers. For instance, while there is substantial competition among wholesalers, the magnitude of price improvement can vary substantially across orders. These liquidity providers are likely to rationally incorporate the potential information embedded in marketable retail orders and offer price improvement only up to the point at which they can still profit from the trade. On the one hand, if they infer that there might be relevant information in the marketable retail order, they might offer less price improvement; on the other hand, if they conclude that the marketable retail order is unlikely to contain relevant information, they might be willing to offer more price improvement. To the extent this is true, the predictive power of marketable retail order imbalances should be greater for marketable retail trades with less price improvement.

In the earlier sections, we group all orders with subpenny prices between 0.6 and 1 as marketable retail-initiated buy orders and those between 0 and 0.4 as marketable retail-initiated sell orders. In this section, we further divide orders into "less price improvement" and "more price improvement" types. For transactions with less improvement, we define buyer-initiated trades as transactions with prices between 0.8 and the round penny, and seller-initiated trades as trades with transaction prices between the round penny and 0.2 cents. For the "more price improvement" category, we define buyer-initiated trades as trades with transaction prices between 0.6 and 0.8, and seller-initiated trades as trades with transaction prices between 0.2 and 0.4. We compute marketable retail order imbalances following equations (1) and (2). We compare the predictive power of marketable retail order imbalances for "more" versus "less" price improvement by estimating equation (6) on each order imbalance measure separately.

Recall that the distribution of subpenny price improvements is displayed in Figure 1, Panel B. Most transactions occur at a round penny or a half-penny. Based on the other bins, each covering 0.1 cent, there is slightly more trading volume for the "less price improvement" category than for the "more price improvement" category. Regression results for the cross-section of future returns are reported in Table XI, Panel G. For the "less price improvement" type, the coefficients range from 0.0004 to 0.0007, with t-statistics above five. For the "more price improvement" type, the coefficients range from 0.0001 to 0.0002, with t-statistics below four. Clearly, both sets of marketable retail order imbalances have predictive power for future stock returns, but marketable retail trades with less price improvement have stronger predictive power, indicating that internalizers/wholesalers successfully price-discriminate against marketable retail orders with potentially more information content. Similar to the presence of large realized spreads, this observation also supports the viability of the business model, particularly for internalizers and wholesalers who can successfully distinguish between more- and less-informed order flows.

G. Earnings Announcements and Marketable Retail Order Flow

Kelley and Tetlock (2013) use the Dow Jones news archive to identify marketable retail order flows that are informed about cash flow news, they find that marketable retail market orders can predict earnings surprises.

Here, we examine whether marketable retail order flow becomes more predictive around earnings news. Specifically, we estimate a variant of equation (6) that allows the predictive relationship to differ based on the variable event day, an indicator that takes a value of 1 if day t is an earnings announcement day and zero otherwise. The results, as reported in Table XI, Panel H, show that the predictive power of marketable retail order flow is greater on announcement days, but the difference is not statistically significant. In Internet Appendix Table IA.VII, we directly replicate the results in Kelley and Tetlock (2013) and find that our marketable retail order flows can predict earnings news positively. The predictive power is statistically significant at the one-day horizon but insignificant over longer horizons. We are thus able to partially confirm Kelley and Tetlock's (2013) results. The difference may be attributable to the different sample periods and coverage we use. Specially, while we cover all subpenny trades for most stocks over the 2010 to 2015 period, Kelley and Tetlock (2013) cover about one-third of all marketable retail trades between 2003 and 2007.

H. Controlling for Overall Order Imbalances

Previous studies such as Chordia and Subrahmanyam (2004) find that overall order imbalances (calculated using all reported transactions, including individual and institutional types) can predict future stock returns. We use the Lee-Ready (1991) algorithm to compute the overall order imbalance from TAQ data. In our data set, overall order imbalances and marketable retail order imbalances are significantly correlated at around 30%. An interesting question is whether overall and marketable retail order imbalances are relatively orthogonal to each other. Specifically, if we control for the overall order imbalance, can the marketable retail order imbalance still predict future stock returns?

We address this question by proceeding in two steps, the results of which are reported in Table XI, Panel I. In the first step, we reestimate equation (6) using the overall order imbalance from the previous week rather than marketable retail order imbalance as a key predicting variable. Consistent with the literature, we find that overall order imbalances significantly predict future stock returns, with a coefficient of 0.0004 and a significant t-statistic of 3.32.

In the second step, we estimate equation (6) using the marketable retail order imbalance variables as key predicting variables, and we include the overall order imbalance as a control. With both marketable retail and market order imbalances in the model, marketable retail imbalances are significantly positive, and they completely drive out the effect of overall order imbalances. The predictive power of the marketable retail order imbalance therefore seems to be stronger than that of the overall order imbalance measure.

Here, we want to be cautious about the interpretation in the sense that this finding does not necessarily indicate that the retail order flow that we identify is more informed than order flow from institutional investors. First, due to different calculation methods for the two measures, the difference between the overall *oib* and the marketable retail order imbalance is not the order imbalance from institutional investors. Second, we only calculate the order flow from marketable retail orders, which accounts for about half of the trades from retail investors, and the overall *oib*'s weaker predictive power might be partially a result of uninformed trading by other participants in the market.

I. When the Effective Spread Is Less Than 1 Cent

Our identification for buy and sell orders relies on the implicit assumption that price improvements are always a small fraction (less than half) of a cent. If price improvements are larger, our method may not correctly sign trades. For example, if a stock has a bid price of \$50.01 and an ask price of \$50.04, and a marketable retail market buy order arrives and is improved by 0.75

cents, the reported transaction price would be $50.0325 and our trade-signing approach would erroneously conclude that this is a sell order. We investigate the reliability of our identification method is reliable in three ways. First, as we discuss above, we cross-validate using the 2010 NASDAQ TRF sample. We find a trade sign error rate of only about 2% using this sample. Second, we examine intraday quote data from TAQ. For all 2015 trades that we can sign using our approach, we compare our buy-sell assignment to the trade sign from the Lee and Ready (1991) algorithm. We find that the trade signs match for 89.9% of the observations. Last, we impose a strict filter that requires the average effective spread from the previous month to be at most 1 cent and we reexamine our results. For stocks with a 1-cent spread, our trade-sign approach for subpenny trades should match the Lee-Ready (1991) algorithm exactly and should be virtually error-free overall. This strict filter excludes more than 80% of the data leaving us with only the most liquid stocks in the sample. The results are reported in Table XI, Panel J. We find that the marketable retail order imbalance still significantly predicts the next week's stock returns, with a coefficient of 0.0008 and a significant t-statistic of 4.48, consistent with the findings in Table III.

IV. Conclusions

In this paper, we exploit the fact that most marketable retail order flows in U.S. equity markets are internalized or sold to wholesalers. As a part of this routing process, marketable retail orders are typically given a small fraction of a penny per share of price improvement relative to the NBBO price, and this price improvement can be observed when the trade is reported to the consolidated tape. Institutional orders almost never receive this kind of price improvement, and thus, it is possible to use subpenny trade prices to identify a broad swath of marketable retail order flow. It is also straightforward to identify whether the marketable retail order is buying or selling stock—transactions at prices that are just above a round penny are classified as marketable retail sales, while transactions that are just below a round penny are marketable retail purchases.

We use this methodology to characterize the trading behavior and information content of marketable retail orders. We find that marketable retail order flows are on average contrarian over weekly horizons, buying stocks that have experienced recent price declines and selling stocks that have risen in the past week. More significantly, we find that marketable retail order flow can predict the cross-section of future stock returns. Over the next week, stocks with more positive marketable retail order imbalances outperform stocks with relatively negative marketable retail order imbalances by about 10 bps, which is on the order of 5% annualized. This predictability extends to about 12 weeks before dying off. Through an empirical decomposition exercise, we attribute less than half of the predictive power of marketable retail order imbalances to the order imbalance's persistence and potential liquidity provision by marketable retail investors' contrarian trading. The remainder of the predictive power (over half of it) is consistent with the view that the marketable retail order flow contains valuable information about future returns. Focusing on the information content of marketable retail trades, we provide some suggestive evidence that marketable retail order flows contain relevant information about short-term future earnings news that is not yet incorporated into price, but our examination of a standard news database does not find evidence that our retail investors can anticipate that particular set of future public news.

An important advantage of our methodology is that it is based on widely available intraday transaction data: Anyone with access to TAQ can easily identify marketable retail buys and sells using our approach. Our approach has many possible research applications. For example, future researchers can investigate certain behavioral biases to determine whether individual traders as a group exhibit them. Another possibility is studying the seasonality of and time-series variation in marketable retail order flow, including tax-related and calendar-driven trading, as well as activity around corporate events, such as dividends, stock splits, and equity issuance.[21]

Initial submission: November 1, 2017; Accepted: October 9, 2020
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

REFERENCES

Arif, Salman, Azi Ben-Rephael, and Charles M.C. Lee, 2016, Short-sellers and mutual funds: Why does short-sale volume predict stock returns? Working paper, Stanford University.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773−806.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Do retail trades move markets? *Review of Financial Studies* 22, 151−186.

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

Battalio, Robert, Shane A. Corwin, and Robert Jennings, 2016, Can brokers have it all? On the relation between make-take fees and limit order execution quality, *Journal of Finance* 71, 2193–2238.

Blume, Marshall E., and Robert F. Stambaugh, 1983, Biases in computed returns: An application to the size effect, *Journal of Financial Economics* 12, 387–404.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang, 2008, Which shorts are informed? *Journal of Finance* 63, 491–527.

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.

[21] Our measure is already used in a few studies. For instance, Farrell et al. (2020) find that our retail order imbalances are strongly correlated with the sentiment of Seeking Alpha articles and that the ability of retail order imbalances to predict returns is roughly twice as high on days with SeekingAlpha posts. McLean, Pontiff, and Reilly (2020) use our methodology to identify retail investor trading and study how different market participants trade with respect to 130 different stock return anomalies. Israeli, Kasznik, and Sridharan (2021) also identify retail investor trading using our methodology and use abnormal retail trading volume and Bloomberg searches as specific measures of retail and institutional investor attention.

Chakrabarty, Bidisha, Pamela C. Moulton, and Charles Trzcinka, 2017, The performance of short-term institutional trades. *Journal of Financial and Quantitative Analysis* 52, 1403–1428.

Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485−518.

Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709–740.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607−636.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Farrell, Michael, T. Clifton, Russell Green, Jame, and Markov Stanimir, 2020, The democratization of investment research and the informativeness of retail investor trading, Working paper, University of Virginia.

Fong, Kingsley Y.L., David R. Gallagher, and Adrian D. Lee, 2014, Individual investors and broker types, *Journal of Financial and Quantitative Analysis* 49, 431–451.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2018, Trading costs, Working Paper, AQR Capital Management.

Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829–853.

Israeli, Doron, Ron Kasznik, and Suhas A. Sridharan, 2021, Unexpected distractions and investor attention to corporate announcements, Working Paper, IDC Herzliya.

Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman, 2012, Individual investor trading and return patterns around earnings announcements, *Journal of Finance* 67, 639–680.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor sentiment and stock returns, *Journal of Finance* 63, 273−310.

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.

Kwan, Amy, Ronald Masulis, and Thomas H. McInish, 2015, Trading rules, competition for order flow and market fragmentation, *Journal of Financial Economics* 115, 330–348.

Lee, Charles M.C., and Balkrishna Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, *Journal of Financial Markets* 3, 83–111.

Lee, Charles M.C., and Mark J. Ready, 1991, Inferring investor behavior from intraday data, *Journal of Finance* 46, 733–746.

Lee, Yi-Tsung, Liu Yu-Jane, Roll Richard, and Subrahmanyam Avanidhar, 2004, Order imbalances and market efficiency: Evidence from the Taiwan Stock Exchange, *Journal of Financial and Quantitative Analysis* 39, 327–341.

McLean, R. David, Jeffrey Pontiff, and Christopher Reilly, 2020, Taking sides on return predictability, Working paper, Georgetown McDonough School of Business.

Menkveld, Albert J., Bart Z. Yueshen, and Haoxiang Zhu, 2017, Shades of darkness: A pecking order of trading venues, *Journal of Financial Economics* 124, 503–534.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703−708.

O'Hara, Maureen, Chen Yao, and Mao Ye, 2014, What's not there: Odd lots and market data, *Journal of Finance* 69, 2199–2236.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.