Heterogeneous Retail Investors and their Trading Behavior

Andrew Glover

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Ed deHaan, Chair

Elizabeth Blankespoor, Chair

Charles Lee

Program Authorized to Offer Degree:

Foster School of Business

University of Washington

Abstract

Heterogeneous Retail Investors and their Trading Behavior

Andrew Glover

Chairs of the Supervisory Committee:

Ed deHaan

Graduate School of Business, Stanford University

Elizabeth Blankespoor

Foster School of Business, University of Washington

Efforts to understand retail investors and their trading behavior often consider them as a homogeneous group. I examine the extent and implications of differences in trading judgments within the retail class. I identify a sophisticated subset of retail traders (the “Savvy”) and an unsophisticated subset (the “Naïve”) who trade in starkly different ways in response to the same information. The Savvy type appear to respond to and process firm disclosures, buying good news and selling bad news. The Naïve type buy bad news and sell good news. I also explore the judgments each type of retail investor makes regarding the timeliness and costliness of their trades. I find that even the subset that appears to process and understand firm disclosures fail to earn returns that exceed their transaction costs. My findings suggest generalizations about “retail investor behavior” by academics and regulators may overlook meaningful and predictable differences within this investor class.

Acknowledgements

I would like to thank the University of Washington, the Foster School of Business, and the Deloitte Foundation for financially supporting my research. I am also thankful for suggestions from helpful scholars at workshops conducted at University of Toronto, University of Colorado Boulder, Massachusetts Institute of Technology, Stanford University, Boston College, University of Texas at Austin, University of Michigan, and Pennsylvania State University.

I am grateful for the guidance of my dissertation committee, Ed deHaan (co-chair), Elizabeth Blankespoor (co-chair), Charles Lee, and Estelle Lingo (Graduate School Representative). Ed and Beth both gave many, many hours (that they didn't have to spare) to help me refine ideas and be precise in conveying them. They also provided the life/career coaching I needed. I feel very lucky to have them as advisors.

Next, I owe a debt to the faculty and PhD students in the Accounting Department at Foster. The group helped me be a better student, teacher, researcher, and person. In particular, Darren Bernard went above and beyond as the PhD coordinator, and Mallory Bell and Jake Everhart are the only cohort I could imagine surviving year one with.

I am very grateful for my family, both immediate and extended, both by blood and by law, and especially for my parents (Steve and Tina), who provided every opportunity for me to love learning. They expected me to both 'do good' and 'be good' and showed me how by example.

Finally, and eternally, for Sophie. With you, I feel I can do it all. Without you, I wouldn't much care to try.

Table of Contents

1. INTRODUCTION

Retail equity investors are increasingly important market participants. Numerous studies find that retail investors actively trade on firms' disclosures but tend to make poor trading decisions that ignore or misinterpret accounting news and fundamental signals.[1] I argue there are two underexplored dimensions of retail trading in response to financial news. First, retail investors are often considered as a monolith, conflating very sophisticated, well capitalized, and experienced individual traders with less experienced traders often associated with platforms like Robinhood and Reddit. This aggregation has the potential to obscure meaningful differences in trading behavior. Second, conclusions about whether retail trades are well-informed are typically based on simply whether the trades are directionally consistent with the news, which fails to consider both the timeliness and costliness of the trades. I extend the literature by re-examining retail trading around disclosures after considering both investor heterogeneity and the multi-dimensional nature of what constitutes an informed trade.

When faced with new, value-relevant information from firm disclosures, I argue that heterogeneous retail traders must make three fundamental judgments. First, does the information indicate firm value should increase or decrease (i.e., what is the sign of the news)? Second, how much of the return has already been priced (i.e., how timely is the trade)? Third, does the remaining unpriced return exceed transaction costs (i.e., how costly is the trade)?[2] Prior work has focused on the first of these judgments (the sign of the news) and largely considered retail traders as a homogeneous group, often concluding that they make uninformed trades in response to attention-grabbing events such as financial disclosures (e.g., Blankespoor et al. 2019; Barber et

[1] This is attributed to documented behavioral biases (for a review, see Barber & Odean 2013) as well as retail investors' limited ability to process information (for a review, see sections 3 and 4 of Blankespoor et al. 2020).
[2] Retail traders are unlikely to explicitly consider each judgment prior to trading. Rather, much as Friedman's billiards player behaves "as if" he mathematically computes the angles and forces for a complicated shot (Friedman 1953), retail traders intending to make profitable trades behave "as if" they considered these judgments.

al. 2022). I contend that the second and third judgments of timeliness and costliness can be equally important in determining the quality of an investor's trading decision, and that considering retail traders in aggregate overlooks within-group heterogeneity that has the potential to change inferences.

Partitioning retail investors by their degree of financial sophistication may reveal a subset who can successfully process financial information. Theory suggests that only some traders act as "information processors" who convert public disclosures into private signals and trade accordingly (Kim & Verrecchia 1994). I rely on theory that posits information is more valuable to investors with more wealth (Ohlson 1975), suggesting wealthier traders likely have greater capabilities and incentives to process value-relevant information.

I propose that larger trades are more likely to come from retail investors with more wealth, experience, and financial sophistication. Very small "microtrades" are more likely to come from less wealthy retail traders, with less experience and financial sophistication. For expositional simplicity, I refer to the more capable type (identified by their large trades) as the "Savvy type" and the less sophisticated type (identified by their microtrades) as the "Naïve type." My trade size partition classifies relatively large Savvy trades as between 100 – 1000 shares and Naïve microtrades as 10 shares or less.[3]

I focus on the trading behavior of each type in response to firm disclosures, which provides a setting to assess differential information processing and related outcomes. Specifically, I compare the Savvy and Naïve types' revealed judgments about 1) the sign of the news, 2) the timeliness of the trade, and 3) the costliness of the trade. An investor can be correct in their assessment of some of these elements, and incorrect about others. Importantly, I limit my

[3] Trade size is no longer a valid proxy for partitioning retail traders from institutional investors due to the advent of high frequency algorithmic trading in the early 2000s (Cready et al. 2014; O'Hara et al. 2014). I use trade size as a partitioning variable exclusively within the set of retail investors.

examination to only instances in which both the Savvy and Naïve types choose to trade on the same information, ensuring observed differences in trading outcomes are not a function of investment selection (i.e., "habitat"), but are rather driven by differences in judgments.

First, to make a correct judgment about the sign of the news, an investor must be willing and able to incur the processing costs necessary to determine the implications of the disclosure for firm value (Blankespoor et al. 2020). To assess this judgment, I follow prior work and consider the association between trade imbalance and returns (e.g., Chordia et al. 2002; Kumar & Lee 2006; Kelley & Tetlock 2013). Event returns serve as a proxy for the "goodness" or "badness" of the news. I find that the Naïve type trades are in the opposite direction of the sign of the news; in other words, their trade imbalance is significantly and negatively associated with the sign of the return.[4] In contrast, Savvy trades in response to firm disclosures appear consistent with the sign of the news. When news is positive, Savvy traders buy, and when it is negative, they sell. These results suggest that the Savvy type retail traders process information about the sign of the news "correctly," meaning directionally consistent with the return. The descriptive evidence in Figure 1 illustrates these differences around firm disclosures. In economic terms, a one standard deviation increase in Naïve (Savvy) buy-sell imbalance is associated with a -30 (+32) basis point cumulative 5-day return.

The second investor judgment is the timeliness of trading in response to the news. If most of the eventual price adjustment occurs before the trade takes place, the return to the investor will be close to zero, even if the security return is large. Consistent with prior research (Patell & Wolfson 1984; Lee 1992; Wertz 2021), I find that the pricing response to information events occurs rapidly in the period following the disclosure. For accounting information events like

[4] One explanation for this contrarian trading is that less sophisticated traders in recent years follow the popular internet investing forum adage to "buy the dip" (i.e., attempt to time the market).

earnings announcements (EAs), the response is mostly completed within the first few minutes of open market trading. Only a small fraction of retail trades occur during this brief window, meaning that many retail investors choose to trade despite missing most of the price response to information releases. Both the Naïve and Savvy retail traders appear to disregard price movements between an information release and the time of their trades, which is inconsistent with standard notions of informed trading (e.g., Grossman & Stiglitz 1980). For the Savvy type, the rapid adjustment of price means they are not able to fully benefit from their correct judgment about the sign of the news. On the other hand, the rapid price adjustment protects the Naïve type from losses they may have otherwise incurred. Thus, the Savvy type does not "earn" the full return when trading in response to firm disclosures and the Naïve type does not "lose" the full return. After considering the timeliness of their trades, a one standard deviation increase in Naïve buy-sell imbalance is associated with a -16 basis point cumulative 5-day return, while the Savvy type association is indistinguishable from 0.

The third investor judgment is whether the expected value of the trade exceeds transaction costs. While retail brokerages have driven down explicit costs to trade over the past several years (e.g., $0 commission trading, no account balance minimums, no trading fees, etc.), investors demanding immediacy must still pay the costs implicit in the bid-ask spread. To explicitly recognize these costs, I compute returns to buy trades and sell trades separately. After accounting for the cost of their trades, both Naïve and Savvy trades earn negative returns following firm disclosures, although the Savvy type continues to outperform the Naïve type.

In sum, and as shown in Figure 2, there are significant differences between the "representative trades" of the Naïve and Savvy traders in response to disclosures. These differences persist (though in smaller magnitudes) after considering both the timeliness and costliness of their trades. In cross sectional examination, I find that the Savvy type's outperformance of the Naïve type is greatest among stocks that are most difficult to trade; in other words, where an information advantage could be most exploited.[5]

My study contributes to the broad literature on the behavior of individual investors by examining heterogeneity among retail investor trading behaviors.[6] I complement recent work that acknowledges investor sophistication as an important determinant of equity market outcomes using alternative proxies for financial sophistication, such as brokerage preference (Fong et al. 2014; Eaton et al. 2022; Liu 2022; Michels 2024) or information source (Blankespoor et al. 2018; Blankespoor et al. 2019; Gomez et al. 2024) or a combination of the two (Chau 2025).[7] I also complement the literature that benefits from richly detailed international data sets of individual equity trading – albeit at the cost of some generalizability – (e.g., Grinblatt & Keloharju 2000; Barber et al. 2009a; Barrot et al. 2016; Jones et al. 2024) and seek to draw inferences from publicly available U.S. market data in the modern investing environment. In addition, the focus of my study is intra-firm and intra-event heterogeneity in retail investor judgments, trading behavior, and outcomes (i.e., holding constant the selection of the investment, how do different types of retail investors trade on the same information?). This contrasts with recent studies that identify cross-sectional differences in the types of firms or news retail investors trade in (e.g., Laarits & Sammon 2025; Betermier et al. 2025; Balasubramaniam et al. 2023), and others that assert conflicting findings can be reconciled by considering

[5] Kim & Verrecchia (1994) suggest traders more capable of processing information are incentivized to trade even in more expensive and illiquid markets when they are relatively well-informed.

[6] Prior work calls for more research on this issue. Kelley & Tetlock (2017) write, "Future empirical studies should test [hypotheses] based on heterogeneity in investor sophistication within groups of retail investors… our findings suggest that within-group heterogeneity could be just as important as accounting for differences between retail and institutional investors."

[7] Some recent studies have also explored investor heterogeneity as an explanation for differential outcomes in an alternative setting: the options market. Both trade complexity (e.g., multi-leg strategies) and trader characteristics (e.g., median holding period, median trade size, median trade frequency) have been used to proxy for financial sophistication (Hu et al. 2024; Beckmeyer et al. 2024; Bogousslavsky & Muravyev 2025).

differences in how intensively retail investors trade in different types of firms (e.g., Barber et al. 2023). I illustrate that different types of retail investors trade in very different ways given the same disclosure. Understanding this heterogeneity may also help address some conflicting answers to questions about aggregate retail investor behaviors, capabilities, and market impacts.[8]

Another contribution of my study is to extend how the literature considers the informedness of retail trades. Many studies operationalize informedness based solely on trade direction. I show that the additional dimensions of trade timeliness and costliness can change inferences. Specifically, I show that a conclusion as to the informedness of the Savvy type's trades varies based on whether you consider only the sign of the news (e.g., appear well-informed), the timeliness of the decision (e.g., appear neutral), or the costliness of the trade (e.g., appear poorly informed). I show that a trade may be in the "correct" direction while simultaneously being too late and too expensive to benefit the trader. Thus, considering the quality of individual investor judgments about trade timing and implicit trading costs can be as important as considering the directional response.

This study should prove helpful to regulators in the ongoing policy debate of how to "protect the everyday retail investor" (Avakian 2020). Recent SEC leadership has stated a desire to "update [the SECs] rules and drive greater efficiencies in our equity markets, particularly for retail investors" (Gensler 2022). Updating rules without considering the heterogeneity inherent in the broad category of "retail investors" may lead to mis-specified regulation.[9] Further, the findings of this study suggest that much of the risk to unsophisticated investors is due to their

[8] E.g., Do retail traders provide stabilizing support to the market (Welch 2022) or seek increased volatility (de Silva et al. 2025)? Are retail traders contrarians (Grinblatt & Keloharju 2000) or momentum traders (Boehmer et al. 2021)? Do retail traders earn poor returns (Barber & Odean 2000) or do their trades positively predict future returns (Hvidkjaer 2008)? If they earn positive returns, is it compensation for liquidity provision (Barrot et al. 2016) or due to private information (Kelley & Tetlock 2017) or a mix of both (Boehmer et al. 2021)?

[9] Leuz et al. (2025) provide evidence of this issue in Germany, observing that different types of traders have different propensities to engage in manipulative investment schemes and different responses to supervisory warnings. They opine that "understanding investor heterogeneity is likely critical for effective investor protection."

own judgments about public information. While there is general agreement that regulators should protect retail investors from fraud and unfair trading practices, it is less clear whether regulators should (or even can) protect individual investors from their own errors. In addition to concerns about individual investor outcomes, regulators are also intent on maintaining fair, orderly, and efficient markets that firms can access for capital. Retail trades comprise an increasing fraction of overall equity trading volume, and understanding how these trades influence price formation, market liquidity, and firm disclosure choices requires first understanding the retail traders.[10]

2. BACKGROUND, THEORY, AND MOTIVATING PREDICTIONS

Prior work largely considers retail investors as unsophisticated capital market participants suffering from both behavioral biases and heightened processing costs. As a few examples of the former, researchers have documented evidence of retail investor overconfidence (Odean 1999; Barber & Odean 2000), sensation- or entertainment-seeking behavior (Grinblatt & Keloharju 2009; Dorn & Sengmueller 2009), and attention-induced trading (Barber & Odean 2008; Barber et al. 2022). As some examples of the latter, studies find investors ignore or misinterpret financial information (Blankespoor et al. 2019; deHaan et al. 2023), are distracted by irrelevant information (Hirshleifer et al. 2009; deHaan et al. 2015; Drake et al. 2016; Israeli et al. 2021), and are influenced by disclosure channel (Engleberg & Parsons 2011; Blankespoor et al. 2018).

Considering retail investors as a homogeneous group implicitly "averages out" potentially meaningful differences in investor sophistication and capabilities. I posit that retail investors vary in their financial sophistication and information processing capabilities, and that these differences yield distinct judgments, behaviors, and outcomes.

[10] As examples, consider the Cianciaruso et al. (2023) analytical model which suggests firm valuation and disclosure choices are tied to the activity of noise traders, or the Li et al. (2026) structural model which suggests "small trader" fundamental-based trading is a driving factor of asset pricing anomalies. This suggests it is important to understanding different types of retail traders' liquidity needs, attention trading tendencies, and disclosure processing capabilities.

2.1. Heterogeneity in investor sophistication

Given the cost of transforming public disclosures into private signals, it is likely only a subset of more informed traders can accomplish this task (Kim & Verrecchia 1994). An extensive literature (both analytical and empirical) divides traders into more and less sophisticated types. When considering how to identify different types of traders, some theorize that the value of information is increasing in wealth (Ohlson 1975; Cready 1985).[11] Cready (1988) proposed that investors with more wealth are more likely to obtain and use information in making trading decisions. He demonstrates that larger trades (a proxy for wealth) cluster more tightly around earnings announcements than small trades, consistent with the idea that financial information events are more valuable to wealthier investors. Lee (1992) likewise partitions traders into individuals (making small trades) and institutions (making large trades). He finds that the institutional response occurs quickly and in the "correct" direction following earnings news, but that individual traders tend to buy regardless of the direction of the news. In these studies, as well as others dividing investors by sophistication (e.g., Lee & Radhakrishna 2000; Battalio & Mendenhall 2005; Hvidkjaer 2008; Malmendier & Shanthikumar 2007; Campbell et al. 2009) the purpose of the partition is to distinguish retail traders from institutional traders. The literature largely identifies retail as the unsophisticated "noise-trading" type, while identifying institutions as having the capabilities to make sophisticated trades.[12]

I argue that the same trade-size theory underlying many of the studies that partitioned retail traders from institutional traders can be applied to a subset of only retail traders (i.e., trade

[11] Wealth and financial sophistication are also linked empirically (Calvet et al. 2009).
[12] Some studies do suggest that retail trades predict returns because the traders have private information. For example, Kaniel et al. (2012) suggest that retail traders may "serendipitously come across what turns out to be valuable information in their day-to-day activities." Kelley & Tetlock (2013) similarly find evidence consistent with retail orders "aggregating novel information about firms' cash flows."

size proxies for the wealth and financial sophistication of a retail trader).[13] My aim is to illustrate differences within the retail investor class, rather than differences between retail traders and institutions. Much as partitioning institutional investors into granular types uncovers useful insights about their trading behaviors (e.g., Bushee 1998), partitioning the retail investor class into distinct types may uncover insights about their trading decisions and performance.[14]

Prior literature suggests that differences in sophistication plausibly impact both 1) market conditions and 2) retail investors' financial outcomes. Eaton et al. (2022) provides evidence on the former, finding that reduced trading by the least sophisticated retail investors (proxied for by brokerage choice) results in increased market liquidity and lower return volatility. Jones et al. (2024) provide evidence on the latter via proprietary trading data for Chinese retail investors, finding that less wealthy investors display overconfidence and gambling tendencies, while the wealthiest investors' trading predicts future earnings surprises and returns. In an alternative setting, deHaan & Glover (2024) find that lower income (i.e., less financially sophisticated) retail traders tend to overtrade and earn lower capital gains when they have greater market access.

I focus on disclosure response. I predict different types of retail traders will have differential abilities and incentives to process disclosures. These differences in information processing capabilities likely yield observably different trading judgments, decisions, and outcomes.[15] I predict that more sophisticated traders can more completely process both the information disclosed by a firm and the information available in market data (e.g., price,

[13] Using retail trades from the 1990s, Barber et al. (2023) illustrate positive associations between trade size and financial knowledge, investing experience, wealth, and income (see Fig. A1 of their supplementary materials).
[14] Some prior work has divided individual investors on ex-post realizations to identify "skilled" retail traders (e.g., Coval et al. 2021, Hwang et al. 2024). In contrast, I attempt to ex-ante classify investors by sophistication and examine responses to disclosure.
[15] My prediction is distinct from the literature examining ownership composition and trading differences across firms based on disclosure attributes (e.g., Allee et al. 2007; Miller 2010; Lawrence 2013; Kalay 2015). This literature suggests investors' differential processing capabilities shape investment selection. I abstract from the selection choice and examine how heterogeneous retail investors process the same information differently.

spread).[16] This should lead to differences between the Savvy and Naïve types' judgments regarding the sign of the news, the timeliness of their decisions, and the costliness of their trades.

These predictions are not without tension. It could be the case that, regardless of financial sophistication, retail investors perform poorly due to pervasive behavioral biases or structural disadvantages relative to professional traders. deHaan et al. (2023) document that retail bond traders (generally considered a more sophisticated group than retail equity traders) consistently make poorly informed trades in the opposite direction of accounting fundamentals. Other studies document that some professional traders exhibit behavioral biases such as loss aversion in their trading behavior (Coval & Shumway 2005), and that market participants with perfect foresight of disclosures still struggle to make profitable trades (Xie 2025), suggesting that even sophisticated traders are not simply expected-value-maximizing automatons.

2.2. Processing disclosures to determine the sign of the news

Investors wishing to trade on firm disclosures must make a judgment regarding whether the disclosed information increases or decreases the value of the firm. I examine the intersection of retail investor sophistication and firm disclosure to determine whether either type of retail investor appears to extract the sign of the news from a disclosure.[17]

The literature on the informedness of retail trading tends to focus on whether retail investors as a whole correctly interpret the sign of the news. Correctness is inferred based on trading consistently with the direction of the news. For example, Lee (1992) classifies earnings

[16] Note that sophisticated retail traders are not simply "small institutions," as retail traders are non-professional, part-time market participants who have neither the resources nor the agency problems of even the smallest institutions.
[17] A significant body of work examines associations between trading and information events. I focus exclusively on differential responses to information events within the retail class, rather than between retail and institutional traders (e.g., Shanthikumar 2004; Mikhail et al. 2007) or among retail as a whole (e.g., Li et al. 2024; Moss et al. 2024; Luo et al. 2025). I also have little to say regarding volume response (e.g., Bhattacharya 2001; Asthana et al. 2004) and instead focus on directional trading, which reveals more information about the investors' underlying judgments. Finally, I complement work examining the market impacts of retail trade in response to firm disclosures (e.g., Hirshleifer et al. 2008; Chen et al. 2015; Friedman & Zeng 2024) by focusing on the consequences for the traders.

news based on unexpected earnings relative to analyst forecasts and finds that retail investors buy regardless of the sign, which is interpreted as less-informed trading. Blankespoor et al. (2019) use a similar approach and find that retail investors largely disregard the positive or negative nature of a disclosure and instead trade on trailing returns. Similar to Lee (1992), Michels (2024) finds that extreme earnings news – whether positive or negative – is associated with only more retail buying. I extend this literature by examining differential responses to the same firm disclosure within the class of retail traders. I predict that there exists a type of retail trader that can correctly process the sign of the news.

2.3. Price adjustment and trade timeliness

In addition to correctly identifying the return sign implied by a firm disclosure, retail investors must also assess the extent to which the news is already priced. This requires a judgment regarding the anticipated "full" price adjustment and the proportion of that adjustment that is already reflected in prices.[18] Understanding that changing prices convey information to market participants is an element of financial sophistication. Price adjustments provide public (if imperfect) revelations of private information (Grossman & Stiglitz 1980). Retail traders may underestimate the speed at which price adjustments occur and misjudge the timeliness of their trades, leading to a poor trading decision even if they correctly identified the sign of the news.

Empirical evidence suggests that retail traders respond too late to information signals, and instead simply "chase" past returns (e.g., Frazzini & Lamont 2008), and that retail trading on high-attention events occurs only after price adjustments have occurred (e.g., Barber & Odean 2008). I predict the more financially sophisticated Savvy type make more timely trades than their Naïve counterparts. However, they may still lag professional participants, leaving it unclear

[18] The timeliness judgment is closely related to a belief about the magnitude of the price adjustment. If a trader incorrectly believes security XYZ's price will increase 8% when in actuality it will increase only 4%, they may incorrectly judge their trade to be timely when observing a price adjustment of "only" 4%.

whether either type of retail trader makes timely trades relative to the overall price response.

2.4. Market conditions and trade costliness

An individual may correctly identify the sign of the news and understand the timeliness of their trade and still fail to make a profitable trade after accounting for transaction costs. Odean (1999) finds that the securities bought by retail investors do not outperform the securities they sell by enough to offset trading costs. Barber & Odean (2000) go on to attribute most of the negative performance of individual traders to trading costs and frequency. While explicit costs have declined significantly over the last two decades, costs inherent in the bid-ask spread are still paid by active retail traders, and spreads have been widening in recent years (Barber et al. 2024). Processing information about trading costs should alter trading decisions.

Costs to trade are particularly salient in my setting as I consider marketable retail trades paying for immediacy via the bid-ask spread. I also focus on trades in response to firm disclosures. Spreads widen in the days and hours immediately surrounding a firm disclosure. For example, Amiram et al. (2016) find that spreads increase significantly just before and during earnings announcements, and then decrease to pre-announcements levels within 2 days.[19] Kim & Verrecchia (1994) analytically model this increase, and argue that traders who process information face a tradeoff between being more informed and trading in less liquid markets or being less informed and trading in more liquid markets. In my context, a retail trader must assess whether it is to their advantage to make a timelier (but more costly) trade immediately following a disclosure, or to wait until spreads return to pre-disclosure levels or lower.[20]

[19] See also Lee et al. 1993 and, for a more recent sample, Table 1 of Blankespoor et al. 2020. For descriptive illustrations of the size of the spread relative to prices in my sample, see the Online Appendix 1.7.
[20] i.e., for an uninformed trader, or one trading exclusively for liquidity reasons, it may be advantageous to wait for spreads to fall rather than trading in immediate response to firm disclosure.

3. DATA, VARIABLE CONSTRUCTION, AND DESCRIPTIVES

3.1. Sample Selection

The primary analyses use data from the Trade and Quote (TAQ) database. To isolate likely marketable retail trades from all other trades in TAQ, I follow the Boehmer et al. (2021) approach for identification and rely on sub-penny price improvement. Since 2005, trades were permitted to execute at amounts smaller than a penny only for marketable retail trades. However, odd lots (i.e., trades of less than 100 shares) have only been reported on the consolidated tape since December of 2013, so my sample ranges from January 2014 – December 2023.

I begin with all retail-identified trades occurring during market hours. I flag subsets of retail trades as either Naïve (1 – 10 shares) or Savvy (100 – 999 shares) and collapse counts of trades and shares to the Firm–Day level.[21,22] For my primary analyses, I require at least 5 trades from both the Naïve and Savvy type for a Firm–Day observation to be retained. I also exclude observations with buy – sell imbalances of exactly |1.00| (i.e., only buys or only sells) for either Savvy or Naïve, as this outcome is more likely to be indicative of low independent volume than high consensus. I limit my sample to trades executing at a price over one dollar. Because I classify type based on the count of shares in the trade, I also limit extreme nominal prices by trimming firms in the top and bottom 1% of nominal prices each year (e.g., Berkshire Hathaway's A shares – which trade at a nominal price of over $500k – are excluded).[23]

[21] The proposed cutoffs are based on mix of social media observations, interviews with money managers, and discussions with small retail traders. Portfolios with <10 share positions are often shared by retail traders on social media. In terms of the Savvy cutoffs, a 100 (1000) share trade of the median priced US equity in 2023 corresponds to a dollar value trade of ~$2,500 (~$25,000). This range seems within reason for wealthier individual traders. I confirm the robustness of my results to alternative thresholds, as well as after considering the impacts of fractional share trading (e.g., Bartlett et al. 2024; Lin 2024; Da et al. 2024) in the Online Appendix 1.1 and 1.9.

[22] In later analyses that focus on intraday trading and returns, I follow the same procedure but collapse trades to five-minute trading increments for each ticker.

[23] The fixed effect structure I employ examines within firm and date variation, holding constant mean nominal price. I could avoid this nominal price problem with a classification based on dollar values rather than shares, but at the cost of introducing firm selection issues. Firms with certain nominal prices would automatically be excluded for small dollar classifications, or oversampled for large dollar classifications. This is problematic as return-relevant fundamentals (e.g., BTM, MVE) as well as momentum effects (e.g., Du et al. 2026) vary predictably with nominal

I require firm returns from CRSP and some fundamentals from Compustat. Using the TAQ – CRSP – Compustat merged sample of Firm–Days, I then identify firm press releases and earnings announcement in Ravenpack and IBES, respectively. The final dataset consists of 5.4 million Firm–Day observations.

3.2. Variables and Proxies

My primary independent variable is the trade imbalance of each type of retail trader. I follow the approach of Barber et al. (2024) and sign retail trades based on the quote midpoint (Lee & Ready 1991).[24] After trades are signed, I compute the trade imbalance for the Savvy (*Savvy IB*) and Naïve (*Naïve IB*) types as:

$$\textit{Retail Trade IB}_{i,t,j} = \frac{(\textit{Total Buy Trades}_{i,t,j} - \textit{Total Sell Trades}_{i,t,j})}{(\textit{Total Buy Trades}_{i,t,j} + \textit{Total Sell Trades}_{i,t,j})}$$

For each firm i on day t for retail investor type j, the trade imbalance measure can range between -0.99 and 0.99 after the data restrictions outlined in section 3.1. To facilitate comparisons, I then standardize each type's imbalance by subtracting the mean and dividing by the standard deviation. Prior work has varied in constructing trade imbalance using either trades, shares, or dollars, but generally concludes the alternate constructions produce highly correlated imbalance measures (Chordia et al. 2002; Boehmer et al. 2021).[25] I use a trade-based imbalance construction for two reasons. First, it better captures a single individual's "vote" on the news, whereas a share- or dollar-based imbalance construction weights larger positions more heavily

prices. See the Online Appendix 1.2 for illustrations of the issues, as well as test results for an alternative, dollars-based threshold for classification.

[24] In a current working paper, Battalio et al. (2023) illustrate that the Type I error rate in the Boehmer et al. (2021) identification method is higher than suggested by prior studies using the measure. This represents a violation of Regulation NMS, which limits sub-penny fulfillment (other than at the half penny) to retail market orders. The SEC declined to provide comments on these purported violations when I contacted them, but reiterated the minimum pricing increments for NMS securities. I conduct several robustness tests based on concerns raised in the Battalio et al. (2023) working paper in the Online Appendix 1.8, and find my results are unlikely to be driven by violations of Regulation NMS.

[25] The O'Hara et al. (2014) recommendation against a trade-based imbalance because of the outsized impact of missing odd-lots does not apply as I consider only periods after odd-lots appear on the consolidated tape.

than small ones. Since the range of shares used to proxy for the Naïve type is significantly narrower than that for the Savvy type, using a share- or dollar-based construction would yield different interpretations of the Savvy imbalance relative to the Naïve imbalance. Second, a trade-based imbalance measure is not impacted by either stock splits or price changes.

I examine trading imbalances and returns in response to firm disclosures. I create indicator variables for both firm disclosure days and the subset of disclosure days that are earnings announcement days. If firm *i* issues a press release or earnings announcement on day *t,* the indicator variable *Firm Disclosure Day* takes a value of 1. If firm *i* issues an EA on day *t,* the variable *EA Day* takes a value of 1. In both cases, disclosures made after market hours are coded to the following trading day. All variable constructions are outlined in Appendix A.

The primary dependent variable is stock returns. I construct the cumulative return (*CRET*) for the initial tests as the summation of returns for a 5-day trading week from $t = 0$ to $t = 4$.[26] Alternative return specifications based on average trade times and execution prices are detailed in sections 4.2 and 4.3. I employ a time-based fixed effect structure to control for cross sectional market returns and thus do not market-adjust returns before cumulating.

3.3. Some descriptive evidence

Descriptive statistics for the primary dataset of 5.4M Firm–Day observations are presented in Table 1. The cumulative return for the trading week (*CRET*) has a slightly positive mean and median of 0.002. Approximately 12% of all Firm–Days are firm disclosure days, and around 2% are classified as EA days.

In examining retail trade imbalance (pre-standardization), immediate heterogeneity

[26] To study retail investor information processing and trading decisions in response to disclosure, I need both returns and imbalance measures to include day 0. This differs from some prior work on retail investment that investigates whether retail investors have private information that predicts future returns. As in other studies which use trade imbalance and contemporaneous returns (e.g., Kumar & Lee 2006; Kaniel et al 2008; Barber et al. 2009b) I do not attempt to identify the causal direction of the association.

across types can be noted in the differences in means and standard deviations. The Naïve type has a relatively strong buy bias (mean = 0.038) and a wider distribution across Firm–Days (SD = 0.332). The Savvy type has a slight sell bias (mean = -0.011) and less variation across Firm–Days relative to the Naïve type (SD = 0.254).[27] Over the full sample, there are typically more Savvy trades per Firm–Day (mean = 339 trades) than Naïve trades (mean = 227 trades). However, the frequency of Naïve type trades has been steadily growing, as illustrated in Figure 3. Figure 3 also illustrates the jump in trading activity in late 2019 and early 2020 as retail brokerages shifted to $0 commission trading (Even-Tov et al. 2026). This also aligns with the introduction of fractional share trading at several retail brokerages, the beginnings of the COVID-19 pandemic, and distribution of economic impact payments (i.e., stimulus checks) in the United States. Each of these shocks to participation likely impacted smaller, less capitalized, and less experienced traders more than large and experienced traders (e.g., the elimination of a $6.95 commission is far more salient for a 5-share trade than a 500-share trade). These participation changes over time are initial descriptive evidence that my selected cutoffs capture different types of retail investors, with the prevalence of very small trades increasing at a far greater rate than large trades as barriers to trading decline.[28] While Figure 3 suggests differences between types, it cannot speak to differential retail investor response to new information. Section 4 examines retail trading in response to firm disclosures.

[27] In the Chinese setting, Jones et al. (2024) find the opposite pattern after parsing by wealth. Exploring this difference is beyond this paper's scope, but illustrates the potential for varied results based on institutional setting (e.g., differences between the retail-dominated Chinese equity market and the institution-heavy US equity market).

[28] As an additional validity check, I manually collect Robinhood trade imbalances from 2024 – 2026 and compare it to my proposed Naïve trader imbalance proxy. They are positively associated. See Online Appendix 1.10 for results.

4. MAIN ANALYSES

4.1. Sign of the news

4.1.1. Firm Disclosure Days

I follow prior work and examine the association between trade imbalance and returns to infer whether traders of either type appear to be informed. I estimate Eq. 1 as follows:

$$\begin{aligned} CRET_{i,t(0,4)} = \; & \beta_1 Retail\ Trade\ IB_{i,t,j} + \beta_2 Firm\ Disclosure\ Day_{i,t} \\ & + \beta_3 Retail\ Trade\ IB_{i,t,j} \times Firm\ Disclosure\ Day_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4)} \end{aligned} \quad (1)$$

CRET is the cumulative return to firm *i* over window *t(0,4)*. I use a 5-day window to allow time for the information that induced the trade to be fully reflected in price, and implicitly assume returns beyond the 5-day window are unassociated with the day (0) news.[29] *Retail Trade IB* is the imbalance of either the Naïve trades or the Savvy trades as defined in Section 3.2. Firm fixed effects hold constant time invariant firm characteristics.[30] Date fixed effects limit the analysis to within-day variation across firms, controlling for the effects of macroeconomic trends and daily market-wide returns. Standard errors are clustered both by firm and date.

The indicator *Firm Disclosure Day* takes a value of 1 if firm *i* issued a press release or earnings announcement on day *t=0*. The interaction between *Retail Trade IB* and *Firm Disclosure Day* captures the incremental association between retail trade imbalance and returns related to trading on new firm-disclosed information.[31] A positive association between investor trading behavior and the return following an information release suggests the investor processes

[29] The results are robust to a tighter return measurement window (see the Online Appendix 1.3). Prior literature has varied in the length of the event response window, but periods between 3 days (e.g., Blankespoor et al. 2019) and 6 days (e.g., Twedt 2016) are common.
[30] The firm fixed effect necessarily introduces information about future firm realizations. This "look ahead" attribute of fixed effects is problematic for a study hoping to develop a predictive model or trading strategy. As I do neither, this design choice succinctly controls for firm-specific attributes.
[31] An alternative to this interactive approach is to retain only Firm–Days with disclosures. I elect to retain all sample days to control for the baseline association between type imbalance and returns. This also provides ample data for estimating the fixed effects. In untabulated tests, I limit the sample to disclosure days and draw similar inferences.

the information in the disclosure to extract the sign of the news.[32]

Table 2A, columns (1) – (3), presents results. Column (1) presents the interaction effect between the Naïve trade imbalance and firm disclosure as the primary variable of interest. The β_3 coefficient is negative and significant (β_3 = -0.0023; t-stat – 12.70), suggesting that the association between Naïve trade imbalance and returns on firm disclosure days is incrementally more negative than their baseline relation. In economic magnitude, a one standard deviation increase in Naïve imbalance is associated with a decline in *CRET* of (β_1=) 7 basis points on a non-disclosure day, and (β_1 + β_3=) 30 basis points on a firm disclosure day. The implication is that the Naïve traders typically trade in the opposite direction of returns, and that this negative association is stronger when new, firm-specific information is released.

Column (2) examines the Savvy trade imbalance and *Firm Disclosure Day.* The interaction of Savvy trade imbalance and firm disclosure moves positively and significantly in the opposite direction of the Naïve type results. A one standard deviation increase in the Savvy imbalance is associated with an increase in *CRET* of (β_1=) 24 basis points on a non-news day, and (β_1 + β_3=) 32 basis points on a firm disclosure day. This result suggests that the Savvy traders typically trade consistently with the direction of the return, and that this positive association is stronger when their trades are in response to firm disclosures.

In column (3), I include the trade imbalance of both types, as well as their respective interactions with *Firm Disclosure Day*. The interaction effects are slightly stronger and more significant, suggesting that conditioning on the trade imbalance of one type makes the resulting effects of the other type more pronounced. In sum, the results in Table 2A suggest that the Naïve type trade in the "wrong direction" following a firm disclosure, and Savvy type trade make

[32] A significant β_2 coefficient represents the return on firm disclosure days when retail trade imbalance is exactly zero. As I retain a Firm–Day observation only if both types trade, a trade imbalance exactly equal to zero exists only when there are identical numbers of buy and sell trades. This scenario is not directly relevant to my predictions.

trades consistent with the return after processing information to extract the sign of the news.

This pattern is depicted in event time in Figure 1. I plot the standardized trade imbalances of each type on both "Good news" and "Bad news" disclosures released on day $t = 0$. I infer the type of news based on the firm-specific market reaction (i.e., the firm's return less the market's return) to the news (Basu 1997; Nofsinger 2001). Upon the release of bad news (Figure 1A), the Naïve type buys while the Savvy type sells. Upon the release of good news (Figure 1B), the Naïve type sells while the Savvy type buys. This result contrasts with prior work that considers retail investors in aggregate and suggests individual investors simply buy all extreme news (e.g., Lee 1992; Michels 2024).

The findings for the Naïve type are generally consistent with the literature; these retail investors appear to make poorly informed trades that are not aligned with the sign of the firm disclosure. On the other hand, the results for the Savvy type differ from prior work and the typical understanding of retail investor information processing and trading behavior. The strong alignment between their trade imbalance and *CRET* – particularly on disclosure days – suggests they can process information into trading decisions.

4.1.2. Alternate approach: The Representative Trade

The previous section followed prior literature in measuring the association between returns and trade imbalance. I also construct an alternative measure for each disclosure day consisting of a "representative trade" and its associated "representative return", and test whether this representative situation differs between the Naïve and Savvy traders. While the regression results in 4.1.1. estimate linear relationships between imbalances and returns, the representative trade has the advantage of providing a simple summary outcome for my sample. To calculate the representative trade, I weight the return by the relative proportion of buy trades and sell trades:

$$\left(\left(\frac{Buy\ Trades_{i,t,j}}{Total\ Trades_{i,t,j}}\right) \times CRET_{i,t}\right) - \left(\left(\frac{Sell\ Trades_{i,t,j}}{Total\ Trades_{i,t,j}}\right) \times CRET_{i,t}\right)$$

This approach combines the trade imbalance and the return (with a positive sign on the buy trades and a negative sign on the sell trades) into an aggregated dependent variable. After computing the representative trade for each type on each firm disclosure day, I average across the sample to calculate μ_j and conduct a simple difference in means test.

The representative trades are reported in Table 2B. The Naive mean *CRET* across all disclosure days is -9 basis points, while the Savvy mean *CRET* is +5 basis points. Both are significantly different than 0, and the Savvy mean *CRET* is +14 basis points greater than the Naïve *CRET*. In my sample, the representative Naïve trade earns a significantly negative *CRET* in response to disclosure, while the representative Savvy trade earns a significantly positive *CRET* trading in response to the same information. The simplicity of the representative trade test lends itself to graphical summary, found in Figure 2.

4.2. Timeliness

4.2.1. Earned Return Specifications

Thus far, the results suggest there are identifiable types of retail investors, and that these types trade differentially in response to the same firm disclosures. In this section, I examine each type's ability to determine how much of the eventual price response has already taken place by the time they trade (i.e., their timeliness judgments).

First, Figure 4A illustrates the cumulative distributions of returns and retail trading on firms' disclosure days. The black hashed line plots the percentage of the total absolute return reached during each five-minute increment, which is conceptually similar to the "intra-period timeliness" curve used in prior work (e.g., Blankespoor et al. 2020).[33] I also plot trading activity for the Savvy and Naïve types by dividing the number of shares traded by each type during each five-minute increment by the type's total shares traded during the day.

The Savvy type have a slightly steeper curve at the extremes of the day and a flatter curve in the middle of the day, indicating a larger percentage of their trading activity occurs near the open and the close (the difference is statistically significant at both the beginning and end of the trading day; untabulated). The Naïve type have a slightly more linear curve, but with a similar pattern overall. More striking is the different pattern of the returns curve relative to the trading curves. While most of the price adjustment occurs in the first few minutes of the trading day, retail trades are spread throughout the day. Specifically, when around 50% of retail trading for the day has been completed, around 80% of the day's total return has been priced on firm disclosure days. This differential is even more extreme on EA days (see Figure 4B).

The implications are two-fold. First, most Savvy and Naïve trades occur too late to capture the majority of the return when they trade in response to the firm disclosures – particularly earnings announcements. The firm-specific news contained in the disclosure has largely been priced before their trades take place. This is likely a good thing for the financial outcomes of the Naïve type who tend to trade in the opposite direction of the return, as they are largely price protected by the speed of the price adjustment. However, the same speed that helps the Naïve trader erases any advantage of the more informed Savvy type, who appear to trade too slowly to capture much of the day's return even with directionally "correct" trades.

Second, and more generally, studies examining both returns and trading in response to an information event cannot simply use the close-to-close return as representative of gains or losses

[33] Like prior work, I exclude smaller disclosure day returns of less than 1% to avoid a small denominator problem. I also retain only firm disclosures that occur outside of regular trading hours. This is the large majority of press releases, and more than 95% of earnings announcements.

to trade. Because of the general efficiency of information in price, very few investors trading in response to information events capture the full price response. Some reckoning is required between the speed of trades relative to the speed of the price response.

To further investigate these points, I compute an alternative return designed to more closely reflect the return actually earned by traders, as opposed to the close-to-close return typically considered. *EarnedCRET* is computed as the return from the average executed trade on $t = 0$ to the midpoint close on $t = 0$, plus the return on days t=1 to t=4.[34] It is computed separately for each retail trader type to account for any differences in their typical trades. I then modify Eq. 1 by replacing *CRET* with *EarnedCRET*:

$$EarnedCRET_{i,t(0,4),j} = \beta_1 Retail\ Trade\ IB_{i,t,j} + \beta_2\ Firm\ Disclosure\ Day_{i,t} + \beta_3 Retail\ Trade\ IB_{i,t,j} \times Firm\ Disclosure\ Day_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4),j} \qquad (2)$$

Table 3A presents results from estimating Eq. 2. Column (1) regresses *EarnedCRET – Naïve* on the Naïve type imbalance and its interaction with *Firm Disclosure Day.* As in Table 2A, column (1), both the main effect and interaction are significantly negative, indicating the trading of the Naïve type continues to have a negative association with returns after accounting for the timeliness of their trades. However, the size of the coefficient and strength of the statistical relationship on the *Firm Disclosure Day* are much smaller in Table 3A than in Table 2A. This suggests much of the *Firm Disclosure Day*-related 5-day cumulative return accrues early on day 0, such that the Naïve type's tendency to sell good news and buy bad news does not harm them to the extent suggested in Table 2A. Column (2) presents the same analysis for the Savvy type. The decline in the association is even more precipitous than with the Naïve. In contrast to Table 2A, the Savvy type's trade imbalance now has an insignificant association with returns. This

[34] For example, consider a Firm–Day with an opening price of \$1.00 and a closing price of \$1.10. Suppose only three trades occur during the day: one immediately at the open for \$1.00, one halfway through the day at \$1.05, and one at the close for \$1.10. The day's return is (1.10 – 1.00)/1.00 = 10%, but the average return earned on day $t = 0$ is based on the average execution price of \$1.05 (1.10 – 1.05)/1.05 = 4.8%. The day t=0 return is then cumulated with the returns for days t=1 to 4.

suggests that after accounting for the timeliness of their trades, the Savvy type trades appear neutral, neither trading advantageously nor to their detriment.[35]

As in Section 4.1, I compute a representative trade for each type and compare the univariate means but with *EarnedCRET* rather than *CRET*. Table 2B presents the representative trades on firm disclosure days. The Savvy trade is marginally negative, while the Naïve trade is significantly negative. As predicted, the difference between the types persists after accounting for the timeliness of their trades, with the Savvy type outperforming the Naïve type on a relative basis (diff. = 0.0005; t-statistic = 4.25). However, the difference between the representative trade of each type in terms of *EarnedCRET* is much smaller than in terms of *CRET*. The representative trades are plotted in Figure 2. Considering the timeliness of each type's trades relative to the speed of the price adjustment yields opposite effects by type. The Savvy type have a lower mean return, while the Naïve type have a higher mean return.

4.2.2. Stacked Specifications

Due to the differing dependent variables across specifications when estimating Eq. 2, there is no way to examine the "full information" model that includes the trading behavior of both types. An alternative approach is to stack the dataset, allowing each type to have a separately computed *EarnedCRET* dependent variable and *Retail Trade IB* independent variable for each Firm–Day observation. I add an indicator variable, *Savvy,* that is equal to 1 when the Firm–Day observation is for the Savvy type, and equal to 0 when the Firm–Day observation is a Naïve type. This allows *EarnedCRET* to differ between Firm–Days in the same model.[36] This alternative stacked approach is estimated using a modified Eq. 2 as follows:

[35] A limitation of my study – and others relying on TAQ data – is that only judgments yielding a "trade" decision are observed. The "not-trade" decision is unobservable and is the correct timeliness judgment in some instances. Online Appendix 1.6 provides preliminary evidence on the "non-traders."

[36] In other words, there are two observations for each Firm–Day. Doubling observations with the same Firms and Dates does not impact the t-statistic calculation due to standard error clustering.

$$\begin{aligned} EarnedCRET_{i,t(0,4),j} = \; & \beta_1 Savvy_{i,t,j} + \beta_2 Imbalance_{i,t,j} + \beta_3 Savvy_{i,t,j} \times Imbalance_{i,t,j} \\ & + \beta_4 Firm\ Disclosure\ Day_{i,t} + \beta_5 Savvy_{i,t,j} \times Firm\ Disclosure\ Day_{i,t} \\ & + \beta_6 Firm\ Disclosure\ Day_{i,t} \times Imbalance_{i,t,j} \\ & + \beta_7 Savvy_{i,t,j} \times Imbalance_{i,t,j} \times Firm\ Disclosure\ Day_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4),j} \end{aligned} \tag{3}$$

Column (1) provides a baseline across all observations and regresses *EarnedCRET* on only *Savvy, Imbalance,* and their interaction. The Naïve type imbalance (the main effect on *Imbalance*) is negatively and significantly associated with the earned return (β_2 = -0.0008; t-statistic -13.10). The association between the Savvy type imbalance and the earned return (the linear combination of β_2 and β_3) is positive but insignificant ($\beta_2 + \beta_3$ = 0.0001; t-statistic 1.25; untabulated). Column (2) tests the full model using the stacked construction. The Naïve type's trade imbalance and the earned return are more negatively associated on firm disclosure days (β_6 = -0.0008; t-stat -6.03). The trade imbalance of the Savvy type has no significant incremental association with returns on firm disclosure days ($\beta_6 + \beta_7$ = -0.0001; t-statistic -1.30; untabulated). The triple interaction term (β_7) indicates the difference between the Naïve and Savvy imbalance-to-returns relationship widens significantly on firm disclosure days; a divergence almost entirely driven by the worsening performance of the Naïve type. The Savvy type clearly outperforms the Naïve type after considering the timeliness of their trades, but Savvy trading activity has an association with *EarnedCRET* of very near 0.

The results in this section suggest that by the time a trader has processed the information in the news and made a trading decision, much of the price adjustment appears to have already occurred. This provides some price protection to the Naïve type and attenuates the Savvy's positive performance to an association not significantly different than 0.

4.3. Costliness

Even if an individual investor correctly extracts the sign of the news and makes a correct inference about the timeliness of their trade, they must still account for transaction costs. To examine the extent of heterogeneity across types after considering the cost of the bid-ask spread, I construct a new dependent variable, *AfterCostCRET.* I separately compute the return to each type's buy trades (which are executed at the ask price) and sell trades (which are executed at the bid price). I combine these post-spread returns using the same procedure used in Sections 4.1 and 4.2 to compute representative trades for each type. Specifically, *AfterCostCRET* equals:

$$\left(\left(\frac{\textit{Buy Trades}_{i,t,j}}{\textit{Total Trades}_{i,t,j}}\right) \times \textit{BuyTradeReturn}_{i,t,j}\right) - \left(\left(\frac{\textit{Sell Trades}_{i,t,j}}{\textit{Total Trades}_{i,t,j}}\right) \times \textit{SellTradeReturn}_{i,t,j}\right)$$

Consistent with the construction of the prior dependent variables, I sum the after-cost return on $t = 0$ with the returns from $t = 1$ through $t = 4$. A numerical example of this construction is available in the Online Appendix 1.5.

Because returns to buys and sells are now calculated separately, I can no longer use a trade imbalance independent variable to determine the "correctness" of trading behavior (i.e., the buy-sell weighting is now implicit in the dependent variable). Instead, I use a stacked dataset construction (as in Section 4.2) to regress *AfterCostCRET* on indicator variables for *Savvy* and *Firm Disclosure Day* as follows:

$$\begin{aligned} \textit{AfterCostCRET}_{i,t(0,4),j} = \quad & \beta_1 \textit{Savvy}_{i,t,j} + \beta_2 \textit{Firm Disclosure Day}_{i,t} \\ & + \beta_3 \textit{Savvy}_{i,t,j} \times \textit{Firm Disclosure Day}_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4),j} \end{aligned} \quad (4)$$

Table 4A presents results. Column (1) includes only the *Savvy* indicator and firm and date fixed effects, and provides the baseline difference in means between Savvy and Naive *AfterCostCRET*. The return difference between types persists after considering the costliness of their trades, but is smaller in magnitude (β_1 = 0.0004; t-statistic 4.13). Column (2) presents the full model from Eq. 4. The indicator variable *Firm Disclosure Day* captures the decline in *AfterCostCRET* for the Naïve type on disclosure days relative to their baseline (β_2 = -0.0002; t-statistic -3.53). The insignificant interaction term (β_3 = 0.0000; t-statistic 0.75) indicates the difference between Savvy and Naïve after-cost performance is no wider on disclosure days.

Table 4B presents the representative trade difference-in-means test analogous with Tables 2B and 3B. The level of the return for both types is significantly negative after considering the cost of their trades. The representative trade of the Savvy type is significantly less negative than the Naïve type (diff. = 0.0004; t-statistic 3.94). While the Savvy type are relatively better off than the Naïve type, they still earn a negative return after considering the costliness of their trades.

4.4. Ancillary Analyses and Robustness Tests

4.4.1. Earnings Announcement Days

To examine differential processing of verifiable, value relevant financial disclosures, I reperform my main tests using the subset of disclosure days that are EA days. I modify Eq. 1:

$$CRET_{i,t(0,4)} = \beta_1 Retail\ Trade\ IB_{i,t,j} + \beta_2\ EA\ Day_{i,t} + \beta_3 Retail\ Trade\ IB_{i,t,j} \times EA\ Day_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4)} \tag{5}$$

Table 5A presents results of Eq. 5 with *EA Day* as the independent indicator variable. In columns (1) and (2), I find consistent, though more extreme, results to those in Table 2A. Column (3) presents the full model for EA days. *Naïve IB* is negatively associated with *CRET*, while *Savvy IB* is consistent with the sign of *CRET*. The association between each type's trade imbalance and the return becomes incrementally stronger on EA days, as well as when conditioning on the trade imbalance of the other type.

Table 5B replicates the full model specifications from Tables 3C and 4A. Column (1) examines *EarnedCRET* and finds directionally consistent but more extreme differentials in the association between Savvy and Naïve imbalance and returns on EA days relative to the corresponding specification in Table 3C.[37] Column (2) examines *AfterCostCRET* and finds a different result than the corresponding test in Table 4A; when limiting the firm disclosure to only

[37] Figure 4B provides descriptive evidence on the evolution of the absolute return and trades of each type on the EA day. The gap between the speed of the price adjustment and trading is even wider than on disclosure days overall.

EAs, the Savvy type performs statistically significantly better when new information is released.

Table 5C illustrates the representative trades of each type on only EA days. Inferences are similar to the full disclosure day sample (though differences between types are wider), and the overall pattern from Figure 2 remains. Collectively, the results in Table 5 suggest that the differential between Savvy and Naïve trading behavior in response to disclosure may be stronger when the disclosure is planned, relatively standardized, and focused on accounting information.

4.4.2. Cross Sectional Differences in AfterCostCRET

If the Savvy type is processing and trading on information while the Naïve type is largely reacting to attention signals, it should be the case that the Savvy type's relative outperformance is wider in less covered, more difficult to trade firms. I test this prediction by examining the differential relationship between Savvy and Naïve *AfterCostCRET* for firms above and below median size, trading volume, and spread. I use Eq. 4, but replace the indicator *Firm Disclosure Day* with indicators for *Big Firm*, *High Volume*, and *High Spread*, which each equal 1 for observations above the median of the respective metric and 0 for observations below.[38]

Results are reported in Table 6A. Column (1) indicates that Naïve traders have relatively higher *AfterCostCRET* in larger firms (β_2 = 0.0008; t-statistic 10.95), while the Savvy type's relative outperformance declines (β_3 = -0.0011; t-statistic -14.21).[39] A similar result is found in column (2) when examining trading volume. Column (3) has opposite signs, but consistent inferences; in more difficult to trade firms, the Savvy type's relative outperformance grows. In short, the Savvy type performs relatively better in smaller, less liquid, and more expensive to trade firms, while the Naive type has the opposite relative performance changes.

[38] Median splits are identified by firm based on quarterly averages of market value of equity, trading volume, and quoted spreads relative to prices.
[39] The β_3 interaction term provides stronger evidence of the relative performance change between the types than the main effect on β_2 given minimal within-fixed-effect variation in *Big Firm* (and *High Volume* and *High Spread*).

4.4.3. Timeliness vs. Costliness

There is an inherent tension between the timeliness and costliness of a trade. The longer a trader waits to act on new information, the less of the return reaction they can capture. On the other hand, spreads widen significantly around planned disclosures such as earnings announcements, meaning those who demand immediacy in their trading pay heightened costs to trade. Ex-ante, it seems the Savvy type may benefit more from timely trades while the Naïve type may benefit from less costly trades.

To examine this question, I compare the *AfterCostCRET* each type earns on firm disclosure days to that earned on the day immediately following. I introduce a new indicator variable, *Next Day*, which is coded as 1 for the Firm–Day immediately following a firm disclosure day, and 0 for the firm disclosure day itself. I limit the sample to only firm disclosure days and the immediately subsequent days. I regress *AfterCostCRET* on:

$$AfterCostCRET_{i,t(0,4),j} = \beta_1 Next\ Day_{i,t} + \beta_2 Savvy_{i,t,j} + \beta_3 Next\ Day_{i,t} \times Savvy_{i,t,j} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4),j} \qquad (6)$$

Column (1) of Table 7 retains only the indicator for *Next Day* and the firm and date fixed effects. Without differentiating between Naïve and Savvy traders, the main effect of waiting to trade until the day after a firm disclosure is estimated to be a positive and significant increase in *AfterCostCRET,* suggesting the cost factor matters more than the timing factor. However, the results in column (2) of Table 7 indicate a differential effect by type. The interaction of *Savvy* and *Next Day* is negative and significant. The effect is that the Savvy do not earn significantly higher *AfterCostCRET* on the day following a disclosure day ($\beta_2 + \beta_3 = 0.0002$; t-statistic 1.54; untabulated). This suggests the Naive type is relatively better off by waiting for lower costs, while the Savvy type is not.

4.4.4. Extreme Return Attention vs. Accounting Information based trading

Prior literature identifies attention-induced trading as a driver of retail investor underperformance. As firm disclosure days (particularly EA days) tend to have more extreme returns than typical trading days, some of the trading behaviors I observe may merely be a response to extreme price movement rather than a reaction to new, value-relevant information.

I first provide descriptive evidence that the preference for extreme returns documented in prior literature differs between the Savvy and Naïve types. Figure 5 plots average retail imbalance by type for Firm–Days with returns ranging from -20% to +20%. The Naïve trade imbalance is starkly V-shaped and centered on zero (similar to Barber & Odean 2008). As returns become more extreme – either positive or negative – the Naïve type buys more intensely. However, the Savvy type's trade imbalance changes in a more linear fashion. They buy (sell) more intensely when returns are above (below) zero.

I next conduct a test similar to that in Taylor (2010). I match each firm disclosure day and EA day observation with a non-disclosure or non-EA observation from the same firm in the same year with a similar return (+/- 1%). I then examine abnormal trading volume response:

$$AbVolume_{i,t,j} = \beta_1 Firm\ Disclosure\ Day_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t,j} \quad (7)$$

If retail investors are trading on disclosed information rather than simply trading on extreme returns, trading volume will be higher for firm disclosure days and EA days. Column (1) and (2) present results for Naïve and Savvy *AbVolume* for firm disclosure days relative to matched non-disclosure days. Both types have significantly more trading activity on disclosure days with similarly extreme returns. Columns (3) and (4) repeat the same test with *EA Day* as the indicator variable and a new set of matched observations. The response difference between the Naïve and Savvy is much wider, suggesting the Savvy type prefer trading on verifiable, value relevant financial information. The *EA Day* indicator also provides more explanatory power in the Savvy model than in the Naïve model (i.e., the adjusted R-squared value increases from 0.119 to 0.299).

Table 8B replicates the tests from columns 1 – 3 of Table 2A using the return-size-matched sample. I find largely consistent results, though the incremental negative association between the Naïve imbalance and CRET is only marginally significant. I conclude that my findings are not simply an artifact of differentially sized returns on disclosure days.

5. CONCLUSION

Prior literature largely treats retail investors as a homogeneous group. I investigate differences in the trading behavior of different retail investor types; specifically, around their ability to extract the sign of the news, judge the timeliness of their trade, and account for transaction costs. I find that the least capitalized, financially sophisticated, and experienced retail traders (the Naïve type) make trades in the opposite direction of the contemporaneous return, and that this tendency is significantly stronger around firm disclosures. The Savvy type, on the other hand, trade consistently with the sign of the news, suggesting they successfully process firm disclosures and value relevant financial information. These results are attenuated towards zero after considering the timeliness of each type's trading. Once transaction costs are factored in, both types tend to earn a negative return, though the Savvy type performs better than the Naïve type, particularly in less visible, more difficult to trade firms where they are better able to exploit an information advantage.

Understanding differences in retail investor attributes and behaviors is increasingly important as they comprise an ever-growing share of market activity. I contribute to the academic literature on retail investment by highlighting both the heterogeneity of this investor class and the importance of considering more than just the direction of their trades when evaluating informedness. I provide a simple heuristic for identifying different retail investor types in publicly available data. I also contribute to the regulatory discussion on investor protections.[40] For the least sophisticated individual investors, a focus on enhancing liquidity and lowering trading costs could shield them from the negative effects of behavioral biases while still encouraging participation in capital markets.

[40] Regulators have made some investor protection policies based on wealth. For example, FINRA's "Pattern Day trader rule" prohibits frequent day trading by individuals with less than $25,000 in their account. As another illustration, SEC's Regulation D limits unregistered offerings to accredited investors who are "financially sophisticated and able to fend for themselves… thus rendering less necessary the protections that come from a registered offering."

References

Allee, K. D., N. Bhattacharya, E. L. Black, and T. E. Christensen. 2007. Pro forma disclosure and investor sophistication: External validation of experimental evidence using archival data. *Accounting, Organizations and Society*.

Amiram, D., E. Owens, and O. Rozenbaum. 2016. Do information releases increase or decrease information asymmetry? New evidence from analyst forecast announcements. *Journal of Accounting and Economics*.

Asthana, S., S. Balsam, and S. Sankaraguruswamy. 2004. Differential Response of Small versus Large Investors to 10-K Filings on EDGAR. *The Accounting Review*.

Avakian, S. 2020. SEC.gov | Protecting Everyday Investors and Preserving Market Integrity: The SEC's Division of Enforcement.

Balasubramaniam, V., J. Y. Campbell, T. Ramadorai, and B. Ranish. 2023. Who Owns What? A Factor Model for Direct Stockholding. *The Journal of Finance*.

Barber, B. M., X. Huang, P. Jorion, T. Odean, and C. Schwarz. 2024. A (Sub)penny for Your Thoughts: Tracking Retail Investor Activity in TAQ. *The Journal of Finance*.

Barber, B. M., X. Huang, T. Odean, and C. Schwarz. 2022. Attention-Induced Trading and Returns: Evidence from Robinhood Users. *The Journal of Finance*.

Barber, B. M., Y.-T. Lee, Y.-J. Liu, and T. Odean. 2009a. Just How Much Do Individual Investors Lose by Trading? *The Review of Financial Studies*.

Barber, B. M., S. Lin, and T. Odean. 2023. Resolving a Paradox: Retail Trades Positively Predict Returns but Are Not Profitable. *Journal of Financial and Quantitative Analysis*.

Barber, B. M., and T. Odean. 2000. Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors. *The Journal of Finance*.

——— 2008. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *The Review of Financial Studies*.

——— 2013. Chapter 22 - The Behavior of Individual Investors. In *Handbook of the Economics of Finance*, edited by G. M. Constantinides, M. Harris, and R. M. Stulz.

Barber, B. M., T. Odean, and N. Zhu. 2009b. Do Retail Trades Move Markets? *The Review of Financial Studies*.

Barrot, J.-N., R. Kaniel, and D. Sraer. 2016. Are retail traders compensated for providing liquidity? *Journal of Financial Economics.*

Bartlett, R. P., J. McCrary, and M. O'Hara. 2024. Tiny trades, big questions: Fractional shares. *Journal of Financial Economics*.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings1. *Journal of Accounting and Economics*.

Battalio, R. H., R. H. Jennings, M. Saglam, and J. Wu. 2023. Difficulties in obtaining a representative sample of retail trades from public data sources. SSRN Scholarly Paper.

Battalio, R. H., and R. R. Mendenhall. 2005. Earnings expectations, investor trade size, and anomalous returns around earnings announcements. *Journal of Financial Economics*.

Beckmeyer, H., N. Branger, and L. Gayda. 2024. Retail Traders Love 0DTE Options... But Should They? SSRN Scholarly Paper.

Betermier, S., L. E. Calvet, S. Knüpfer, and J. S. Kvaerner. 2025. Investor Factors. *The Journal of Finance*.

Bhattacharya, N. 2001. Investors' Trade Size and Trading Responses around Earnings Announcements: An Empirical Investigation. *The Accounting Review*.

Blankespoor, E., E. deHaan, and I. Marinovic. 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*.

Blankespoor, E., E. deHaan, J. Wertz, and C. Zhu. 2019. Why Do Individual Investors Disregard Accounting Information? The Roles of Information Awareness and Acquisition Costs. *Journal of Accounting Research*.

Blankespoor, E., E. deHaan, and C. Zhu. 2018. Capital market effects of media synthesis and dissemination: evidence from robo-journalism. *Review of Accounting Studies*.

Boehmer, E., C. M. Jones, Xiaoyan Zhang, and Xinran Zhang. 2021. Tracking Retail Investor Activity. *The Journal of Finance*.

Bogousslavsky, V., and D. Muravyev. 2025. An Anatomy of Retail Option Trading. SSRN Scholarly Paper.

Bushee, B. J. 1998. The Influence of Institutional Investors on Myopic R&D Investment Behavior. *The Accounting Review*.

Calvet, L. E., J. Y. Campbell, and P. Sodini. 2009. Measuring the Financial Sophistication of Households. *American Economic Review*.

Campbell, J. Y., T. Ramadorai, and A. Schwartz. 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics*.

Chau, J. 2025. Accounting Information Usage and Trading by Retail Investors: Evidence from Integrated Trading Platform. *Journal of Accounting Research*.

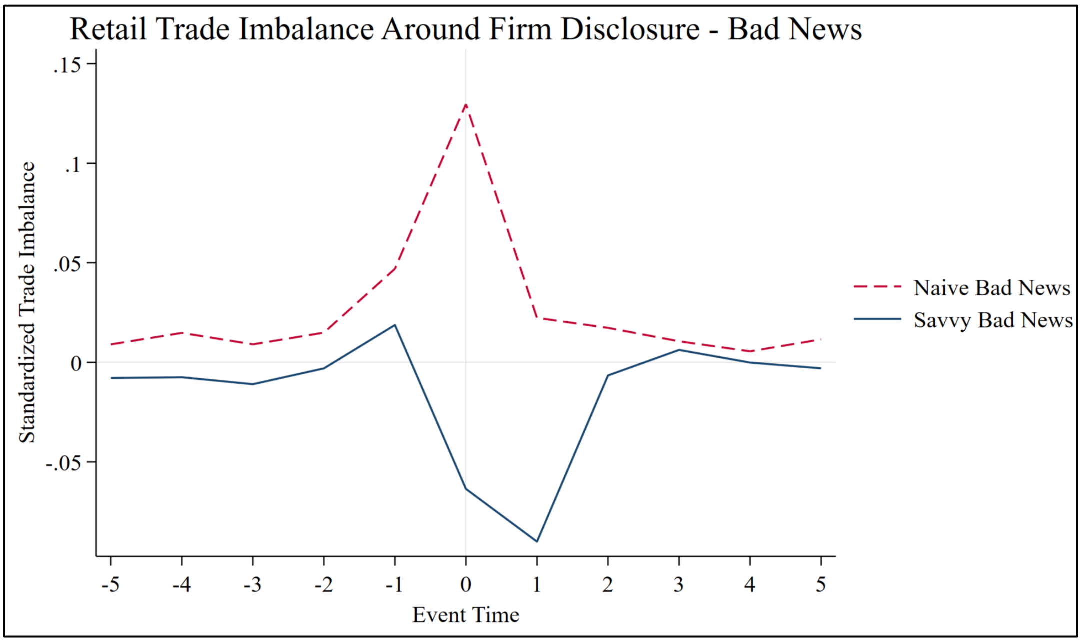
Chen, H.-L., E. H. Chow, and C.-Y. Shiu. 2015. The informational role of individual investors in stock pricing: Evidence from large individual and small retail investors. *Pacific-Basin Finance Journal*.
Chordia, T., R. Roll, and A. Subrahmanyam. 2002. Order imbalance, liquidity, and market returns. *Journal of Financial Economics*.
Cianciaruso, D., I. Marinovic, and K. Smith. 2023. Asymmetric Disclosure, Noise Trade, and Firm Valuation. *Accounting Review*.
Coval, J. D., D. Hirshleifer, and T. Shumway. 2021. Can Individual Investors Beat the Market? *The Review of Asset Pricing Studies*.
Coval, J. D., and T. Shumway. 2005. Do Behavioral Biases Affect Prices? *The Journal of Finance*.
Cready, W. M. 1985. An empirical investigation into the relationship between the value of accounting earnings announcements and equity investor endowment sizes. Ph.D. Dissertation.
——— 1988. Information Value and Investor Wealth: The Case of Earnings Announcements. *Journal of Accounting Research*.
Cready, W., A. Kumas, and M. Subasi. 2014. Are Trade Size-Based Inferences About Traders Reliable? Evidence from Institutional Earnings-Related Trading. *Journal of Accounting Research*.
Da, Z., V. W. Fang, and W. Lin. 2024. Fractional Trading. *The Review of Financial Studies*.
de Silva, T., K. Smith, and E. C. So. 2025. Losing is Optional: Retail Option Trading and Expected Announcement Volatility. SSRN Scholarly Paper.
deHaan, E., and A. Glover. 2024. Market Access and Retail Investment Performance. *The Accounting Review*.
deHaan, E., J. Li, and E. M. Watts. 2023. Retail bond investors and credit ratings. *Journal of Accounting and Economics*.
deHaan, E., T. Shevlin, and J. Thornock. 2015. Market (in)attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics*.
Dorn, D., and P. Sengmueller. 2009. Trading as Entertainment? *Management Science*.
Drake, M. S., K. H. Gee, and J. R. Thornock. 2016. March Market Madness: The Impact of Value-Irrelevant Events on the Market Pricing of Earnings News. *Contemporary Accounting Research.*
Du, J., D. Huang, Y.-J. Liu, Y. Shi, A. Subrahmanyam, and H. Zhang. 2026. Nominal Prices, Retail Investor Participation, and Return Momentum. *Management Science.*
Eaton, G. W., T. C. Green, B. S. Roseman, and Y. Wu. 2022. Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics*.
Engelberg, J. E., and C. A. Parsons. 2011. The Causal Impact of Media in Financial Markets. *The Journal of Finance*.
Even-Tov, O., K. Munevar, S. Kogan, and E. C. So. 2026. No Fees to Play: Trading Fees, Gamification, and Retail Investor Behavior. SSRN Scholarly Paper.
Fong, K. Y. L., D. R. Gallagher, and A. D. Lee. 2014. Individual Investors and Broker Types. *Journal of Financial and Quantitative Analysis*.
Frazzini, A., and O. A. Lamont. 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*.
Friedman, D. M. 1953. A methodology of Positive Economics. *The Philosophy of Economics: An Anthology*. Cambridge University Press.
Friedman, H. L., and Z. Zeng. 2024. Retail Investor Trading and Market Reactions to Earnings Announcements. SSRN Scholarly Paper.
Gensler, G. 2022. SEC.gov | "Market Structure and the Retail Investor:" Remarks Before the Piper Sandler Global Exchange Conference.
Gomez, E. A., F. Heflin, J. R. Moon Jr., and J. D. Warren. 2024. Financial Analysis on Social Media and Disclosure Processing Costs: Evidence from Seeking Alpha. *The Accounting Review*.
Grinblatt, M., and M. Keloharju. 2000. The investment behavior and performance of various investor types: a study of Finland's unique data set. *Journal of Financial Economics*.
——— 2009. Sensation Seeking, Overconfidence, and Trading Activity. *The Journal of Finance*.
Grossman, S. J., and J. E. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *The American Economic Review*.
Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to Distraction: Extraneous Events and Underreaction to Earnings News. *The Journal of Finance*.
Hirshleifer, D. A., J. N. Myers, L. A. Myers, and S. H. Teoh. 2008. Do individual investors cause post-earnings announcement drift? Direct evidence from personal trades. *The Accounting Review*.
Hu, J., A. Kirilova, S. (Gilbert) Park, and D. Ryu. 2024. Who Profits from Trading Options? *Management Science*.
Hvidkjaer, S. 2008. Small Trades and the Cross-Section of Stock Returns. *The Review of Financial Studies*.
Hwang, Y., J. Park, J. H. Kim, Y. Lee, and F. J. Fabozzi. 2024. Heterogeneous trading behaviors of individual investors: A deep clustering approach. *Finance Research Letters.*
Israeli, D., R. Kasznik, and S. A. Sridharan. 2022. Unexpected distractions and investor attention to corporate announcements. *Review of Accounting Studies*.

Jones, C. M., D. Shi, Xiaoyan Zhang, and Xinran Zhang. 2024. Retail Trading and Return Predictability in China. *Journal of Financial and Quantitative Analysis*.
Kalay, A. 2015. Investor sophistication and disclosure clienteles. *Review of Accounting Studies*.
Kaniel, R., S. Liu, G. Saar, and S. Titman. 2012. Individual Investor Trading and Return Patterns around Earnings Announcements. *The Journal of Finance*.
Kaniel, R., G. Saar, and S. Titman. 2008. Individual Investor Trading and Stock Returns. *The Journal of Finance*.
Kelley, E. K., and P. C. Tetlock. 2013. How Wise Are Crowds? Insights from Retail Orders and Stock Returns. *The Journal of Finance*.
——— 2017. Retail Short Selling and Stock Prices. *The Review of Financial Studies.*
Kim, O., and R. E. Verrecchia. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*.
Kumar, A., and C. M. c. Lee. 2006. Retail Investor Sentiment and Return Comovements. *The Journal of Finance*.
Laarits, T., and M. Sammon. 2025. The retail habitat. *Journal of Financial Economics*
Lawrence, A. 2013. Individual investors and financial disclosure. *Journal of Accounting and Economics*.
Lee, C. M. C. 1992. Earnings news and small traders: An intraday analysis. *Journal of Accounting and Economics*.
Lee, C. M. C., B. Mucklow, and M. J. Ready. 1993. Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis. *The Review of Financial Studies*.
Lee, C. M. C., and B. Radhakrishna. 2000. Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets*.
Lee, C. M. C., and M. J. Ready. 1991. Inferring Trade Direction from Intraday Data. *The Journal of Finance.*
Leuz, C., S. Meyer, M. Muhn, E. Soltes, and A. Hackethal. 2025. Who Falls Prey to the Wolf of Wall Street? Investor Participation in Market Manipulation. *Management Science*.
Li, Q., E. M. Watts, and C. Zhu. 2024. Retail investors and ESG news. *Journal of Accounting and Economics*.
Li, Y., S. Sokolinski, and A. Tamoni. 2026. Which investors drive anomaly returns and how? *Journal of Financial Economics*.
Lin, W. 2024. Strategic Informed Trades and Tiny Trades. Ph.D. Dissertation.
Liu, B. Y. 2022. Retail Investors and the Pricing of Accounting Information: Evidence from Brokerage Website Outages. Ph.D. Dissertation.
Luo, C., E. Ravina, M. Sammon, and L. M. Viceira. 2025. Retail Investors' Contrarian Behavior Around News, Attention, and the Momentum Effect. SSRN Scholarly Paper.
Malmendier, U., and D. Shanthikumar. 2007. Are small investors naive about incentives? *Journal of Financial Economics*.
Michels, J. 2024. Retail investor trade and the pricing of earnings. *Review of Accounting Studies*.
Mikhail, M. B., B. R. Walther, and R. H. Willis. 2007. When Security Analysts Talk, Who Listens? *The Accounting Review*.
Miller, B. P. 2010. The Effects of Reporting Complexity on Small and Large Investor Trading. *The Accounting Review*.
Moss, A., J. P. Naughton, and C. Wang. 2024. The Irrelevance of Environmental, Social, and Governance Disclosure to Retail Investors. *Management Science*.
Nofsinger, J. R. 2001. The impact of public information on investors. *Journal of Banking & Finance*.
Odean, T. 1999. Do Investors Trade Too Much? *American Economic Review*.
O'hara, M., C. Yao, and M. Ye. 2014. What's Not There: Odd Lots and Market Data. *The Journal of Finance*.
Ohlson, J. A. 1975. The Complete Ordering of Information Alternatives for a Class of Portfolio-Selection Models. *Journal of Accounting Research*.
Patell, J. M., and M. A. Wolfson. 1984. The intraday speed of adjustment of stock prices to earnings and dividend announcements. *Journal of Financial Economics*.
Shanthikumar, D. M. 2004. Small and large trader behavior: Reactions to information in financial markets. Ph.D. Dissertation.
Taylor, D. 2010. Individual investors and corporate earnings. Ph.D. Dissertation.
Twedt, B. 2016. Spreading the Word: Price Discovery and Newswire Dissemination of Management Earnings Guidance. *The Accounting Review*.
Welch, I. 2022. The Wisdom of the Robinhood Crowd. *The Journal of Finance*.
Wertz, J. L. 2021. The Speed of Price Responses to Individual Signals in a Bundle. Ph.D. Dissertation.
Xie, C. 2025. Informed Trade of Earnings Announcements. *Journal of Accounting Research.*

Figure 1: Retail Trade Imbalance and Firm Disclosure

Figure 1 illustrates retail trade imbalance by type of trader and type of news in event time. Day $t = 0$ is defined as any day with a firm press release or earnings announcement. Trade Imbalance and Investor Types are defined in Appendix A. The type of news is defined based on the adjusted return from CRSP on $t = 0$. If the firm return less the value-weighted market return is positive (negative), the firm disclosure is deemed to be good (bad) news. Any overlapping event-time observations (e.g., press releases within +/- 6 days of one other) are excluded. Imbalances are standardized within firm-year for this figure.

Panel A: Standardized Trade Imbalance – Bad News



Panel B: Standardized Trade Imbalance – Good news

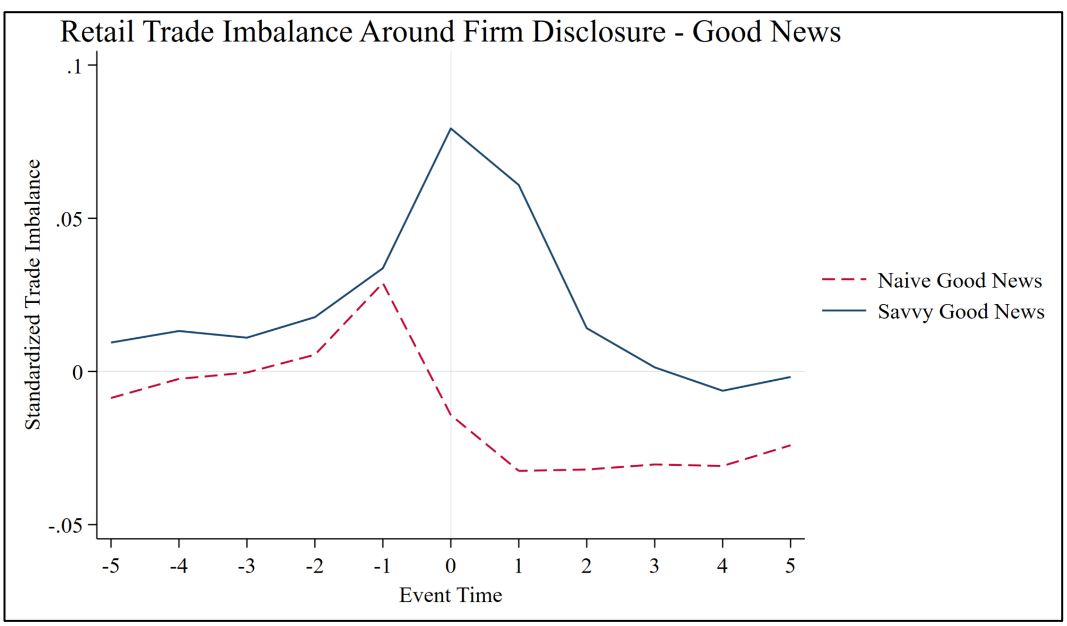


Figure 2: Returns to the Representative Retail Trade

Figure 2 provides a summary of the average return to a "representative trade" of each type on a firm disclosure day for each of the three main dependent variables. Variables are defined in Appendix A.

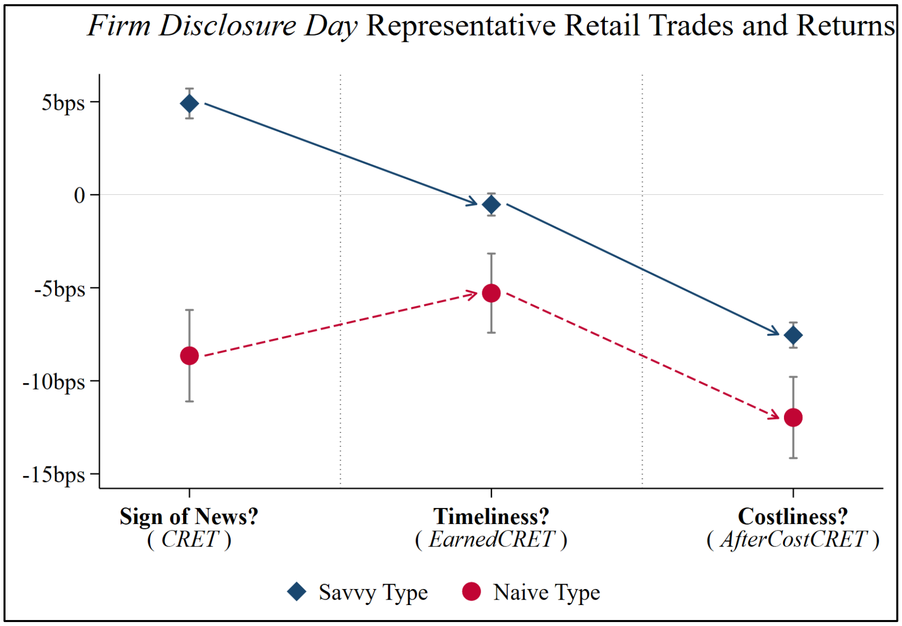


Figure 3: Changing Participation

Figure 3 illustrates changes in the makeup of retail investor trades over time. The figure plots the average per-day, per-ticker count of trades for each type included in the sample. Retail trades are identified in TAQ using the Boehmer et al. (2021) method. Types are defined in Appendix A.

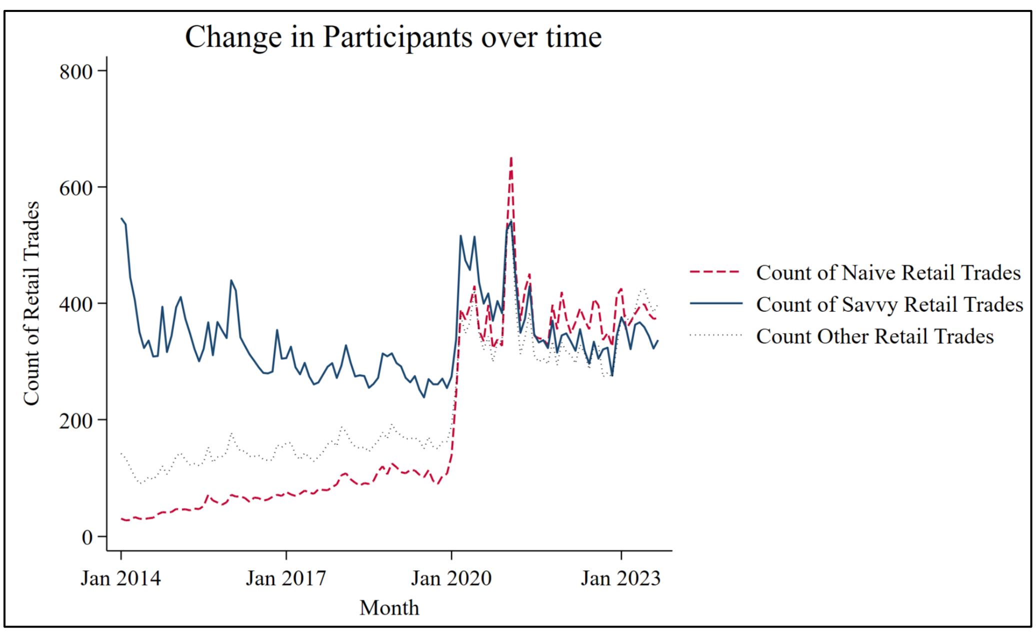
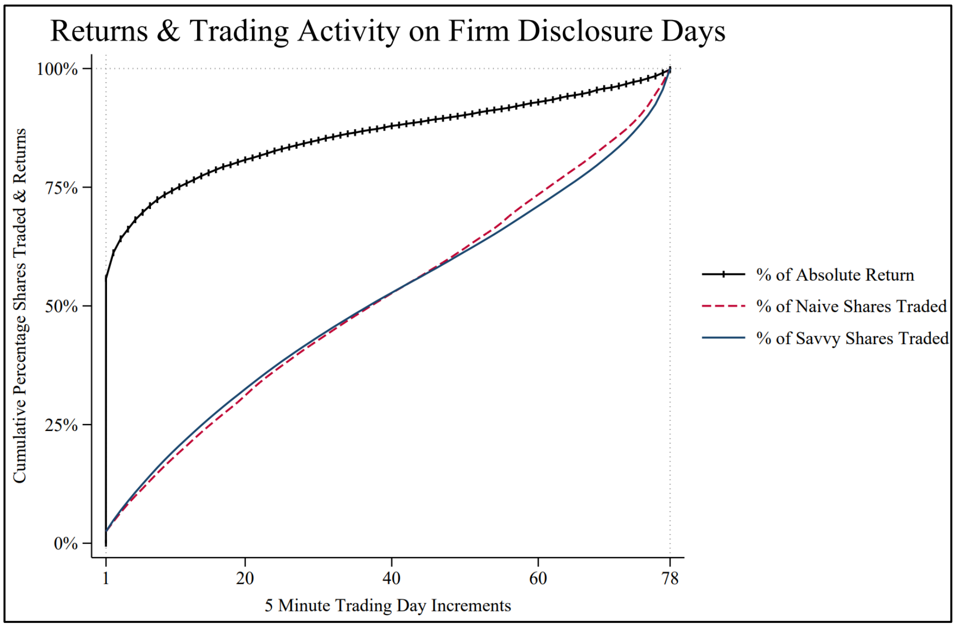


Figure 4: Cumulative Distribution of Returns and Trades

Figure 4 illustrates the typical trading patterns and return progression on firm disclosure days (Panel A) and EA days (Panel B). Shown are the empirical CDFs of both the absolute return and trades from the Naïve and Savvy type. Returns are drawn from CRSP. The return CDF plots the percentage of the eventual, full day return reached during each 5-minute increment of the trading day, winsorized at 1% and 99%. Days with returns less than |1%| are excluded (as in tests of Intra Period Timeliness; see Blankespoor et al. 2020). The trading activity CDFs plot the percentage of the total daily shares traded reached during each 5-minute increment. Investor types are defined in Appendix A.

Panel A: Disclosure Days



Panel B: EA Days

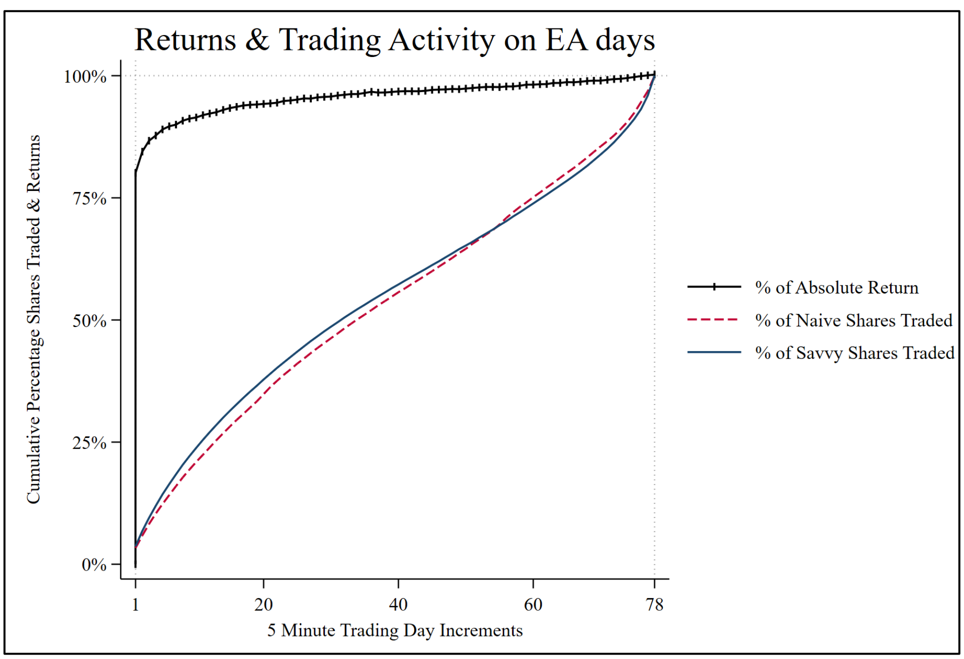


Figure 5: Retail Trade Imbalance and Extreme Returns

Figure 5 illustrates average retail trade imbalance by type and extremity of daily CRSP returns. Returns are rounded to the nearest 0.01 before taking the mean trade imbalance of each type. Trade Imbalance and type are defined in Appendix A.

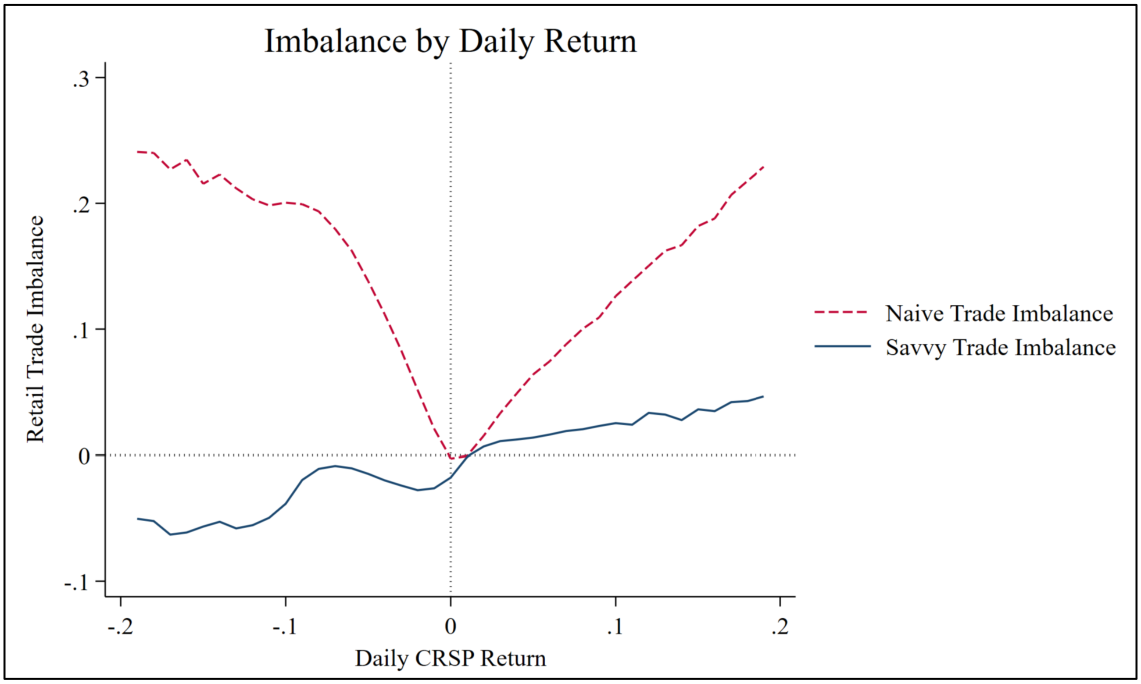


Table 1: Descriptive Statistics

This table presents summary statistics for the primary sample. The unit of observation is Firm–Day. Variables are defined in Appendix A.

	N	Mean	Std. Dev	P25	Median	P75
CRET	5,430,490	0.002	0.086	-0.031	0.002	0.033
Firm Disclosure Day	5,430,490	0.116	0.321	0.000	0.000	0.000
EA Day	5,430,490	0.017	0.127	0.000	0.000	0.000
Naïve IB	5,430,490	0.038	0.332	-0.175	0.044	0.250
Savvy IB	5,430,490	-0.011	0.254	-0.156	0.000	0.138
Total Naïve Trades	5,430,490	227	1888	15	40	122
Total Savvy Trades	5,430,490	339	1497	41	108	277
EarnedCRET – Naïve	5,430,490	0.000	0.077	-0.029	0.001	0.029
EarnedCRET – Savvy	5,430,490	0.001	0.077	-0.029	0.001	0.029
AfterCostCRET – Naïve	5,430,490	-0.001	0.028	-0.007	-0.001	0.005
AfterCostCRET – Savvy	5,430,490	-0.001	0.020	-0.005	-0.001	0.003
Relative Spread	5,430,490	0.004	0.006	0.001	0.002	0.004

Table 2: Firm Disclosures

This table presents an analysis of returns and trade imbalances around firm disclosures. The dependent variable in Panel A is the 5-day cumulative return from day $t = 0$ to $t = 4$. The indicator variable *Firm Disclosure Day* takes a value of 1 if the firm issued a press release or earnings announcement on that date (as reported in Ravenpack and IBES). Fixed effects are noted near the bottom of each specification. Panel B provides return means by type and tests the difference in means. Standard errors are clustered by Firm and Date, and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: CRET and Retail Trade Imbalance

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB	-0.0007***		-0.0010***
	(-9.04)		(-12.37)
Savvy IB		0.0024***	0.0025***
		(38.01)	(39.85)
Firm Disclosure Day	0.0021***	0.0020***	0.0021***
	(13.44)	(12.70)	(13.32)
Naïve IB × Firm Disclosure Day	-0.0023***		-0.0024***
	(-12.70)		(-13.41)
Savvy IB × Firm Disclosure Day		0.0008***	0.0011***
		(5.67)	(8.05)
Observations	5,430,434	5,430,434	5,430,434
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.168	0.168	0.169

Panel B: Representative trades: CRET

	μ_{Savvy} (1)	$\mu_{\text{Naïve}}$ (2)	$\mu_{\text{Savvy}} - \mu_{\text{Naïve}}$ (3)
Disclosure Day Averages	0.0005*** (12.00)	-0.0009*** (-6.90)	0.0014*** (10.50)

Table 3: Timeliness of Trades

This table presents analysis of earned returns and trade imbalances around firm disclosures. In Panel A, the earned return is computed separately for each investor type, yielding a different dependent variable in each specification. The indicator variable *Firm Disclosure Day* takes a value of 1 if the firm issued a press release or earnings announcement on that date (as reported in Ravenpack and IBES). Panel B provides return means by type and tests the difference in means. Panel C stacks the Savvy and Naïve observations and adds an indicator (*Savvy*) equal to 1 if the observation is of the Savvy type and 0 if the observation is of the Naïve type. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: EarnedCRET and Retail Trade Imbalance

	(1)	(2)
Dependent Variable =	*EarnedCRET – Naïve*	*EarnedCRET – Savvy*
Naïve IB	-0.0008***	
	(-11.74)	
Savvy IB		0.0001
		(1.63)
Firm Disclosure Day	-0.0005***	-0.0006***
	(-3.61)	(-4.33)
Naïve IB × Firm Disclosure Day	-0.0008***	
	(-6.12)	
Savvy IB × Firm Disclosure Day		-0.0001
		(-1.26)
Observations	5,430,434	5,430,434
Fixed Effects	Firm, Date	Firm, Date
Adj. R-squared	0.169	0.169

Panel B: Representative trades: EarnedCRET

	μ_{Savvy} (1)	$\mu_{\text{Naïve}}$ (2)	$\mu_{\text{Savvy}} - \mu_{\text{Naïve}}$ (3)
Disclosure Day Averages	-0.0001* (-1.72)	-0.0005*** (-4.87)	0.0005*** (4.25)

Panel C: Stacked construction – EarnedCRET

	(1)	(2)
Dependent Variable =	*EarnedCRET*	*EarnedCRET*
Savvy	0.0001***	0.0001***
	(8.81)	(10.19)
Imbalance	-0.0008***	-0.0007***
	(-13.10)	(-11.81)
Savvy × Imbalance	0.0009***	0.0008***
	(12.19)	(11.23)
Firm Disclosure Day		-0.0005***
		(-3.51)
Savvy × Firm Disclosure Day		-0.0001***
		(-9.58)
Firm Disclosure Day × Imbalance		-0.0008***
		(-6.03)
Savvy × Firm Disclosure Day × Imbalance		0.0006***
		(4.35)
Observations	10,860,980	10,860,980
Fixed Effects	Firm, Date	Firm, Date
Adj. R-squared	0.169	0.169

Table 4: After-Cost Returns

This table presents analysis of after-cost returns around firm disclosures Panel A stacks the Savvy and Naïve observations and adds an indicator (*Savvy*) equal to 1 if the observation is of the Savvy type and 0 if the observation is of the Naïve type. Panel B illustrates conditional means and tests of difference in means for *AfterCostCRET*. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Stacked construction – AfterCostCRET

Dependent Variable = *AfterCostCRET*	(1)	(2)
Savvy	0.0004***	0.0004***
	(4.13)	(4.08)
Firm Disclosure Day		-0.0002***
		(-3.54)
Savvy × Firm Disclosure Day		0.0000
		(0.75)
Observations	10,860,980	10,860,980
Fixed Effects	Firm, Date	Firm, Date
Adj. R-squared	0.011	0.011

Panel B: Representative trades: AfterCostCRET

	μ_{Savvy} (1)	$\mu_{\text{Naïve}}$ (2)	$\mu_{\text{Savvy}} - \mu_{\text{Naïve}}$ (3)
Disclosure Day Averages	-0.0008*** (-21.81)	-0.0012*** (-10.75)	0.0004*** (3.94)

Table 5: EA days

This table presents an analysis of returns and trade imbalances around EA days. The indicator variable *EA Day* takes a value of 1 if the firm announced earnings on day $t = 0$, or after hours on day $t - 1$. In Panel A, the dependent variable is the 5-day cumulative return from day $t = 0$ to $t = 4$. In Panel B, the dependent variables are *EarnedCRET* and *AfterCostCRET* as used in Tables 3 and 4. Panel C illustrates conditional means and tests of difference in means for *CRET*, *EarnedCRET*, and *AfterCostCRET* on EA days. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: CRET and Imbalance on EA days

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB	-0.0008***		-0.0010***
	(-9.32)		(-12.72)
Savvy IB		0.0024***	0.0025***
		(38.63)	(40.57)
EA Day	0.0020***	0.0004	0.0021***
	(3.97)	(0.74)	(4.15)
Naïve IB × EA Day	-0.0116***		-0.0120***
	(-20.20)		(-20.56)
Savvy IB × EA Day		0.0015**	0.0034***
		(2.52)	(5.73)
Observations	5,430,434	5,430,434	5,430,434
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.168	0.168	0.169

Panel B: Stacked Construction – EarnedCRET and AfterCostCRET

	(1)	(2)
Dependent Variable =	*EarnedCRET*	*AfterCostCRET*
Savvy	0.0001***	0.0004***
	(8.86)	(4.06)
Imbalance	-0.0008***	
	(-12.93)	
Savvy × Imbalance	0.0009***	
	(11.97)	
EA Day	0.0006	-0.0003*
	(1.49)	(-1.76)
Savvy × EA Day	-0.0001*	0.0005**
	(-1.92)	(2.41)
EA Day × Imbalance	-0.0007**	
	(-2.13)	
Savvy × EA Day × Imbalance	0.0012***	
	(2.63)	
Observations	10,860,980	10,860,980
Fixed Effects	Firm, Date	Firm, Date
Adj. R-squared	0.169	0.169

Panel C: Representative trades on EA Days

	μ_{Savvy} (1)	$\mu_{\text{Naïve}}$ (2)	$\mu_{\text{Savvy}} - \mu_{\text{Naïve}}$ (3)
CRET Averages	0.0007*** (5.56)	-0.0042*** (-13.43)	0.0049*** (15.00)
EarnedCRET Averages	0.0001* (1.68)	-0.0008*** (-3.04)	0.0009*** (3.43)
AfterCostCRET Averages	-0.0007*** (-10.20)	-0.0017*** (-6.63)	0.0009*** (3.68)

Table 6: Cross Sectional Differences

This table uses a stacked construction and includes indicator variables for Firm–Days above the median quarterly market value of equity (*Big Firm*), trading volume (*High Volume*), and relative spread (*High Spread*). Standard errors are clustered by Firm and Date and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Dependent Variable = *AfterCostCRET*	(1)	(2)	(3)
Savvy	0.0010***	0.0009***	-0.0002*
	(7.44)	(10.96)	(-1.89)
Big Firm	0.0008***		
	(10.95)		
High Volume		0.0011***	
		(11.74)	
High Spread			-0.0009***
			(-14.09)
Savvy × Big Firm	-0.0011***		
	(-14.21)		
Savvy × High Volume		-0.0009***	
		(-13.35)	
Savvy × High Spread			0.0012***
			(17.10)
Observations	10,860,980	10,860,980	10,860,980
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.011	0.011	0.011

Table 7: Timeliness vs. Costliness

This table presents results of regressing *AfterCostCRET* on indicators for the day following a firm disclosure day (*Next Day*) and for whether the observation is for a Savvy type (*Savvy*). The omitted category is the return on the firm disclosure day itself. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

	(1)	(2)
Dependent Variable =	*AfterCostCRET*	*AfterCostCRET*
Next Day	0.0002***	0.0003***
	(5.52)	(6.12)
Savvy		0.0004***
		(3.94)
Next Day × *Savvy*		-0.0003***
		(-4.67)
Observations	2,186,388	2,186,388
Fixed Effects	Firm, Date	Firm, Date
Adj. R-squared	0.018	0.018

Table 8: Extreme Returns vs. Accounting Information

This table presents results from the return-size-matched sample. Panel A uses abnormal volume measures as the dependent variables. Panel B uses *CRET,* and replicates results from Table 2 using the matched sample. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Abnormal Volume

	(1)	(2)	(3)	(4)
Dependent Variable =	*AbVolume – Naïve*	*AbVolume – Savvy*	*AbVolume – Naïve*	*AbVolume – Savvy*
Firm Disclosure Day	0.0775***	0.1953***		
	(22.33)	(46.44)		
EA Day			0.2603***	0.9050***
			(30.33)	(55.81)
Observations	1,174,264	1,174,264	133,810	133,810
Fixed Effects	Firm, Date	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.029	0.053	0.119	0.299

Panel B: Firm Disclosure Days with return-size-matched sample

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB	-0.0014***		-0.0017***
	(-7.93)		(-9.09)
Savvy IB		0.0019***	0.0020***
		(13.93)	(15.13)
Firm Disclosure Day	0.0002	0.0002	0.0002
	(0.36)	(0.29)	(0.33)
Naïve IB × Firm Disclosure Day	-0.0005*		-0.0006*
	(-1.70)		(-1.88)
Savvy IB × Firm Disclosure Day		0.0004**	0.0005**
		(2.14)	(2.53)
Observations	1,174,264	1,174,264	1,174,264
Fixed Effects	Firm, Date	Firm	Firm
Adj. R-squared	0.070	0.070	0.070

Appendix A: Variable Definitions

This table provides variable definitions and data sources. Variables are at the Firm–Day unless noted otherwise.

Variable	Description	Source
Total Naïve Trades	The count of all Boehmer et al. (2021) identified retail trades with <= 10 shares.	TAQ
Total Savvy Trades	The count of all Boehmer et al. (2021) identified retail trades with 99 < shares < 1,000.	TAQ
Naïve IB	Defined as (Total Naïve Buy trades – Total Naïve Sell trades) / *Total Naïve Trades*. For all tests, the variable is standardized by subtracting its mean and dividing by the standard deviation.	TAQ, Calculated
Savvy IB	Defined as (Total Savvy Buy trades – Total Savvy Sell trades) / *Total Savvy Trades.* For all tests, the variable is standardized by subtracting its mean and dividing by the standard deviation.	TAQ, Calculated
Firm Disclosure Day	Indicator equal to 1 if a firm issues a press release or earnings announcement on the date, 0 otherwise. For after-hours disclosures, the following trading day is used.	Ravenpack, IBES
EA Day	Indicator equal to 1 for an earnings announcement date, 0 otherwise. For after-hours announcements, the following trading day is used. A subset of *Firm Disclosure Day.*	IBES
Return	RET from CRSP, winsorized at .01% and 99.99% to account for extreme outliers	CRSP, Calculated
CRET	5-day cumulative return generated by summing Return across days $t = 0$ to $t = 4$	CRSP, Calculated
EarnedCRET	Adapted 5-day cumulative return, where the return on $t = 0$ is calculated as (CloseMidpoint$_t$ – AvgExecution$_{t,j}$)/AvgExecution$_{t,j}$ and is winsorized at .01% and 99.99%.	CRSP, TAQ, Calculated
Savvy	An indicator variable equal to 1 if the Firm–Day observation is for the Savvy type. Used only in the stacked dataset specifications.	Calculated
AfterCostCRET	A composite return generated by summing the return to buy trades and sell trades, weighted by the proportion of buys and sells on $t = 0$. Daily sell returns are computed as (-1)*Return. The $t = 0$ returns are calculated separately for buy and sell trades using the same procedure as for *EarnedCRET* and are winsorized at .01% and 99.99%.	CRSP, TAQ, Calculated
Next Day	An indicator variable equal to 1 if day *t* is the day immediately following a *Firm Disclosure Day*.	Calculated
AbVolume Naïve (Savvy)	Volume of firm *i* shares traded by the *Naïve* (*Savvy*) type on day t =0, less the mean number of shares traded per day for firm *i* by the *Naïve* (*Savvy*) type during the year including day $t = 0$, divided by the same mean shares.	CRSP
Relative Spread	The average quoted bid – ask spread across all trades in firm *i* on day t =0 divided by the nominal price at the end of day $t = 0$.	TAQ, CRSP
Big Firm (High Volume) (High Spread)	An indicator variable equal to 1 if the observation falls in above the median of the partitioning attribute, either average quarterly market value of equity, trading volume, or relative spread.	TAQ, CRSP

Online Appendix

OA.1.1. Alternative cutoffs

The size cutoffs used to define the Naïve and Savvy trader types in the main specifications of my study were developed based on a mix of social media observations, interviews with money managers, and discussions with small retail traders. While the general tendency of trade size to positively covary with wealth and financial sophistication is grounded in theory (e.g., Ohlson 1975) and represented in datasets of retail brokerage accounts (e.g., Barber et al. 2023), the specific cutoffs I selected to demarcate the types are admittedly ad-hoc. To determine the robustness of my results to my particular selection of cutoffs, I redefine the Naïve type to be trades of 50 shares or less, and the Savvy type to be trade of between 100 and 5,000 shares.

Figure OA1 illustrates the relationship between Naïve and Savvy trade imbalance and contemporaneous returns both for the cutoffs used in the main paper (in solid lines) or the proposed alternatives (in dashed lines). The pattern of returns across quintiles is very similar. The alternatively defined Naïve line is slightly flatter, and the alternatively defined Savvy line is slightly steeper. Inferences as to the relationship between imbalance and returns would remain largely unchanged.

Table OA1 illustrates the results from Table 2A and 5A with alternative cutoffs.[1] Panel A replicates the results using *Firm Disclosure Day* interactions. In columns (1) and (2), the results are consistent with those in Table 2. The main effect for the Naïve (Savvy) type is slightly weaker (stronger), while the interaction effect is slightly stronger (weaker). The full model in column (3) produces very similar results.

Panel B replicates the results from Table 5A using the *EA Day* indicator variable interacted with the alternatively defined Naïve and Savvy imbalances. Again, the results are largely consistent with those in the paper. The exception is in column (2), where the interaction of *Savvy Imbalance ALT* and *EA Day*, is positive but only marginally significant. When considering this same interaction with the full model in column (3), the coefficient is positive and significant.

Panel C adds a dollar threshold for the definition of Naïve and Savvy. Specifically, the

[1] The sample size differs due to different dropped observations following the procedure outlined in section 3 of the paper. Specifically, because of the wider share bands, fewer Firm–Day observations have less than 5 trades for each type.

share size cutoffs remain the same as in the main paper, but with the additional restriction that Naïve trades are only considered if they do not exceed a ceiling of $1,000, and Savvy trades are only considered if they do not fall below a floor of $1,000. In this way, there are strictly demarcated trade sizes in dollars (i.e., regardless of nominal price, no Savvy trades and Naïve trades have the same dollar value). Results are similar to those reported in the main paper, with the exception of the *Savvy Imbalance 2 Cuts* interaction with *Firm Disclosure Day*, which is positive but insignificant. The coefficient on the interaction is positive and significant in the full model presented in column (3). However, it is smaller and less significant than in other versions of the same test.

OA.1.2. Problems with dollar-based classification

Why would the additional dollar threshold drive down the Savvy imbalance association with returns in Table OA1 Panel C? The dollar threshold effectively eliminates from the sample low nominal price observations (e.g., 100 shares of a $9.00 stock is no longer considered a Savvy trade, while 100 shares of an $11.00 stock still is). Nominal prices are strongly related to return-relevant factors such as MVE and BTM. Figure OA2 Panel A illustrates this relationship across nominal price deciles. Firm size increases monotonically in nominal price, and book-to-market ratios decline monotonically. As the cross-sectional tests in Table 6 illustrate, the Savvy type's relative outperformance is concentrated in smaller, more difficult to trade firms. As this alternative dual share and dollar threshold eliminates many of those trading observations from the sample, the observed association between imbalance and returns is driven towards zero.

Figure OA2 Panel B illustrates how the composition of Savvy and Naïve classified trades change when using a share-based, dollar-based, or two-threshold cutoff. Across all firm-days for each decile of nominal prices, I plot the percentage of retail trades classified as Savvy and as Naïve. The share-based cutoff used in the main paper produces a relatively smooth distribution of trade classifications across nominal price deciles, with fairly similar ratios of trades classified decile to decile. Both the dollar-based and two-threshold cutoffs produce sharp rises or falls in classification between deciles (e.g., using the two-threshold cutoff approach, just 10% of retail trades are classified as Savvy in the first decile, while more than 55% are classified as Savvy in the third decile). Thus, I retain the use of share-based cutoffs with nominal prices held constant via a fixed effect structure.

OA.1.3. Alternative return windows

Studies of price response to information events frequently examine the short window immediately preceding and following the news release. This accounts for any potential "information leakage" that may occur before an official announcement. In addition, a shorter window around the information event can tighten the identifying assumption that the return response is related to the released information. To address these concerns, I examine a *CRET* window of (-1,1) as well as simply the day 0 return.

I replicate the analysis from Table 2A using the alternatively defined returns windows and report results in Table OA2. All results in Panel A are consistent with those in Table 2A. The explanatory power of the models is slightly smaller (e.g., in column (3) the adj. $R^2 = 0.143$ versus 0.169 in the same specification of Table 2). The specifications in Panel B also yield similar inferences to those from Table 2A.

I conclude that either specification of cumulative return window leads to similar conclusions; namely, that the Savvy type trades have a strong, positive association with the return that grows more positive around firm disclosures, and the Naïve type have a strong, negative association with the return that grows more negative around information events.

OA.1.4. Stability of differences between types over time

The decade examined in my study (2014 – 2023) was one of significant changes in the retail trader landscape (e.g., fractional shares, $0 commission trading, mobile-phone trading optimization, the "meme-stock" events, COVID-19 stimulus money, increased participation rates, etc.) It could be the case that the differences I observe between the Naïve and Savvy type's associations with returns exist only before or after some of these changes. Rather than examining all changes individually, I simply plot the coefficients from a baseline regression of *CRET* on Savvy and Naïve trade imbalance over time. I limit the sample to only *Firm Disclosure Day* = 1, and plot β_1 and β_2 from the following equation in Figure OA3:

$$CRET_{i,t(0,4)} = \beta_1 Naive\ IB_{i,t} + \beta_2 Savvy\ IB_{i,t} + \psi_{firm} + \theta_{date} + \varepsilon_{i,t(0,4)} \qquad \text{(OA Eq. 1)}$$

The Savvy coefficient is significant and positive in all 10 years. The Naïve coefficient is significant and negative in 9 of 10 years, with the 2014 coefficient being insignificantly different from 0. A visual inspection of the coefficients over time does not exhibit any obvious or consistent time trend, though the Naïve type's imbalance does appear to have a more extremely negative association with returns in the latter half of the sample. I conclude that it is unlikely the results I find differentiating the performance of the Naïve and Savvy types are entirely the product of a recent change technology, costs, or trading environment.

OA.1.5. Explanation and illustration of AfterCostCRET *construction*

AfterCostCRET is computed by:

1. Calculating the day $t = 0$ return to buy trades and sell trades separately using the approach outlined in section 4.3 of the main paper. Namely, I compute the return from the average buy (or sell) execution price for type j to the closing midpoint price of firm i.
2. Multiplying the sell trades return by -1 (i.e., a sell followed by a positive price change represents a negative return to the trade).
3. Summing the day $t = 0$ calculated return for firm i and type j with the returns to firm i on days t+1 through t+4. As in step 2, the sell transaction returns are multiplied by -1.
4. Weighting the cumulative 5-day buy (sell) return by the proportion of buy (sell) trades on day $t = 0$.
5. Summing the trade-weighted 5-day buy return and the trade-weighted 5-day sell return for each Type–Firm–Day.

A simple illustration of computing *AfterCostCRET* is mapped in Figure OA4. Consider a Firm–Day with four Savvy-type trades. There are three buys at the Ask (say, \$1.05) and one sell at the Bid (say, \$0.95). The midpoint price when the trades are made is \$1.00 and at the close is also \$1.00, meaning the $t = 0$ return based on price change is exactly 0.00%. Say the return for days $t = 1$ through $t = 4$ is 4%. The variable would be constructed as:

1. Day *t(0)* Buy Return = (1.00 – 1.05) / 1.05 = - 4.76% and Day *t(0)* Sell Return = (1.00 – 0.95)/0.95 = 5.26%
2. Multiply the sell return by -1: (5.26% * -1) = -5.26%
3. Cumulative Buy Return = -4.76% + 4% = -0.76% and Cumulative Sell Return = -5.26% – 4% = -9.26%
4. Weighted Buy Return = -0.76% * (3/4) = -0.57% and Weighted Sell Return = -9.26% * (1/4) = -2.32%.
5. Sum to derive *AfterCostCRET* = -0.57% + -2.32% = -2.89%

OA.1.6. The decision not to trade

An inherent limitation of my empirical design is an inability to observe trader judgments which result in a decision not to trade. In other words, traders who judge the potential benefits of trading in response to a firm's disclosure as less than the relevant costs do not appear in my data.[2] It should be the case that a financially sophisticated trader makes judgments yielding both

[2] Even with this limitation, there is an economically meaningful amount of retail trading activity around disclosures in my sample, and it is increasing over time (see Figure OA5).

"trade" and "do not trade" decisions. I anticipate the Savvy type make the "do not trade" decision more frequently when faced with firms presenting the lowest opportunity to exploit an information advantage (i.e., easy to trade firms with rapid price adjustments to information).

I attempt to infer the existence of the "non-traders" by examining the proportion of retail trades which come from the Savvy type in the cross section of trading ease. A lower proportion of Savvy traders relative to the whole suggests more "do not trade" decisions being made by this type.[3] While transacting in easy to trade firms with efficient prices is relatively inexpensive, the potential benefits from exploiting new information are also lower. I propose that the benefits to trading these firms fall at a steeper rate than the corresponding costs, leading to less participation from more sophisticated traders. I test this prediction using a new dependent variable, *SavvyFraction*, which is equal to the count of Savvy retail trades for a Firm–Day divided by total retail trades on the same Firm–Day. I regress *SavvyFraction* on:

$$\begin{aligned} SavvyFraction_{i,t} = \; & \beta_1 FirmDisclosureDay_{i,t} + \beta_2 [Ease]_{i,t} \\ & + \beta_3 FirmDisclosureDay_{i,t} \times [Ease]_{i,t} + \beta_4 NominalPrice_{i,t} \\ & + \theta_{date} + \varepsilon_{i,t} \end{aligned} \qquad \text{(OA2)}$$

Equation OA2 is the first of the paper to exclude a Firm fixed effect (as I am here interested in cross sectional differences), and as such includes a control for the nominal share price. *Ease* is one of three cross-sectional indicator variables (i.e., *Big Firm, High Volume, High Spread*). The coefficient of interest is β_3, the interaction between firm disclosure days and the ease of trade.

Table OA3 presents results. In each column, the main effect for *Firm Disclosure Day* is positive and significant, suggesting a larger fraction of trades come from Savvy type traders when new information is released by more difficult to trade firms. In column (1) and (2), the interaction effect β_3 is negative and significant, indicating that relatively fewer Savvy traders trade in response to firm disclosures from large, high-volume firms. The interaction effect in column (3) is positive and significant, indicating that less liquid firms have a higher proportion of Savvy traders. The result in column (3) suggests that the higher transaction costs in the bid-ask spread are less than the perceived benefits from informed trade by the more financially sophisticated Savvy type.

[3] Or more "do trade" decisions being made by other retail traders. Either explanation suggests less Savvy "do trade" decisions relative to the overall retail proportion.

OA.1.7. Spreads relative to prices

The purpose of computing *AfterCostCRET* is to account for the cost of trading implicit in the spread. If the size of the spread relative to returns is immaterial, the after-cost return should not differ meaningfully from the return calculated at midpoint prices. To illustrate that spreads are a meaningful fraction of potential returns, particularly around EAs, I plot both the evolution of relative spreads in event time around EAs as well as intraday in Figure OA6 Panel A and Panel B. Panel A illustrates that the bid-ask spread relative to a firm's nominal price spikes at the EA event date. In Panel B, I divide observations into quintiles by average annual volume. Unsurprisingly, more traded equities are less expensive to trade. However, even the most liquid firms can experience spreads of nearly 0.5% at the open on EA days. Regardless of volume quintile, the pattern of higher spreads earlier in the EA day persist. As information asymmetry is resolved and the price reflects most of the new information, spreads shrink.

OA.1.8. Boehmer et al. (2021) method error rate robustness

Battalio et al. (2023) (BJSW) attempt to quantify the rate of both Type I and Type II errors in retail trades identified using the Boehmer et al. (2021) (BJZZ) method. Using one month of trades provided by a subset of wholesalers in May, 2022, BJSW are able to correctly identify 28% of "known" retail trades using the BJZZ method. This is lower than the 35% identified by the Barber et al. (2024) field study, but yields similar conclusions. This is problematic only if the unidentified trades vary from the identified trades in systematic ways. BJSW demonstrate that this can be case in the cross section by examining differences in time-to-trade execution metrics for identified and non-identified trades. As I restrict my analysis to within Firm–Day variation between subsets of retail traders, characteristic differences between identified and unidentified retail traders are not an explanation for my results.

Type I errors are the greater concern. If many of the "retail trades" I identify are actually misidentified institutional trades, it becomes very difficult to interpret my results. However, if the misidentified trades are 1) a relatively small fraction of total "retail trades" identified and 2) equally distributed between the Naïve and Savvy classifications, the misidentification simply adds measurement error to my proxies. To the first point, BJSW state that they can correctly identify 28% of their set of 53 million trades (p. 2-3), yielding approximately 14.8 million BJZZ identified retail trades. They are also provided 2.1 million institutional trades that are misidentified as retail by the BJZZ algorithm. This suggests that (at least for the data providing wholesalers in May of 2022) 88% of BJZZ identified trades are correctly identified retail trades and 12% are misidentified institutional trades (i.e., 2.1m/(14.8m+2.1m) = 12%).[4] While this suggests the vast majority of BJZZ identified trades are correctly identified, the misidentified trades could still be problematic if they contaminate the Savvy and Naïve types at different rates.

BJSW report that 38% of the misidentified institutional trades are for odd lot trades (1-99 shares) and 50% are for trades of 100 – 499 shares. I reexamine my primary results after redefining Naïve to include all odd-lot trades and Savvy to include only trades of between 100 and 499 shares. This redefinition, coupled with the rate of misidentification (12%), suggests that 5 of every 100 odd lot "retail trades" I observe are actually institutional trades. Similarly, 6 of every 100 "retail trades" of between 100 – 499 shares are actually institutional trades. In other words, there is a low rate of institutional trade contamination, and a fairly similar rate in both the redefined Naïve and redefined Savvy categories. I replicate the tests from Table 2A in Table OA4 Panel A and find results consistent with my main findings.

As an alternative robustness check, I examine the subset of Firm–Days least likely to be impacted by Type I errors. BJSW reports a determinates model of Type I misclassification, and reports that misclassification is less prevalent in smaller firms with wider spreads. In Panel B of Table OA4, I replicate the full model from Table 2A on firms with below median MVE, above median spreads, and the intersection of these two subsets. In all instances, I find results consistent with the main findings. I conclude that it is unlikely Type I errors are driving the differential responses I observe.

OA.1.9. Excluding fractional share trades

An emerging literature explores the smallest of trade sizes: fractional share trading. Several studies seek to quantify the effects of fractional trading on the capital markets and to identify fractional traders (Bartlett et al. 2024; Lin 2024; Da et al. 2024). Fractional share trades are "rounded up" to a single share for the consolidated tape, and are thus included in my proxy for the Naïve type of retail trader. This seems generally appropriate, as the ability to trade less than a full share of most securities likely matters more to less experienced, less sophisticated, and less well capitalized investors that I deem the Naïve type. However, fractional share trades can also arise for other reasons (e.g., dividend reinvestment, placing orders in dollars rather than shares, etc.). These sorts of fractional trades likely occur for all investor types, and add noise my

[4] These are my calculations based on the numbers reported in the working paper, not calculations from the authors.

measure of the Naïve type. I reperform the main analyses limiting the Naïve type to 2-10 shares. In untabulated results, I find that my main results are all stronger and more significant for the Naïve type, suggesting the fractional/single share trades add noise to my Naïve proxy. Future work could consider limiting the Naïve identification to 2-10 shares.

OA.1.10. Validity of the naïve trade proxy

I use microtrades of 1-10 shares to identify trades from Naïve type retail investors. The literature has considered alternative measures to proxy for unsophisticated traders, such as brokerage selection. Both Fong et al. (2014) and Eaton et al. (2022) hypothesize that an investor's brokerage type selection signals their type. Eaton et al. uses Robinhood outages as a proxy for a reduction of less sophisticated traders in the market.[5] The advantage of my measure relative to a brokerage outage proxy is that TAQ trades are available for all NMS regulated securities for every trading day. If my microtrades imbalance measure is positively correlated with the trading of the less experienced, less wealthy, and less financially sophisticated Robinhood investors, it bolsters the assertion that the trade size proxy captures the "Naïve Trader" construct.

To test this association, I hand collect Robinhood trade imbalance information from July 2024 – February 2026. Robinhood displays the trade imbalance of its customers by ticker to anyone with a free account (see an illustration for Microsoft's Robinhood page in Figure OA7 Panel A). I regress the Robinhood imbalance on the BJZZ identified Naïve and Savvy trade imbalances. If both the BJZZ Naïve imbalance and Robinhood imbalance capture the trading of less financially sophisticated retail investors, I would anticipate a positive coefficient on the *Naïve IB* term. To the extent BJZZ Savvy traders are mostly absent from the Robinhood population or (as demonstrated in the main paper) trade very differently than the Naïve type, I anticipate either an insignificant or negative coefficient on the *Savvy IB* term.

Table OA5 presents results by retail volume quintile. With low retail trading volume (quintiles 1 and 2), there is an insignificant relationship between *Naïve IB* and Robinhood imbalance. However, as the level of retail trading volume increases (quintiles 3-5), the relationship grows more positive and more significant.[6] At the same time, the *Savvy IB* and

[5] Eaton et al. (2022) note that Robinhood account holders are both younger and less wealthy than account holders in the dataset from Barber & Odean (2000), which is also drawn from a discount brokerage. In addition, Robinhood claims that millions of their customers are first time investors.

[6] This is likely attributable to imbalance measures becoming more precise and less volatile as trade volume grows. Consider a firm-day with 2 trades. The trade imbalance measure can only take values of -1, 0, or 1. Contrast that to a

Robinhood imbalance association grows more negative. In the highest volume quintile, a 1 unit change in Naïve imbalance is associated with a 0.6 unit change in Robinhood imbalance (t-statistic = 26.61). In other words, my BJZZ identified Naïve trader proxy and the actual trade imbalance of Robinhood customers are positively and significantly associated.

Figure OA7 Panel B provides a visual depiction of the highest volume quintile association between imbalance measures. I plot the mean values for *Naïve IB* and *Savvy IB* in each of 50 quantile bins of Robinhood imbalance. The association between the Naïve type and Robinhood imbalance is stronger than the association between the Savvy type and Robinhood imbalance. Note that the Naïve type slope does not approach the 45-degree angle which would represent a 1:1 correspondence. I have no expectation that 1) Robinhood's customer base is exclusively Naïve traders or 2) the 1-10 microtrade proxy exclusively and entirely captures Naïve type trades. Rather, the correspondence provides comfort that my proxy is positively associated with an observable measure of retail traders generally thought to be less financially sophisticated.

firm-day with 200 trades, where the imbalance measure can take any value between -1 and 1 rounded to two decimal places.

Figure OA1: Alternative Cutoffs

This figure illustrates the average *CRET* by trade imbalance quintiles of both the Savvy and Naïve types. The solid lines plot the cutoffs used in the main paper, and the dashed lines plot the alternate cutoffs suggested in OA1.1. The first quintile are the strongest Sell imbalances, and the fifth quintile are the strongest Buy imbalances. Quintiles are computed by firm and year (e.g., the 20% of strongest Sell imbalances for Firm XYZ in Year 20XX are aggregated in Quintile 1).

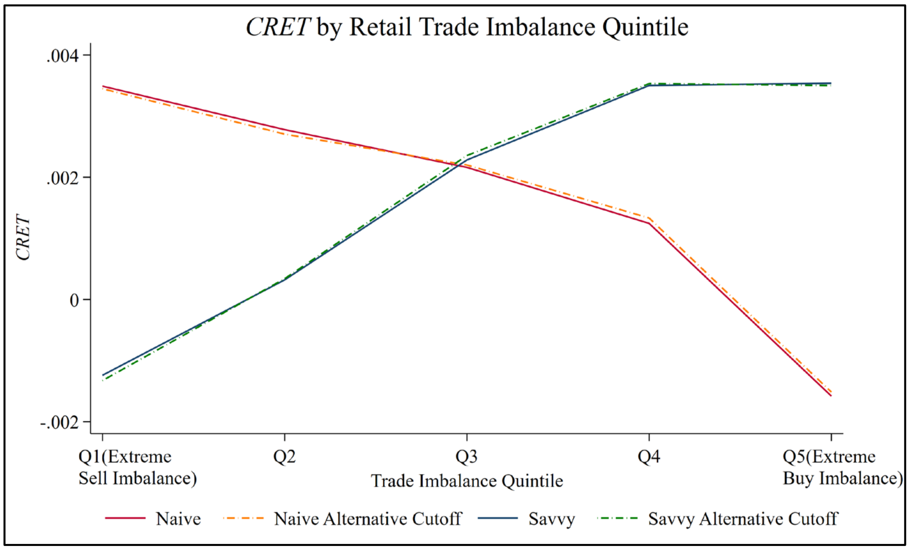
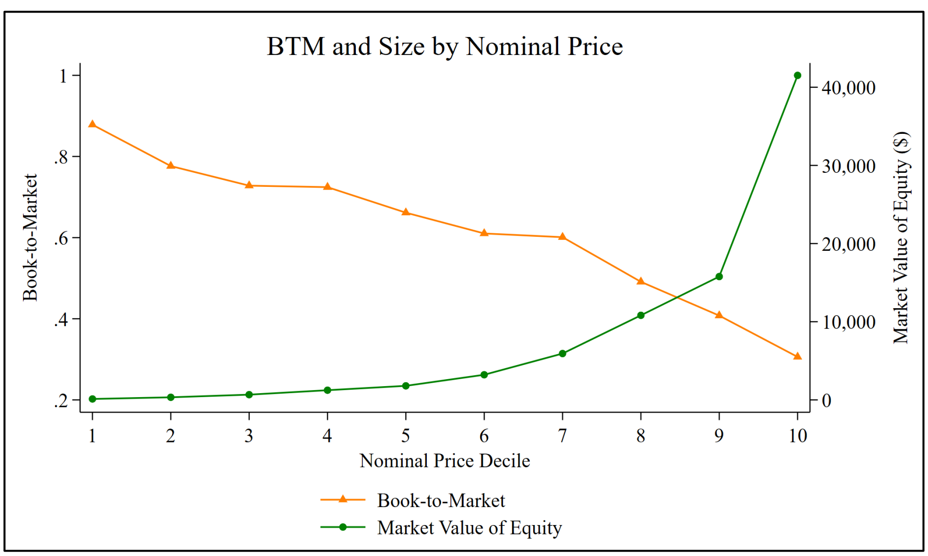


Figure OA2: Nominal Price Deciles and Trade Classification

This figure illustrates the relationship of firm attributes and trade classifications by nominal price decile. Nominal price deciles are based on closing midpoint prices by firm-day. Panel A plots average book-to-market ratios and market value of equity. Panel B plots the percent of retail trades classified as Naïve or Savvy using either share cutoffs (1-10 for Naïve, 100-999 for Savvy), dollar cutoffs (<$250 for Naïve, $1k<Savvy<$5k), or a combination (1-10 and <$1k for Naïve, 100-999 and >$1k for Savvy).

Panel A: Book-to-Market ratio and Market Value of Equity plotted by Nominal Price



Panel B: Percent of retail trades classified as Naïve or Savvy under different cutoff criteria

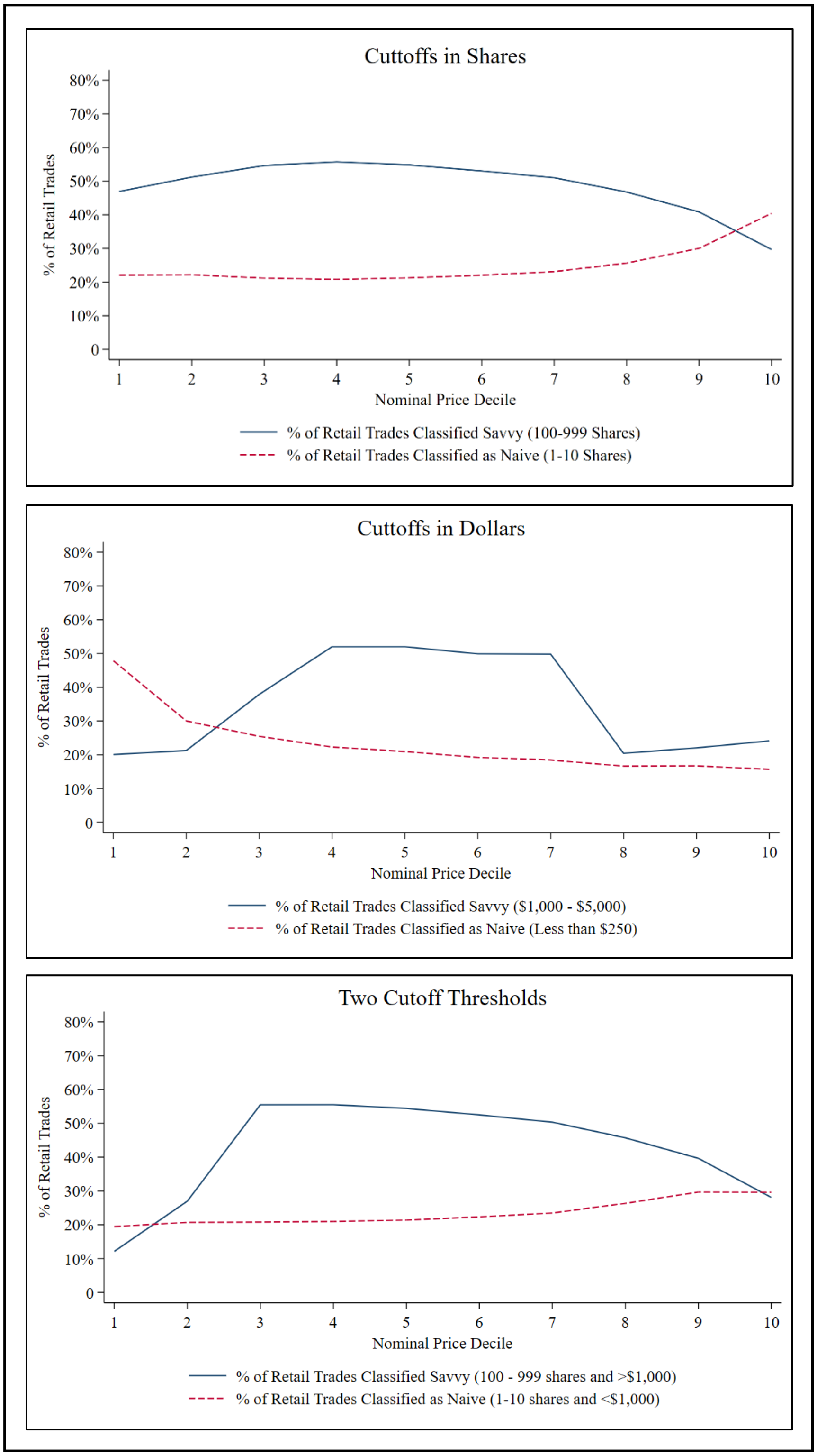


Figure OA3: Coefficient of Baseline Association Over Time

This figure plots the yearly coefficient estimates for the Naïve and Savvy type trade imbalances on firm disclosure days from OA Eq. (1). All variables are defined as in Appendix A of the main paper.

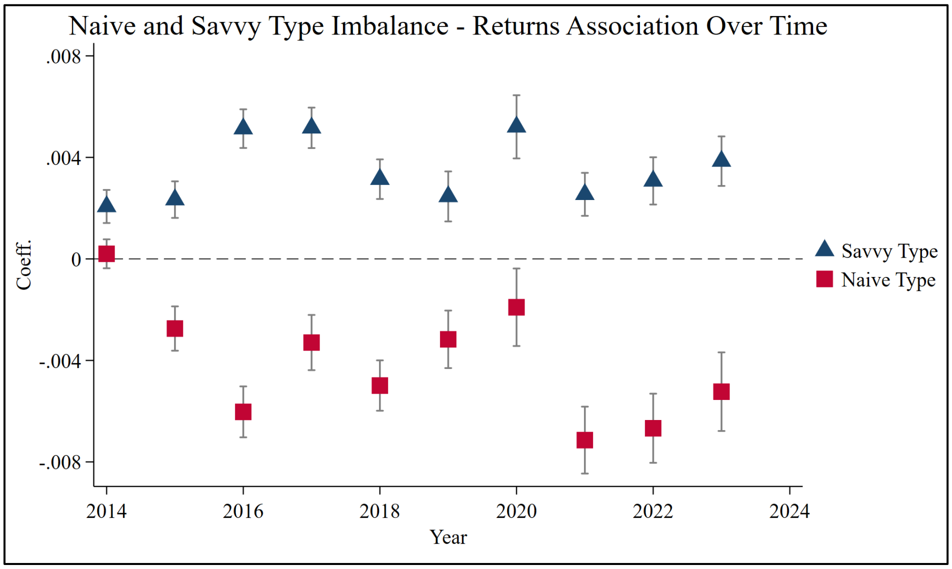


Figure OA4: Calculating *AfterCostCRET*

This figure illustrates the numerical calculation of *AfterCostCRET* for a hypothetical Firm–Day with 4 total trades.

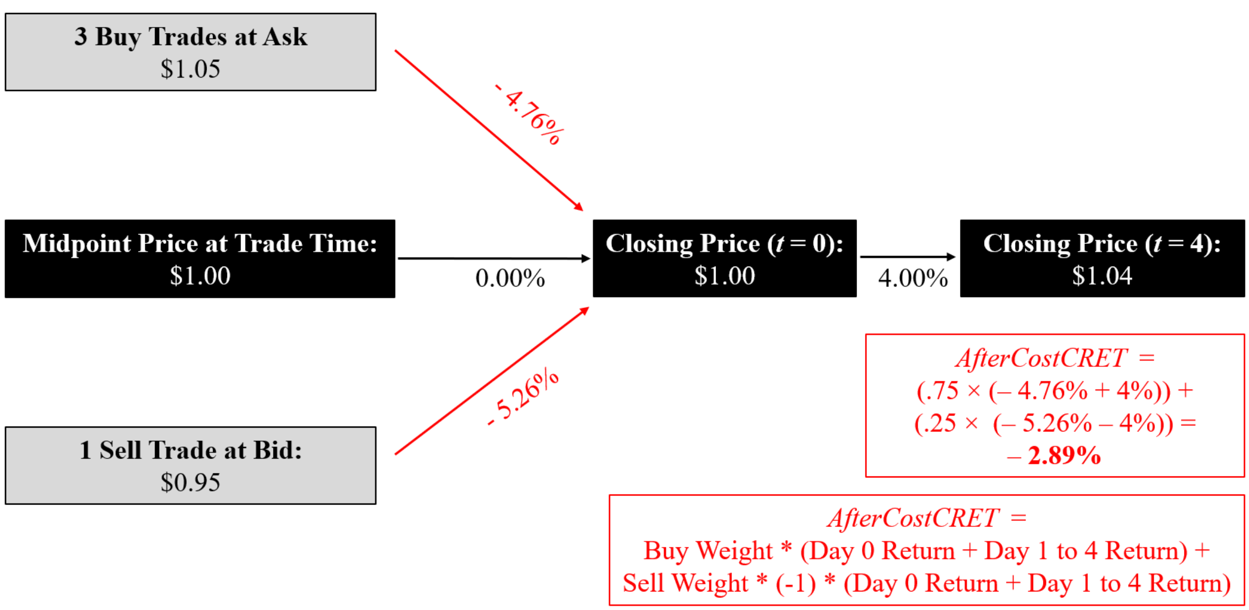


Figure OA5: Dollars of Retail Trade in Sample

This figure illustrates the growth in Boehmer et al. (2021) detected retail trade dollars for the Firm-Days which include a firm disclosure during my sample window.

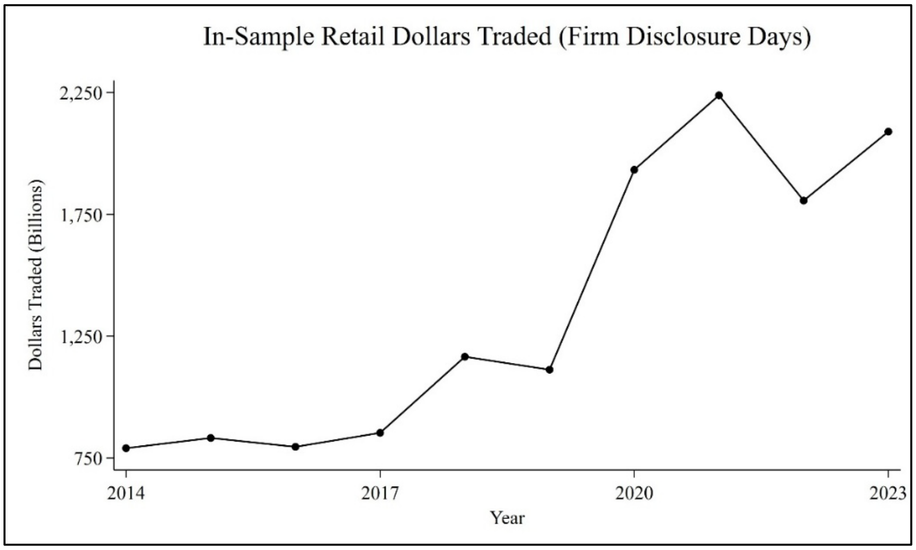
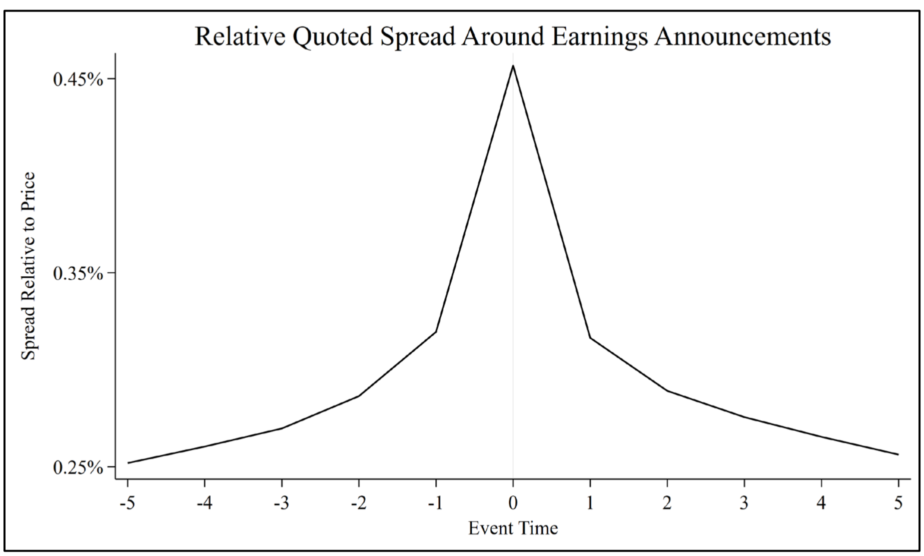


Figure OA6: Spread on EA days

This figure illustrates the evolution of the quoted spread relative to executed prices both around EA days (Panel A) and within EA days (Panel B). In Panel A, relative spread is the average quoted spread for the EA day over the closing price. In Panel B, relative spread is the average quoted spread during each 5-minute increment over the average price of trades executed in the same increment. In Panel B, spreads are plotted based on quintiles of average volume, computed at the annual level.

Panel A: Event time spread around EA



Panel B: Spread by Volume Quintile

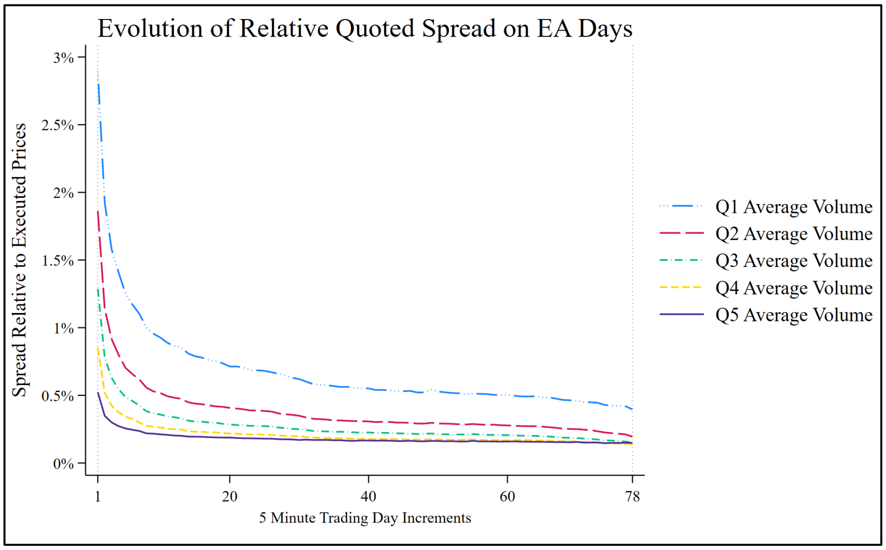


Figure OA7: Validating Naïve Proxy against Robinhood Imbalances

This figure illustrates the relationship between the proxies developed in this paper and an observable measure of Robinhood Retail Trader imbalances. Panel A is an illustration of the Robinhood Trade Imbalance measure. Panel B plots the average Naïve and Savvy trade imbalance per each of 50 quantiles of Robinhood trade imbalance.

Panel A: Example of Robinhood Trade Imbalance information for Microsoft

Panel B: Robinhood imbalance vs. Boehmer et al. identified Naïve and Savvy imbalances

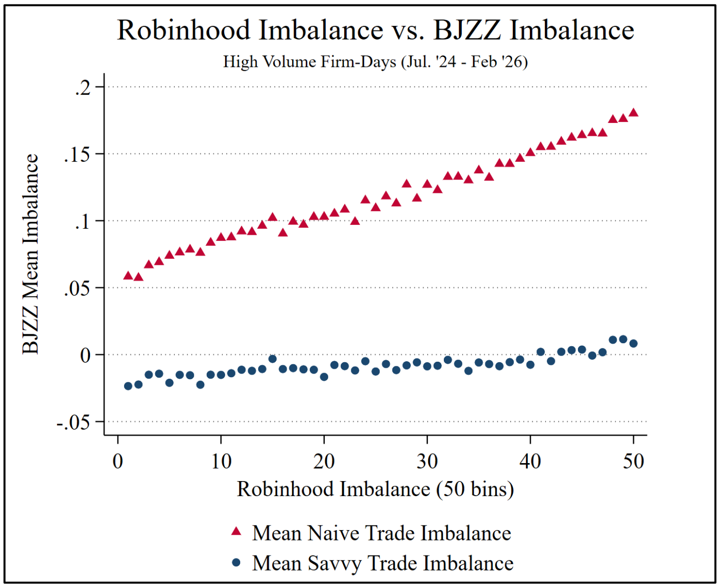


Table OA1: Alternative Trade Size Cutoffs

Panels A and B replicate the analysis from Table 2 and Table 5 of the main paper using alternative definitions for the Naïve and Savvy type traders. Naïve type trades are now defined as any trade of 50 shares or less. Savvy type trades are now defined as any trade of between 100 and 5,000 shares. Panel C uses a double threshold rule, with the same share cutoffs as the main paper, but the additional restriction that Naïve trades have a ceiling of $1,000 and Savvy trades have a floor of $1,000. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date. T-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Alternative Trade Size Cutoffs on Firm Disclosure Days

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB ALT	-0.0003***		-0.0007***
	(-3.32)		(-9.06)
Savvy IB ALT		0.0025***	0.0026***
		(41.51)	(43.97)
Firm Disclosure Day	0.0021***	0.0020***	0.0021***
	(14.18)	(13.37)	(14.01)
Naïve IB ALT × Firm Disclosure Day	-0.0026***		-0.0029***
	(-14.57)		(-15.81)
Savvy IB ALT × Firm Disclosure Day		0.0007***	0.0013***
		(5.50)	(9.99)
Observations	6,146,932	6,146,932	6,146,932
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.164	0.165	0.165

Panel B: Alternative Trade Size Cutoffs on EA days

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB ALT	-0.0003***		-0.0007***
	(-3.94)		(-9.81)
Savvy IB ALT		0.0026***	0.0027***
		(42.15)	(44.81)
EA Day	0.0015***	0.0004	0.0016***
	(3.19)	(0.77)	(3.47)
Naïve IB ALT × EA Day	-0.0124***		-0.0132***
	(-21.02)		(-22.16)
Savvy IB ALT × EA Day		0.0010*	0.0043***
		(1.86)	(8.01)
Observations	6,146,932	6,146,932	6,146,932
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.164	0.165	0.165

Panel C: Combined Share and Dollar Cutoffs

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB 2 Cuts	-0.0007***		-0.0009***
	(-8.51)		(-11.13)
Savvy IB 2 Cuts		0.0021***	0.0022***
		(34.52)	(36.32)
Firm Disclosure Day	0.0022***	0.0021	0.0022***
	(14.03)	(13.30)	(13.91)
Naïve IB 2 Cuts × Firm Disclosure Day	-0.0021***		-0.0021***
	(-11.58)		(-12.00)
Savvy IB 2 Cuts × Firm Disclosure Day		0.0002	0.0004***
		(1.18)	(3.14)
Observations	5,241,884	5,241,884	5,241,884
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.172	0.173	0.173

Table OA2: Alternative Return Window

This table presents an analysis of returns around *Firm Disclosure Days* with alternative return windows. The dependent variable in Panel A is the day 0 return, and the dependent variable in Panel B is the (-1,1) cumulative return. Naïve and Savvy trades and indicator variables are as defined in Appendix A. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date. T-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Return (0) window

Dependent Variable = *RET*	(1)	(2)	(3)
Naïve IB	-0.0009***		-0.0012***
	(-17.27)		(-22.29)
Savvy IB		0.0022***	0.0023***
		(52.60)	(56.79)
Firm Disclosure Day	0.0022***	0.0021***	0.0022***
	(19.95)	(19.40)	(19.89)
Naïve IB × Firm Disclosure Day	-0.0020***		-0.0021***
	(-14.22)		(-15.00)
Savvy IB × Firm Disclosure Day		0.0007***	0.0010***
		(7.40)	(10.48)
Observations	5,430,434	5,430,434	5,430,434
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.140	0.142	0.143

Panel B: Return (-1,1) window

Dependent Variable = *CRET(-1,1)*	(1)	(2)	(3)
Naïve IB	-0.0016***		-0.0020***
	(-19.41)		(-24.21)
Savvy IB		0.0037***	0.0039***
		(57.02)	(60.61)
Firm Disclosure Day	0.0027***	0.0025***	0.0027***
	(16.79)	(16.05)	(16.72)
Naïve IB × Firm Disclosure Day	-0.0020***		-0.0022***
	(-11.23)		(-12.17)
Savvy IB × Firm Disclosure Day		0.0012***	0.0015***
		(8.94)	(11.34)
Observations	5,430,434	5,430,434	5,430,434
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.154	0.156	0.157

Table OA3: The Decision Not to Trade

This table presents an analysis of the fraction of Savvy traders relative to all retail traders around firm disclosure events in various cross sections. The indicator variable *Firm Disclosure Day* takes a value of 1 if the firm issued a press release or earnings announcement on that date (as reported in Ravenpack and IBES). The cross sections are defined by indicator variables for Firm–Days above the median quarterly market value of equity (*Big Firm*), trading volume (*High Volume*), and relative spread (*High Spread*). Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date, and t-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Dependent Variable = *SavvyFraction*	(1)	(2)	(3)
Firm Disclosure Day	0.0239***	0.0263***	0.0057***
	(27.49)	(28.83)	(3.66)
Big Firm	0.0051**		
	(2.14)		
High Volume		0.0423***	
		(22.51)	
High Spread			-0.0329***
			(-16.49)
Firm Disclosure Day × Big Firm	-0.0134***		
	(-8.05)		
Firm Disclosure Day × High Volume		-0.0254***	
		(-15.29)	
Firm Disclosure Day × High Spread			0.0188***
			(10.89)
Nominal Share Price	-0.0012***	-0.0012***	-0.0013***
	(-41.36)	(-49.46)	(-47.45)
Observations	5,430,490	5,430,490	5,430,490
Fixed Effects	Date	Date	Date
Adj. R-squared	0.539	0.550	0.545

Table OA4: BJZZ Type I Error Robustness

This table presents results for a replication of Table 2A using alternate specifications for both Savvy and Naïve cutoffs, as well as a subset of the sample. Battalio et al. (2023) report that of identified Type I errors, 38% are in the odd lots, and 50% are between 100-499 shares. Panel A redefines Naïve and Savvy along these cuts to better balance the incidence of any institutional contamination. Panel B replicates the full model from Table 2A in three subsamples: below median MVE, above median bid-ask spread, and the intersection of these two designations. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date. T-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Balancing Type I Error Incidence

Dependent Variable = *CRET*	(1)	(2)	(3)
Naïve IB	0.0000		-0.0005***
	(0.57)		(-6.04)
Savvy IB		0.0024***	0.0025***
		(42.36)	(44.67)
Firm Disclosure Day	0.0021***	0.0020***	0.0021***
	(13.95)	(13.30)	(13.88)
Naïve IB × Firm Disclosure Day	-0.0025***		-0.0029***
	(-13.95)		(-15.90)
Savvy IB × Firm Disclosure Day		0.0011***	0.0018***
		(7.99)	(12.95)
Observations	6,250,572	6,250,572	6,250,572
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.164	0.165	0.165

Panel B: Subsamples with lower incidence of misidentification

Dependent Variable = *CRET*	(1)	(2)	(3)
Subset	Below Median MVE	Above Median Spread ($)	Below Median MVE & Above Median Spread ($)
Naïve IB	-0.0007***	-0.0003***	0.0001
	(-8.18)	(-3.39)	(1.20)
Savvy IB	0.0026***	0.0030***	0.0031***
	(35.30)	(43.97)	(36.67)
Firm Disclosure Day	0.0036***	0.0025***	0.0039***
	(12.64)	(12.19)	(10.79)
Naïve IB × Firm Disclosure Day	-0.0031***	-0.0020***	-0.0017***
	(-12.22)	(-9.13)	(-5.93)
Savvy IB × Firm Disclosure Day	0.0021***	0.0008***	0.0013***
	(9.74)	(4.78)	(5.38)
Observations	2,716,307	2,714,005	1,341,226
Fixed Effects	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.153	0.168	0.139

Table OA5: Robinhood Imbalance Validation

This table presents the results of regressing Robinhood Trade Imbalance on BJZZ identified Naïve and Savvy trade imbalance. Each column presents the results for each retail trade volume quintile of the sample. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by Firm and Date. T-statistics are reported in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Dependent Variable = *Robinhood IB*	Volume Quintile 1	Volume Quintile 2	Volume Quintile 3	Volume Quintile 4	Volume Quintile 5
Naïve IB	-0.0110	-0.0099	0.0484***	0.2209***	0.6022***
	(-1.14)	(-0.86)	(3.51)	(11.73)	(26.61)
Savvy IB	0.0114	0.0174*	-0.0226**	-0.0907***	-0.1944***
	(1.29)	(1.66)	(-2.03)	(-6.96)	(-10.17)
Observations	128,985	128,956	128,937	128,900	128,877
Fixed Effects	Firm, Date	Firm, Date	Firm, Date	Firm, Date	Firm, Date
Adj. R-squared	0.030	0.031	0.039	0.043	0.097