Anomalies Never Disappeared:

The Case of Stubborn Retail Investors*

Xi Dong †

Cathy Yang‡

Dec 2023

* We would like to thank Brad Barber, Andrew Chen, Alex Chinco, Clifton Green, Lin Peng, Yajun Wang, Morad Zekhnini, Dexin Zhou, and seminar participants at Baruch College, 2023 FMA Consortium, 2023 Greater New York Finance Women Inaugural Symposium for comments and suggestions.

† Zicklin School of Business, Baruch College. E-mail: Xi.Dong@baruch.cuny.edu.
‡ Zicklin School of Business, Baruch College. E-mail: yuqing.Yang@baruch.cuny.edu.

Anomalies Never Disappeared:

The Case of Stubborn Retail Investors

Dec 2023

Abstract

Defying conventional beliefs, our analysis of 260 anomalies over the last half-century reveals that they have not vanished as market efficiency improves, because their alphas predominantly materialize over the long haul—an overlooked timeframe with far reaching financial and real implications. The enduring long-run alphas in recent years are driven by the anomalies that retail investors trade against, yielding staggering value-weighted two-year alphas of 23%. Incorporating retail trading, we develop asset pricing models that surpass existing prominent models in explaining these long-run alphas. We propose and validate a hypothesis: the stubbornness of retail investors underpins long-run alphas, inflicting long-horizon risks on arbitrageurs. Our findings imply that as society advances and other frictions fade, the unyielding nature of financially-naive individuals will remain an enduring impediment to market efficiency.

JEL: G10, G11, G14

Keywords: Anomalies, Retail investors, Horizon, Noise Trading Risk, Limits to Arbitrage, Return Predictability.

1. Introduction

Over the last five decades, academics have uncovered countless patterns of anomalous cross-sectional return predictability that are not well explained by standard asset pricing theories.[1] Many studies report that anomaly returns are disappearing in recent years.[2] The explanations include that anomaly publication attracting significant arbitrage trading, improved fund performance measurement and firm information disclosure, and increased trading and liquidity.[3] These explanations are consistent with the conventional wisdom that market efficiency improves as society advances and frictions fade over time. However, prior literature focuses on the formation-month return rather than the long-run abnormal return of anomalies. Recent studies point out that long-horizon anomaly alphas have far-reaching financial and real implications. On the financial side, they are crucial for understanding asset pricing, market efficiency, and the actual profits of trading strategies. After all, a good asset pricing theory should be able to explain long-horizon moments.[4] On the real side, long-run alphas are the ones that really matter for real inefficiencies as no firms make investment decisions expecting outcomes in a month.[5] Separately, although meta-analyses provide some mixed evidence on the role of institutions in anomalies,[6] there is little evidence on how retail investors

---

[1] The *Journal of Financial Economics* published a special issue on return anomalies as early as 1978. Since that time, hundreds of cross-sectional return anomalies have been documented (e.g. Harvey et al., 2015).

[2] Green, Hand and Zhang (2017) conduct a meta-analysis of 97 anomalies and find that only two anomalies are independently significant return predictors since 2003, and, outside microcaps, the long-short returns have also become insignificant since then. Other studies document the declining trend for a subset of anomalies, e.g., Chordia, Subrahmanyam, and Tong (2014), Fama and French (2021), Ben-David, Li, Rossi, and Song (forthcoming), Smith and Timmermann (2022), and Bowles, Reed, Ringgenberg, and Thornock (2023).

[3] Dong et al. (forthcoming) provides comprehensive evidence that hedge funds and short sellers significantly increase their intensity of exploiting 99 anomalies after anomaly publication. McLean and Pontiff (2015) provide similar evidence using short interest, while Calluzzo et al. (2019) provide hedge fund evidence for a handful of anomalies. Ben-David, Li, Rossi, and Song (forthcoming) shows that a reform to the Morningstar fund performance rating system in 2003 reduces momentum anomaly profits. Bowles, Reed, Ringgenberg, and Thornock (2023) show that 10K statement is released increasingly earlier, allowing arbitrageurs to quickly correct the mispricing in accounting-based anomalies. Kim, Ivkovich, and Muravyev (2021) shows similar results using the introduction of Edgar as a shock that lowers the information costs. Chordia, Subrahmanyam, and Tong (2014) shows that increased trading and liquidity in recent years make a handful of anomalies disappear.

[4] Examples of studies making this argument include Baba-Yara, Boons, and Tamoni (2020), Chernov, Lochstoer, and Lundeby (2022), Gormsen and Lazarus (2023), van Binsbergen, Boons, Opp, and Tamoni (2023), Cho and Polk (forthcoming).

[5] Van Binsbergen and Opp (2019) shows that it is not the magnitude of alphas but the persistence of alphas of anomalies that are important for causing real outcomes.

[6] In contrast to the hedge fund and short seller evidence in Footnote 3, institutions, most of which are mutual and pension funds, tend to trade against anomalies implying they may exacerbate anomalies (Edelen, Ince, and Kadlec,

shape anomalies. The answer to this question is not only necessary for providing a complete picture of the role of different market participants in anomalies, but also a priori unclear as the retail trading literature is debating whether retail trading is smart or dumb.[7]

In this article, we contrast the long vs. short-run alphas of up to 260 established cross-sectional anomalies and explore the role of retail trading in these alphas. We focus on the CAPM alpha because as argued by Jensen, Kelly, and Pedersen (2022), economic theory dictates that the truly anomalous return is the one unexplained by the standard asset pricing theory if we wish to learn about "anomalies."

We start by examining the value-weighted long-short anomaly portfolio alphas over the portfolio formation month (t+1) and those over the 24 months post ranking skipping the formation month (t+2 to t+24), holding the anomaly ranking (computed in month t) for each stock constant throughout the two-year event time. To examine the evolution of alphas over time, we break the last 50 years into three periods: 2009-2022, 1989-2008, and 1969-1988, and compute alphas separately for each. 2009-2022 are the recent years we refer to, for which direct retail trading data are available (in some tests, we reserve the post-Covid period March 2020-December 2022 as a case study). We find that the formation-month alpha indeed gradually declines over time, with the percentage of significant anomalies ($t>1.96$) more than halved from 55% to 24% in recent years and the magnitude of alphas nearly halved to 0.36% per month. This pattern aligns with the conventional wisdom that anomalies are disappearing. However, when examining the 24-month alphas, we find a sharply different picture. These alphas remain similarly strong across all three periods. Furthermore, the total alphas of anomalies predominantly materialize not in the formation month but over the long haul, suggesting that this overlooked timeframe is an economically more relevant

---

2016). Akbas et al. (2015) show correlation-based evidence that flows to mutual and hedge funds induce fund managers to exacerbate and correct a few noninvestment-based anomalies, respectively. Dong, Kang, and Peress (2023) provide direct trading and causal evidence for this argument. These results point towards individual investors as a potential underlying cause of at least some anomalies.

[7] On the one hand, individual investors have been shown to exhibit behavioral biases or are uninformed (e.g., Kumar and Lee 2006; Barber, Odean, and Zhu 2009; Barber, Lin, and Odean 2021; Bryzgalova, Pavlova, and Sikorskaya 2022). On the other hand, other work finds that aggregate retail order flows correctly predict returns (e.g. Kaniel, et al. 2012; Kelley and Tetlock 2013; Boehmer et al. 2021), suggesting that retail investors in aggregate appear informed.

timeframe to understand anomalies. Since we deal with long-horizon returns, we ensure the significance of our findings by using various methods throughout the paper, including the Newy-West standard error adjustment, the Jagadeesh-Titman nonoverlapping return approach, or double-clustered standard errors.

To gain further insight into the alpha pattern, we separate anomalies based on abnormal retail trading intensity (ATI) in individual anomalies during the quarter (i.e., month t-2 to t) before the formation month. Our approach relies on the retail trade algorithm developed by Boehmer, Jones, Zhang, and Zhang (2021) (BJZZ), and we track how retail trading evolves as stocks acquire their anomaly-defining characteristics. We construct the anomalies so that their average long-short returns deliver positive alphas in the formation month. Therefore, if the ATI is negative (positive) in an anomaly, we categorize this anomaly into the anomalies that retail investors trade against. We find that retail investors trade against around 60% of all anomalies. For the remaining 40% of anomalies that they trade with, the significance and magnitude of alphas indeed drop to nonexistence in recent years regardless of return horizons. In contrast, among the anomalies that they trade against, the percentage of significant formation-month alphas also drops from 56% to 34% over the past 50 years, whereas the percentage of significant 24-month alphas increases from 40% to almost 60%. In terms of magnitude, the average formation-month alphas only drop modestly from 0.6% per month to 0.5%, whereas the 24-month alphas increase from 10.29% to 14.44%. Therefore, the improvement of market efficiency over the years appears to only have affected the alphas of the anomalies that retail investors trade with but not the anomalies that they trade against, especially for the long run alphas.

We then evaluate whether anomaly alphas vary with the degree of retail anomaly ATI. We sort anomalies based on their pre-formation (t-2 to t) ATI into five groups from R1 to R5, with R1 being the anomalies with the most positive ATI and R5 the most negative. We compute the average ATI and return across anomalies within each group. Retail investors trade against three groups of anomalies R3-R5 and trade with two groups of anomalies R1-R2, consistent with the aforementioned 60%/40% with/against ratio. The pre-formation alphas are negative for R3-R5 but positive for R1-R2, matching the sign of pre-formation

ATI. Moving from R3 to R5, as ATI becomes more negative, alphas also turn more negative, from -0.13% in R3 to -3.33% in R4 and to -4.25% in R5 per quarter. However, the sign of alphas of R3-R5 reverses post-formation, turning significantly positive across return horizons. The magnitude of alphas dramatically increases as the horizon lengthens. Remarkably, R5 anomalies can achieve a substantial 23.36% 24-month *value-weighted* alpha. In stark contrast, for R1-R2, although the pre-formation alphas are economically and statistically significantly positive, they are insignificant across all event horizons except for the formation month of R2. Overall, the results suggest that the more retail investors trade against an anomaly before portfolio formation, the higher the anomaly alpha will be after portfolio formation, especially in the long term. In contrast, when retail investors trade with an anomaly, any mispricing associated with the anomaly appears to have largely been corrected before portfolio formation.[8] It is those anomalies that retail investors trade against that are responsible for explaining the enduring long-run alphas we observe earlier when examining all anomalies.

To understand what anomaly characteristics retail investors trade against/with and rule out that their trading tendency is random, we group anomalies according to different economic concepts identified in prior literature. We find that retail investors trade against characteristics not in a random way but in a fashion consistent with prior considerations of retail preferences/beliefs. For example, they trade against profitability, trading frictions, and value vs. growth anomalies (Hou, Xue, and Zhang 2020); lottery anomalies (Kumar 2009); and anomalies having speculative short legs (Birru 2018).

Why do the alphas of the anomalies that retail investors trade against survive in recent years while those they trade with disappear? Based on our findings so far, we hypothesize one underlying cause is that retail investors have stubborn preferences/beliefs, inflicting long-horizon risks on arbitrageurs who take the opposite side in those anomalies traded against by retail investors.

---

[8] We interpret anomalies as mispricing-based for expositional purposes. Our paper's conclusion could also be consistent with some risk-based interpretations.

This long horizon risk can persist as *noise* traders' preferences/beliefs, by definition, need not be rational, and thus, the *rational* measures aimed to improve market efficiency over the last half-century need not change retail behavior. Furthermore, retail investors' influence on stock prices may even increase in recent years and in the future due to easier coordination through social media (Barber, Huang, Odean, and Schwarz, 2022), whose effect is also amplified by the convenience offered by the advancement of fintech such as zero-commission and fractional trading. Indeed, before the social media/fintech era, small and noisy retail trades could cancel each other, rendering the net retail trading not concerning. But now this is not the case. This view is shared by the popular press and general public. For example, it is argued in a recent Hollywood biographical movie *Dumb Money* about the GameStop saga that "There was no hope for the little guy. Maybe now there is." Many hedge funds are now actively monitoring retail trades to control risk. Indeed, we show that the trading volume of retail investors are steadily increasing in recent years; Their net trading intensity in the anomalies that they trade against also triples over recent years.

Based on our hypothesis, retail investors pose, and may continue to pose, a significant horizon risk, enabling the longevity of anomalous returns in financial markets. To highlight the significance of retail investors' role in understanding asset pricing, we develop factor models incorporating retail trading information. We introduce two methods to build a retail factor. The first method, statistically motivated, takes the average returns across the quintile of anomalies most negatively traded by retail investors. The second method, rooted in theory, accounts for the marginal utility of arbitrageurs opposing retail investors. The resulting asset pricing models considerably outperform existing prominent models in explaining the long-run alphas of the surviving anomalies in recent years—the anomalies that are still relevant for asset pricing models to explain.

Having established the importance of retail trading for understanding asset pricing, we next further support our hypothesis. If retail investors stick to their misactions or the sentiment underlying their misactions may even increase post formation, this introduces additional risk for arbitrageurs based on three considerations. First, the arbitrageurs will be exposed to the fundamental risk as well as the noise trader risk of their stock holdings for a long spell of time. Indeed, the effective holding costs of an arbitrage position is its life-time risk exposure—a critically important form of arbitrage holding costs (e.g., Pontiff 2006). Second, short sellers' holding costs, including short-selling fees and short-selling fee risk, increase proportionally with the holding horizon (Engelberg, Reed, and Ringgenberg 2018), endangering the viability of arbitrage. Short sellers' concern is the most relevant concern for exploiting anomalies as most anomaly alphas are driven by anomaly short legs (e.g., SYY 2012, 2015; Dong, Li, Rapach, and Zhou 2022), which we confirm in our sample. In equilibrium, short sellers will require a higher expected profit to trade against noise traders if they expect mispricing will take a longer time to correct or can even exacerbate in the short term. Therefore, the anomalies that retail investors trade against should generate higher alphas, particularly in the long run, than other anomalies. Third, contrary to conventional beliefs, the risk-reward tradeoff need not improve as the holding period lengthens. For professional fund managers, who worry about poor performance-induced outflows, risk measures such as Value at Risk (VaR) could be an important determinant of their preferences. Indeed, VaR measures the size of potential loss for a given probability, and such losses increase sharply as holding time lengthens. Many hedge funds (e.g., LTCM and Malvin Capital) eventually shut down following an extremely rare event of 50% losses due to the ensuing outflows.

We conduct several additional tests to validate the three considerations mentioned above. First, we examine the abnormal positions of short sellers and retail investors in anomaly strategies that retail investors trade against, in event time from 24 months before portfolio formation to 24 months after. We find short sellers move in lock steps with retail investors but exactly in the *opposite* direction. As retail investors slowly build up the wrong positions in anomalies, e.g., buy overpriced short-leg stocks, over the 24 months before formation, short sellers also slowly build up the right positions by shorting these stocks in the meantime. Both traders reach their peak level of positions right before the portfolio formation month. Then both sides start reducing their positions in the underlying stocks. The evolution of retail positions is fairly slow: after a long 48 months (4 years) of time, they go back to where they started, which is still half the size of their peak level! The change in retail positions in each event month is comparable in magnitude to that in short positions, supporting that the risk posed by retail trading is an economically relevant concern for short sellers. Overall, the pattern fits well with a general equilibrium where short sellers take the opposite side of retail investors over a long spell of time. Such long-term commitment to the risky position must be a daunting task for short sellers as we find that the average tenure of short positions for an *average* stock is merely 4 months (2 months) on an equal-(value-) weighted basis.

Second, we establish a linkage between the persistence of retail holding of anomaly portfolios post formation with the degree of their trading mistakes before formation. This comprises two findings. The first finding is that as we move from R1 (the least negatively traded anomalies by retail investors) to R5 (the most negatively traded anomalies by retail investors), the percentage of months over the sample period that retail ATI is negative monotonically increases from 29% to 77% of the time. To the extent that trading against an anomaly before formation is wrong, the result suggests that retail misactions before formation in the anomalies that they trade against is fairly persistent. The second finding is that the more retail investors misact before formation, the more persistent their holdings in the anomaly portfolio will be post formation. The pattern is only significant among the anomalies that retail investors trade against. This result suggests that they tend to stubbornly stick to their misactions even after disconfirming evidence starts piling in after formation as anomalies start realizing positive alphas. This stubbornness implies that the more retail investors trade against an anomaly, the longer short sellers will expect to hold a short position, and, thus, the higher long-run alpha the anomaly is expected to deliver.

Third, we examine the risk measures of anomaly strategies. For the long-run alphas of anomalies traded against by retail investors to survive, the risk-reward tradeoff of the anomaly strategies must be not as attractive as what the value of alphas suggests in some way. We find two patterns supporting this prediction. The first pattern is that the 24-month 1% VaR and 0.1% VaR of the quintile of the anomalies most negatively traded by retail investors (R5) is -42% and -62%, respectively. This level of loss would easily lead a hedge fund to shut down due to outflows. The 1% and 0.1% VaR thresholds we employ are conservative, considering, for example, that the event leading to the collapse of LTCM was a rare 10-sigma event.[9] Furthermore, the VaR of R5 anomalies are similar to that of R1 anomalies, which do not generate long-run alphas. This means holding R5 anomaly portfolios for 24 months is not necessarily more attractive than holding an anomaly portfolio amounting to pure noise, provided that VaR is an important determinant of arbitrageurs' risk management practice. The second pattern is that the *t* value of the monthly alphas of individual anomalies in R5 only infrequently cross the 1.96 bar over the 24 months post formation. The *t* value of the 24-month cumulative alpha is, however, considerably higher than 1.96. This difference is caused by the noise in alpha in each event month. The 24-month alpha is similar to a simple average alpha measure, which effectively diversifies away the noise of alpha in each event month, akin to diversification. Therefore, to achieve the substantial 24-month alpha with a high signal-to-noise ratio, arbitrageurs must sustain long-term positions. However, such time diversification need not be attractive in light of the VaR issue we documented earlier.

Finally, for retail trading to be a credible threat to arbitrageurs, there at least needs to be some evidence *suggesting* that retail trading can exacerbate anomalies. Such evidence includes two aspects. First, retail trading is powerful enough to exacerbate mispricing before formation, realizing negative anomaly alpha. Such evidence helps explain why anomalies realize positive alpha after formation. Second, retail investors may sometimes trade against anomalies even more after portfolio formation. This presents another threat for arbitrageurs who build positions in the anomaly ranking month as the noise trader sentiment may later intensify for exogenous reasons (Shleifer and Vishny 1997), implying an even longer holding horizon. We document four sets of findings supporting these conjectures.

The first set of findings pertains to our earlier results. When sorting anomalies, we use ATI, which is an unexpected shock that removes firm and anomaly fixed effects in retail trading. We also focus on a short, three-month pre-formation period, which alleviates persistent economic-wide omitted variables driving

[9] See p. 127 of Lowenstein, Roger. When Genius Failed. Random House, Inc. New York, NY. 2000. Furthermore, the case of LTCM is widely taught in business schools and known on Wall Street. The hedge fund industry's awareness of the critical need to safeguard against black-swan events has notably heightened in the post-LTCM era.

both ATI and returns. Therefore, for the anomalies traded against by retail investors (R3-R5), the finding that retail ATI and anomaly returns is negative before formation and turn positive after formation is consistent with the interpretation that retail investors misact before formation, exacerbating mispricing, which is then corrected by multiple forces slowly after formation. However, this evidence is still correlation-based.

Our second set of findings are based on a difference-in-difference (DiD) test. We utilize the Covid-19 stay-at-home period (March 2020 to December 2020) as a shock to retail trading. The dramatic rise in retail trading during this period helps shed light on the causal effect of retail investors (e.g., Ozik, Sadka, and Shen 2021). The rise can be attributed to several factors. First, the pandemic-induced shutdown placed a hold on many other activities, such as sports and travel. Much of the labor force had to work from home, offering greater flexibility during market hours. Second, the availability of zero-commission brokers with easy-to-use apps, fractional trading, and instant access to capital reduced barriers to investing (Welch, 2022). Third, social media also played a role as individual investors turned to forums such as Reddit\WallStreetBets for advice, which led to an increase in coordinated trading (Barber, Huang, Odean, and Schwarz, 2022). Fourth, stimulus checks further provide capital for retail trading.

To more cleanly support retail investors exacerbating anomalies, we narrow down to a set of treatment anomalies where we use the pre-pandemic persistent retail trading behavior as an instrument. The treatment anomalies (104 anomalies) are the ones that retail investors persistently trade against at least 60% of the months over the pre-pandemic decade. Since retail investors almost always trade against these anomalies before the pandemic, any increase in retail ATI due to the stay-at-home shock can be considered as an exogenous (and expected by an econometrician) increase unrelated to the Covid-related economic conditions. In the remaining anomalies, we remove those that retail investors persistently trade with around 60% of the months during the pre-pandemic decade, leaving the remaining ones as the control anomalies. These are the ones that retail investors show little interest in either trading against or trading with. We find that the ATI of the treatment anomalies is significantly negative during the pre-formation months and then slowly reverses to positive post-formation over multiple quarters. In contrast, there is little significant ATI in the control anomalies for any event-time month. The ATI and alpha patterns of the treatment anomalies are again consistent with the interpretation that retail traders exacerbate mispricing before formation, which then reverse slowly.

We then proceed to the DiD test. We find that the treatment anomalies with such expected increases in retail trading experience striking effects. First, as retail trading volume triples for an average stock, the pre-formation ATI turn three times more negative in the post-period, confirming the relevance of the shock. Second, the pre-formation anomaly alpha also becomes three times more negative. This pattern survives the parallel trend test. Third, the ATI remains significantly negative for the first quarter after portfolio formation, suggesting that the level of retail sentiment of holding overpriced stocks did not abate but even increase after anomaly ranking information is publicly available to potential arbitrageurs. Correspondingly, anomaly alphas remain negative even after portfolio formation for the first quarter. The bulk of the correction and hence positive anomaly alphas only start realizing from quarter 4, suggesting that noise trader risk can significantly increase the horizon risk exposure of arbitrageurs. Fifth, although the short-term alpha of treated anomalies decreases, the long-term alpha is much larger than the baseline period, consistent with our hypothesis that a longer-term risk exposure comes with a higher required rate of return.

In our third set of findings, we further exploit the events during the Covid period to sharpen the evidence that there is an exogenous increase in retail ATI in treated anomalies. We utilize the stimulus check event. The first paycheck was disbursed on April 9th, 2020 (Zimmerman and Divakaruni 2021) and shown to be the most effective one to affect market prices (Greenwood, Laarits, and Wurgler 2023). Every single person receives $1,200. Based on a regression-discontinuation design (RDD), we calculate two stock level ATI based on whether the dollar amount of a retail order falls into the range of [1150,1200] and (1200, 1250], respectively. We hypothesize that before paycheck disbursement, there exists a difference between treatment and control anomalies' ATI within each dollar-cost range; however, the difference should be the same across the two dollar-cost ranges. After the disbursement, as retail investors' wealth increases, they are more likely to increase trading in the dollar-cost range of [1150,1200] because the check payment cannot be used to buy a stock beyond $1200. Therefore, the difference in ATI between treatment and control anomalies should be more pronounced in the lower range. The change in difference should mainly come from the ATI change computed from buy trades. We find that this is indeed the case. A DiD test reveals that the pre-formation month treatment anomaly return indeed becomes significantly more negative at the daily level between April 9 and June 12 when the stimulus check is most likely to affect the stock market (Greenwood, Laarits, and Wurgler 2023).

In our fourth set of findings, we show that the treated anomaly alphas during the formation month, which is usually positive, turn significantly negative during not only the Covid period, but also the tech bubble period (1999-2000). Since both periods are characterized by retail frenzy, these results suggest that retail sentiment can indeed increase, exacerbating mispricing, after formation, inflicting more risk on performance-sensitive arbitrageurs.

Overall, our paper provides food for thought not only for understanding asset pricing, anomalies, and arbitrage profits, but also for policy makers and regulators debating whether and how to regulate retail trading. Many questions come up following recent events such as the GameStop saga. Should retail trading be banned or restricted? How to deal with social media, on which retail investors can coordinate their trades in the same direction, so that they form a formidable force strong enough to scare away the arbitrage sector? Is the development of Fintech such as zero commission trading a good thing for society at large? Our results suggest that certain forms of restriction or at least some financial education is needed for retail participation. After all, if financially unsophisticated investors stick to their beliefs/preferences for a long period of time, the financial market and real economy will both be adversely affected because short sellers can hardly take long-horizon risk and the long-run cost of capital is critical for firms' investment decisions. In countries where it maybe more difficult to educate and change stubborn retail investors, private markets may indeed be an important alternative to the public markets.

However, our results also imply that retail trading plays a dual role in the financial market. If retail view coincides with arbitrageurs' view such as in anomalies traded with by retail investors, regulatory changes that reduced other limits to arbitrage may be particularly effective in recent years. This results in positive alphas before the anomaly portfolio formation but little alphas post-formation.

Literature and Contribution

We contribute to understanding anomalies, retail trading, and their relation in several aspects. First, we document novel patterns of anomaly returns. Anomaly alphas predominantly materialize in the long term rather than in the portfolio formation month that the prior literature focuses on. These long-run alphas stay relatively stable over time even as the formation-month alphas disappear in recent years. These findings call for a fundamental re-focus of the anomaly literature on the long-run alpha of anomalies going forward, given the importance of long runs alphas and the diminishing formation-month returns. Second, to further highlight the importance of our finding for understanding asset pricing, we propose a new factor model, incorporating retail trading information, that outperforms prominent factor models in explaining the anomalies that remain significant in recent years. Our best performed model is the one that adds the retail factor to the mispricing factor model of Daniel, Hirshleifer, and Sun (2020), whose explanation power is largely due to a firm trading factor (i.e., a share issuance factor). To the extent that the stochastic discount factor is determined by market participants, our findings suggest that a future direction of developing asset pricing models could be aggregating information from important groups of traders. This direction differs from the traditional way of building factors from accounting variables (e.g., Fama and French 2015; Hou, Xue, and Zhang 2020). Third, we empirically support a unique horizon-based risk that retail trading imposes on arbitrageurs at the anomaly level. We show that retail investors are stubborn in their misactions in trading anomalies, and such misactions can exacerbate following exogenous retail trading shocks. As a result, post-formation anomaly returns can even turn negative for extended periods before they turn positive again. Given the stochastic discount Fourth, we establish that retail trading plays a crucial role in shaping anomaly returns. The anomalies that retail investors trade against are entirely responsible for the alphas of anomalies in the recent decade, especially for the long-run alpha. We provide causal evidence that retail investors exacerbate a significant number of anomalies before portfolio formation.

Existing evidence that connects retail trading with anomalies is limited. The first study exploring this relation is McLean, Pontiff, and Reilly (2021 MPR), which finds that retail trading is negatively associated with an aggregate anomaly signal. MPR study how well anomaly returns explain retail investor performance. We focus on explaining the recent trend in anomaly risk-adjusted returns, introduce a horizon-based risk for arbitrageurs, and propose a new factor model. We consider individual anomalies and find that retail investors trade with some anomalies and against others. We shed light on the causal effect of retail trading on anomaly performance and thus characterize which anomalies are exacerbated by retail investors (and which ones aren't).

The retail trading literature is debating whether retail trading is smart or dumb. In particular, the retail trading measure we use has been shown to positively predict short-term returns with no obvious reversal. We shed light on this debate by examining short-run and very long-run *risk-adjusted* returns. Our findings need not be against any of the mixed results provided in the literature. Indeed, we find retail investors trade with and against 40% and 60% of the anomalies, respectively, providing evidence of them being smart and dumb under different circumstances. Since individual anomalies account for a small portion of the return variance, combining all evidence across anomalies may not significantly affect the aggregate retail trading smartness. However, our findings shed light on how retail trading can be smart or dumb w.r.t different anomaly characteristics. We also explore why examining risk-adjusted returns makes a difference.

2. Sample Construction and Descriptive Statistics

2.1 Anomaly Sample Construction and Measuring Anomaly Performance

We consolidate a dataset of 260 anomalies sourcing from those studied by Dong, Li, Rapach, and Zhou (2022), Jensen, Kelly, and Pedersen (2023), Stambaugh, Yu, and Yuan (2012, 2015), and Green, Hand, and Zhang (2017) and remove any duplicates with similar construction methods. We replicate anomaly portfolio returns using publicly available data from CRSP (Center for Research for Security Prices), Compustat, and

I/B/E/S. To avoid microstructure issues, we only focus on common stocks listed on the NYSE, AMEX, and Nasdaq and exclude stocks priced below $1 in the preceding month.

For each month t (ranking month), we calculate the sorting variable for each anomaly based on the information at the end of month t. This calculation presumes the availability of annual accounting data if a firm's fiscal year concluded at least three months prior, and quarterly accounting data is considered immediately accessible upon the issuance of the quarterly earnings report. Given that some accounting information is available during earnings announcement, which usually happens before the 10-K report [10], a three-month time lag between the fiscal year-end date and signal formation ensures timeliness. Monthly stock returns are adjusted for delisting. The long-short anomaly portfolio goes long on the tenth decile and short on the first decile, with the expectation that the long leg will generate relatively high returns and the short leg will generate low returns.

To better characterize anomaly returns dynamics both before and after anomaly formation, we use the event-time approach by holding anomaly portfolio composition constant a quarter before and two years after anomaly formation and examining anomaly portfolio risk-adjusted returns. We recognize the potential for inaccuracies in statistical inference due to overlapping holding periods, a concern that persists despite the application of Newey-West adjustments for serial correlation. Consequently, we incorporate a long horizon return calculation method as proposed by Jagadeesh and Titman (1993). In this alternative method, for any given month t, the h holding-horizon monthly return is calculated as the average monthly risk-adjusted return at month t from ranking date t-h, $\alpha_{t-\mathrm{h,t}}$ to the ranking date t-1, $\alpha_{t-1,\mathrm{t}}$. The h-holding monthly return is constructed from h long/short portfolios. We then compound h-holding monthly returns by h months to get the cumulative holding period return.

Table 1, Panel A compares anomaly portfolio returns for different sub-samples. Specifically, it examines the shift in cumulative CAPM-adjusted anomaly returns from the pre-pandemic to the during

[10] See, for instance: Bowles, Reed, Ringgenber, and Thornock (2023).

COVID-19 pandemic periods. The cumulative CAPM-adjusted anomaly returns a quarter before anomaly formation (α_{t-2tot}) declines from the pre-pandemic value of 0.07% to the during-pandemic value of -2.75%. The average formation month α_{t+1} declines from the pre-pandemic value of 0.29% to the during-pandemic of -0.40% under the value-weight method. Furthermore, this downtrend persists beyond the formation of the anomaly, with risk-adjusted returns remaining negative on average six months post-anomaly formation. This contrasts sharply with the pre-pandemic period, where after-ranking risk-adjusted anomaly returns were positive. Raw anomaly portfolio returns display a similar pattern. Pre-pandemic anomaly portfolio raw returns are generally higher compared to during-pandemic numbers, for a holding horizon of up to 12 months. In sum, during the pandemic, the reversal shows up in much later periods and is stronger than the pre-pandemic value.

2.2 Measuring Retail Trading and Holding at the Stock Level

Retail investor trade data comes from TAQ (Trade and Quote). According to Boehmer, Jones, Zhang, and Zhang (2021), Reg NMS (Regulation National Market System) requires a broker/dealer to give a slight price improvement relative to the National Best Bid or Offer (NBBO) to retail orders. We follow the method of price improvement to isolate retail investors' marketable orders from institutional orders. Specifically, we identify a transaction as a retail buy if the sub-penny price is between 60 and 100 basis points and a retail sell if the sub-penny price is between 0 and 40 basis points. We then take the difference between buy and sell volumes to get retail net buying, denoted as NB_t. The abnormal retail trading measure takes the difference between the current month NB and the average NB in the past 12 months and is scaled by shares outstanding.

$$RetailTrading_t = \frac{NB_t - \sum_{j=t-1}^{j=t-12} NB_j/12}{ShrOut_t}$$

We detrend retail trading using the past 12-month average net buy precisely because there has been an upward trend in retail trading volume in the recent decade. In addition, given the unprecedented retail volume increase during the pandemic (from March 2020 to December 2020), we calculate abnormal trading during the COVID-19 pandemic by detrending the 2019 average monthly NB. To avoid the problem of outliers, we winsorize stock level trading using the whole panel at 0.5% and 99.5%. Our sample period, defined as anomaly ranking months, starts from January 2009 to December 2020[11].

Table 1, Panel B shows monthly stock level retail trading intensity. On average, retail investors tend to purchase stocks included in our sample. The average retail stock level monthly trading intensity is 0.003%, with the median value of -0.002%.

We also model the behavior of 'noise traders' in line with the demand-based frameworks established by De Long, Shleifer, Summers, and Waldmann (1990) and Shleifer and Vishny (1997). We compute retail abnormal holding as a weighted average of retail trading activity.

Let H_t denote retail holding of a stock at time t, and NB_t denote retail net buying. The transformation from the weighted average of trading to the 12-month abnormal holding can be expressed as follows.

$$RetailHolding_t = H_t - \frac{(H_{t-1}+H_{t-2}+\cdots+H_{t-12})}{12}$$

$$= H_t - H_{t-1} + \left(\frac{11}{12}H_{t-1} - \frac{11}{12}H_{t-2}\right) + \left(\frac{10}{12}H_{t-2} - \frac{10}{12}H_{t-3}\right) + \cdots + \left(\frac{1}{12}H_{t-11} - \frac{1}{12}H_{t-12}\right)$$

$$= NB_t + \frac{11}{12}NB_{t-1} + \frac{10}{12}NB_{t-2} + \cdots + \frac{1}{12}NB_{t-11}$$

The above abnormal holding formulation captures the cumulative impact of retail trading over 12 months, weighted to reflect more recent transactions more heavily. It provides a nuanced view of retail holding patterns, crucial for understanding the dynamics of retail behavior in the market.

Table 1, Panel B shows monthly stock level retail holding intensity. On average, retail investors have a positive holding position of 0.098%.

2.3 Measuring Retail Trading and Holding at the Anomaly Level

Our study primarily focuses on how retail investors trade on anomalies. To assess this, we construct retail anomaly trading and holding intensity, defined as the average trading or holding in decile-10 (underpriced)

[11] As written in BJZZ (2021, P.2253), "During the first few years under Reg NMS, there is a strong upward trend in the number of sub-penny trades. The upward trend disappears and stabilizes after 2009".

stocks of the anomaly minus that of the decile-1 (overpriced) stocks. Specifically, for each anomaly, retail anomaly trading and holding intensity are constructed as below.

$$ATI_{it}/AHI_{it} = L_{it} - S_{it}$$

where L_{it} and S_{it} are the average $RetailTrading$ or $RetailHolding$ in the long and short leg stocks. A negative value of ATI_{it} suggests that retail investors buy or hold more (or sell less) decile-1 stocks than decile-10 stocks. Retail investors appear to trade in the "wrong" direction to exacerbate anomaly mispricing. Similarly, a positive value can be interpreted as retail investors trading in the "right" direction.

Table 1, Panel C shows the retail ATI and retail AHI a quarter before anomaly formation with the sample period from January 2009 to December 2020, by summing up monthly ATI or AHI. Prior to the pandemic, the average ATI is -0.017%, with a median of -0.007%, suggesting that average retail investors trade against anomalies before anomaly formation. However, a significant shift is observed during the pandemic: the average ATI falls to -0.145%, which is approximately eight times the pre-pandemic level. Additionally, there is a notable increase in the standard deviation of ATI during the pandemic, which increases from the pre-pandemic value of 0.18% to 0.57%. Such an increase in standard deviation underscores a heightened arbitrage risk, suggesting greater unpredictability in retail trading patterns during the pandemic.

2.4 Measuring Short Seller Holding at the Stock and Anomaly Level

We construct short-seller anomaly intensity using Compustat short interest holding data to understand how retail investors pose a long-horizon risk to arbitrageurs. First, at the stock level, short-seller abnormal holding intensity is calculated using the difference between current mid-month short interest and average short interest in the past 12 months and is scaled by shares outstanding. We invert the sign of short interest directly from Compustat; thus, a more negative value indicates higher short interest.

$$ShortSellerHolding_t = -\frac{Short\ Interest_t - \sum_{j=t-1}^{j=t-12} Short\ Interest_j/12}{ShrOut_t}$$

Consistent with our methodology for retail stock level trading, the short seller holding during the pandemic is calculated by normalizing against the average monthly short interest from 2019.

The short-seller anomaly holding intensity is constructed as follows:

$$AHI_{it} = L_{it} - S_{it}$$

Where L_{it} and S_{it} are the average $ShortSellerHolding$ in the long and short leg stocks. A positive value of short seller AHI_{it} suggests that short sellers hold more (or are less inclined to sell) decile-10 (underpriced) stocks than decile-1 (overpriced) stocks. That is, short sellers are acting in alignment with correcting the anomalies, trading in the 'right' direction by targeting underpriced stocks for holding and overpriced stocks for selling. Conversely, a negative value implies that arbitrageurs are exacerbating anomaly mispricing by trading in the 'wrong' direction.

Table 1, Panel B shows short seller stock level holding intensity. On average, short sellers in our sample are predominantly engaged in selling activities. The average short interest is -0.14% with a standard deviation of 3.37%. Panel C shows the short seller AHI a quarter prior to anomaly formation. This intensity is calculated by aggregating the monthly short seller AHI. Overall, short sellers tend to trade in the "right" direction of anomalies, selling overpriced stocks and buying underpriced stocks. The average short-seller AHI is 0.91% before the pandemic and increases to 1.77% during the pandemic, potentially indicative of an increase in mispricing amid heightened volatility and uncertainty.

3. Long-Horizon Alpha and Retail Anomaly Trading

In this section, our analysis starts with an exploration of the performance of anomalies over the past half-century. We discover that long-horizon anomaly alpha has never disappeared, a phenomenon largely due to those anomalies traded 'wrong' by retail investors. Using retail trading data from 2009, we next provide more direct evidence by sorting 260 anomalies according to retail ATI in recent years. This detailed analysis reveals a striking insight: anomaly only works in the long horizon if retail investors trade against them. Finally, we refine our analysis by categorizing anomalies into different anomaly groups from existing literature to better characterize retail anomaly trading.

3.1 Anomaly Short and Long-Term Performance Over Past Half-Century

Many studies have discovered anomaly portfolio returns have become statistically or economically insignificant in recent years (Green, Hand, and Zhang, 2017) or largely disappear after publication (Mclean and Pontiff, 2016). However, the anomaly literature generally focuses on the formation month while largely ignoring the long-run profitability (Baba-Yara, Boons, and Tamoni (2020)). We first document that most anomaly alpha is realized in the long horizon. Figure 1, Panel A shows the historical value weighted CAPM alpha of 260 anomalies. We divide the sample period, defined as the months when anomalous signals are ranked, into three distinct intervals. The first interval is from 1969 to 1988, a 20-year period that begins with the empirical testing of market beta as an explanatory variable for the cross-section of expected returns, as noted by Harvey, Liu, and Zhu (2016). The second interval covers the subsequent 20 years, from 1989 to 2008. The final interval extends from 2009 to February 2018. We intend to exclude the COVID-19 period in the recent decade, reserving it for a separate event study. We find the formation-month (t+1) alpha indeed monotonically declines over time with the percentage of significant anomalies (t>1.96) being more than halved from 55% to 23%, and Figure 2, Panel B shows the magnitude of alpha also decreases to 0.35% per month. However, the 24-month alpha remains stable both in number and magnitude. The number of anomalies with significant 24-month alpha is around 40% throughout, and the magnitude remains roughly at 9%. This finding is consistent with Jensen, Kelly, and Pedersen's (2022) argument that one should focus on alpha but not return. Overall, the result suggests that the vast majority of anomaly alpha is realized in the months after the formation month. In the long horizon, anomalies can realize a large alpha, making them anomalous to the CAPM model.

We then separate anomalies based on the retail ATI observed a quarter prior to the formation month. This separation uses retail trading data from 2009, thereby positioning the periods 1969-1988 and 1989-2008 as out-of-sample tests. If retail investors buy more (less) stocks in the short leg than in the long leg of an anomaly, we classify this anomaly into the group where retail investors trade against (with). We find that retail investors trade against 60% of all anomalies. Figure 2, Panel A1 illustrates that among the 40%

of anomalies that retail investors trade with, there is a marked decrease in the number of anomalies with significant one-month alpha, falling sharply from 52% to 9%. A similar trend is observed for long-term alpha; the number of anomalies with significant 24-month alpha drops from 41% to 13%. Contrastingly, Panel A2 indicates a more modest reduction in the one-month alpha of anomalies traded against by retail investors, declining from 57% to 33% between 1969-1988 and 2009-2018. However, there is an increase in the percentage of significant 24-month alpha, rising from 40% to 59% during the same periods12.

Regarding magnitude, Panel B1 reveals a decline in the one-month alpha for anomalies that are traded positively by retail investors, dropping from 0.53% to 0.10%. The 24-month alpha similarly diminishes, falling from 6% to a negative number of -0.62%. In sharp contrast, Panel B2 highlights that for the anomalies traded against by retail investors, the average formation-month alpha experiences a slight decline from 0.60% per month in 1969-1988 to 0.51% in 2009-2018. More notably, the 24-month alpha increases from 10.26% to 14.44%. These findings suggest that increasing market efficiency over time impacts the short-term and long-term alphas of anomalies traded positively by retail investors. However, it has little effect, especially on the long-term profitability of the anomalies that retail investors trade against.

3.2 Anomaly Performance and Retail Anomaly Trading in Recent Years

Beginning our analysis in 2009 and utilizing direct retail trading data, we examine the comparative performance of anomalies that are either favored or opposed by retail investors. We start by calculating t-statistics from the time-series retail ATI a quarter before anomaly formation for each anomaly. This analysis uses the sample period from 2009 to February 2018, defined as the anomaly ranking months, *deliberately omitting the COVID-19 period for a dedicated event study. Anomalies are then sorted into five groups,*

[12] This pattern emerges because, first, as evidenced in Figure IA2 of the Internet Appendix, retail ATI becomes more negative in recent years, and long-horizon alpha is more pronounced. Second, the anomalies themselves are selected into retail trade with vs. against using the retail trading date in recent years.

labeled R1 through R5, according to their t-statistics. We then examine cumulative risk-adjusted anomaly portfolio returns from a quarter before to two years after formation.

Table 2, Panel A column (2) shows the average ATI for each group. From R1 to R5, retail ATI decreases monotonically. Retail investors trade with anomalies in R1 and R2 and trade against those in R3 to R5. *By averaging retail ATI and anomaly portfolio returns across anomalies within each group, we aim to reduce individual anomaly noise and discern more distinct patterns.* **Column (1)** sheds light on the persistence of retail anomaly trading by counting the months with negative trading. For instance, in group R5, where retail investors trade mostly negatively, they trade against the anomalies 77% of the time on average. This column also shows *that increased negative trading by retail investors is associated with more persistent negative trading behaviors.*

CAPM-adjusted anomaly portfolio returns are calculated by excluding the part of the return that can be attributed to market risk. For each h-period monthly return at the ranking month t, denoted as $\mathrm{Ret_{t+h}}$, *we first align the anomaly portfolio return with the market premium as of the return date.* Then, we regress anomaly portfolio raw returns on market premium to get risk loading on the market risk β_h. The alpha is calculated as the difference between raw return and CAPM-explained return $\alpha_{t+h} = \mathrm{Ret_{t+h}} - \beta_h \times MktRf_{t+h}$. *Cumulative monthly risk-adjusted returns are obtained by multiplying these monthly values, represented as* $\alpha_{t+1tot+h} = \prod_{j=1}^{h}(1 + \alpha_{t+j})$. *To account for serial correlation, we adjust the t-statistics of* $\alpha_{t+1tot+h}$ *using the Newey-West method, with adjustments for 6, 12, and 24 lags* for $\alpha_{t+1tot+6}$, $\alpha_{t+1tot+12}$, and $\alpha_{t+1tot+24}$, respectively.

Table 2, Panel A columns (3) to (8) show cross-sectional anomaly performance. We report results based on equal breakpoint sorting, with robustness checks using NYSE breakpoint. Column (3) demonstrates retail investors contribute to mispricing before anomaly formation. The cumulative alpha a quarter before anomaly formation decreases monotonically as retail ATI becomes negative. For anomalies in the R1 group, α_{t-2tot} is significantly positive at 4.03%. In contrast, for R4 and R5 groups, α_{t-2tot} are significantly negative at -3.54% and -3.68%.

Columns (4) to (7) show anomaly alphas at the formation month and at 6, 12, and 24 months post-formation. We observe an immediate reversal in anomaly portfolio alphas following their ranking. Cross-sectional anomaly portfolio alphas post-formation are positively significant only when retail investors trade against them. For instance, the formation month alpha for R5 anomalies stands at 0.83%, compared to an insignificant 0.07% for R1 anomalies. Moreover, the 24-month anomaly portfolio alphas exhibit a monotonic pattern: the more negative the retail $\mathrm{ATI}_{\mathrm{t-2tot}}$, the higher the alpha. The highest 24-month alpha reaches 23.49%. This pattern is particularly striking, given the expectations of immediate mispricing correction due to rapid information distribution, increased arbitrage trading, and improved liquidity. However, our findings indicate that most of the corrections occur over an extended period. Therefore, our results suggest retail traders help contribute to anomaly mispricing before formation, followed by a pronounced and sustained reversal in anomalies predominantly traded negatively by retail investors.

We also apply the methodology outlined by Jagadeesh and Titman (1993), as detailed in section 2.1, to address issues related to overlapping holding periods. We compound *h*-holding monthly returns by *h* months to get the cumulative holding period return, and t-statistics are derived from monthly alpha. Panel B shows the result. The results align closely with those in Panel A, both in terms of magnitude and significance levels. This consistency reinforces the robustness of our findings against potential biases in statistical inferences from overlapping holding periods.

In light of the results discussed in this section, a pertinent question arises: Do the long-term profits of anomalies in group R5 diminish when transaction costs are taken into account? Muravyev, Pearson, and Pollet (2023) report that anomalies cease to be profitable when high-fee observations are excluded. Following their methodology, we employ the 'Indicative Fee' from the Makit short-selling dataset for short-selling costs. We filter out stocks whose fee exceeds 1% per annum as of the end of the previous month. As shown in Table IA2 of the internet appendix, although the monthly alpha declines when high-fee stock observations are excluded—with R5 anomaly formation-month alpha decreasing from 0.83% to 0.53%—the 24-month alpha for the R5 group anomalies remains robust at 17.29%.

3.3 Retail Anomaly Trading Group Characteristics

To understand what anomaly characteristics retail investors tend to trade against or with, and to address the possibility that their trading might be arbitrary, we group anomalies according to different economic concepts identified in previous studies. This involves computing the group-level retail ATI by averaging the ATI a quarter before anomaly formation across anomalies within each group.

Table 3 shows the result. Initially, we segregate all anomalies into subgroups following the categorizations of Hou, Xue, and Zhang (2020). Our analysis reveals that retail investors strongly trade against anomalies related to profitability, trading frictions, and value versus growth, while their engagement with anomalies related to momentum, investment, and intangible-related anomalies appears comparatively weaker. The results suggest that retail investors may have biased beliefs toward profitability anomalies, consistent with the theory of Bouchaud, Krueger, Landier, and Thesmar (2019). Additionally, retail investors may also have mistaken preferences and beliefs in trading frictions-based anomalies, which are inherently challenging for arbitrageurs to correct. The results also suggest retail investors trade against and with several dimensions of characteristic concepts with no single concept being the dominant driver.

Daniel, Hirshleifer, and Sun (2020) argue that the mispricing of some anomalies is more persistent than others, categorizing anomalies into short- and long-horizon types. The short-horizon is earnings surprise factors, whereas the long-horizon exploits the information in the manager's decisions to issue or repurchase equity. Our results align with this classification: retail investors trade against the long-horizon anomalies while trade with short-horizon ones and is consistent with our earlier finding that the anomalies traded against by retail investors exhibit significant long-horizon alpha. Furthermore, our finding that retail investors tend to trade with earnings surprise factors corroborates the insights of Kaniel, Liu, Saar, and Titman (2012). They demonstrated that retail investor buying (or selling) is predictive of positive (or negative) abnormal returns surrounding earnings announcement dates.

For anomalies that Akbas et al. (2015) identify as investment and non-investment anomalies in the 11 Stambaugh, Yu, and Yuan (SYY, 2015), retail investors strongly trade against non-investment anomalies while trading with investment anomalies. This suggests that retail trading resembles mutual fund flow-induced trading, which also shows a tendency against noninvestment anomalies (Akbas et al. 2015; Dong, Kang, and Peress 2023), consistent with the fact that mutual fund flows are retail-driven.

We also divide our anomalies into those identified as lottery-demand-driven anomalies (e.g., Kumar 2009) and the rest, which we label as non-lottery anomalies. Retail investors trade against both groups of anomalies but more strongly against lottery anomalies, consistent with the literature finding that retail investors have a lottery preference (Han and Kumar 2013). Finally, we find that among the anomalies that Birru (2018) identifies as having clear speculative short and long legs, retail investors strongly trade against the anomalies having speculative short legs, such as idiosyncratic and profitability anomalies. In contrast, they trade with those having speculative long legs, like size and illiquidity anomalies.

In summary, our findings suggest that retail investors engage in trading against and with various dimensions of characteristic concepts, indicating that no single concept predominantly drives their trading behavior.

4. Explaining Long-Horizon Returns: Two Retail-Related Factor Models

The long-horizon return is crucial for several reasons. First, the typical holding horizon for investors extends beyond one month. Second, as Cohen, Polk, and Vuolteenaho (2009) argue, firms' capital budgeting decisions hinge on discounting long-term cash flows; inaccuracies in the discount rate can lead to significant capital misallocation. Pricing long-horizon returns, however, can be challenging. Recent research indicates that commonly used multifactor models fail to adequately explain expected returns over extended periods (Chernov, Lochstoer, and Lundeby, 2022; Gormsen and Lazarus, 2023; Binsbergen, Boons, Opp, and Tamoni, 2023; Bara-Yara, Boons, and Tamoni, 2023). Our previous findings underscore the influence of retail trading on the long-term performance of anomalies. Specifically, we observed that the 24-month CAPM-adjusted alpha for anomalies predominantly traded against by retail investors can reach nearly 24%. Building on the insights, we propose two retail-related factors. The first factor is derived from the formation-month return of R5 anomalies, while the second factor is based directly on retail trading. We then add a market risk premium to make them factor models to capture the long-horizon return dynamics.

4.1 R5F Retail Factor Construction

The first retail factor, labeled R5F, is derived from the formation-month return of R5 anomalies, predominantly traded against by retail investors. To construct this factor, we first compute the average anomaly rankings within R5 anomalies. To control for the size effect, we employ a double sorting method following Fama and French (1993). Specifically, we independently sort size into 2 groups—small (S) and big (B)—based on the NYSE breakpoint and sort the average rankings from R5 anomalies into low (L), middle (M), and high (H) groups. A size-controlled, value-weighted factor return is obtained as the average return of the high-average-ranking portfolio minus that of the low-average-ranking portfolio:

$$R5F = (r_{SH} + r_{BH})/2 - (r_{SL} + r_{BL})/2$$

The intuition is straightforward. The R5 anomalies predominantly experience persistent buying by retail investors. Consequently, it is expected that these anomalies will prominently reflect the long-horizon retail risk and its corresponding risk premium. By averaging across anomalies, we effectively cancel out noise and potential other risk premiums that are not related to long-horizon retail risk.

4.2 TF Retail Factor Construction

The second retail factor, denoted TF, is derived using market-wide retail trading, largely following the methodology established by Pastor and Stambaugh (2003). We construct the TF factor using daily trading data. Aligning with monthly detrended retail trading, we detrend daily trading. This involves detrending the current day's trading—quantified as the difference between the number of shares bought and sold—against the prior 252-day moving average and scaling it by number of shares outstanding. To address potential noise introduced by daily overnight trading, we ensure that the volume of overnight shares bought or sold does not surpass that of intraday. Additionally, we winsorize the cross-section at 0.5% and 99.5% levels to avoid outliers. Our TF factor yields even stronger results with retail daily intraday trading. We posit that systematic retail risk is more likely to be priced in the market compared to idiosyncratic risk, given arbitragers' ability to diversify.

Firstly, we calculate daily market-wide average retail trading (MWRT) across all stocks i:

$$MWRT_t = \sum_i RetailTrading_{it}$$

We then derive the market-wide trading shock using the residual from an AR(1) model, denoted by η:

$$MWRT_t = \gamma MWRT_{t-1} + \eta_t$$

Utilizing daily data within a month, we compute the monthly sensitivity of each stock's return to the market-wide trading shock, denoted as β. This involves running a regression for each stock:

$$R_t = \beta_{0t} Mkt + \beta_t \eta_t + \epsilon_t$$

Next, at each ranking month, we sort on the retail trading β and form monthly long-short portfolio 24 months after ranking. The TF factor at each return date is calculated as the average return of these 24 portfolios (Jagadeesh and Titman, 1993).

The TF factor is fundamentally rooted in understanding the interplay between retail beta and the marginal utility of short sellers. Stocks with high retail beta are more susceptible to fluctuations in retail sentiment. In the scenario where retail sentiment is heightened, arbitragers— who typically act as a counterforce to retail trading—experience losses. This increase in losses leads to a rise in their marginal utility. Consequently, to offset this heightened risk, arbitragers demand higher compensation for shorting these high retail beta stocks. As a result, we anticipate that stocks with a higher retail beta will generally yield lower returns. These dynamics form the core rationale for the TF factor, linking retail trading directly to stock expected return through the lens of arbitrage activity.

4.3 Factors and Factor Models Performance

We first run a spanning test to examine how well other commonly used factor models can explain our two retail-related factors, R5F and TF, and vice versa. Using sample period from 2011 to 2022, defined as the portfolio return dates, we run a time-series regression of FIN, a long-horizon behavioral factor from Daniel, Hershleifer, and Sun (2020), R5F, and TF on different factor models: capital asset pricing model, denoted CAPM; the three- and five-factor model of Fama and French (1993) and Fama and French (2015), denoted FF3 and FF5; the four-factor model of Carhart (1997), denoted Carhart4; the short and long-horizon behavioral model (DHS); and our two retail-related factor models. Since the FIN factor in DHS is only available up to 2018, we constructed a size-controlled FIN factor following the DHS methodology. This process entails averaging the rankings of two anomalies: the 1-year net share issuance and the 5-year composite share issuance; and applying a double sort based on size, using the NYSE breakpoint and the average rankings. Monthly returns for the FIN factor are determined by calculating the difference between the average returns of portfolios with low and high share issuance.

Table 4, Panel A reports the alphas from the spanning test. Columns (6) and (7) reveal that our two retail-related models, R5F and TF, effectively explain the FIN factor, as indicated by their insignificant alphas. In contrast, Column (5) shows that the long and short-term behavioral model fails to account for these two retail-related factors. Furthermore, the R5F factor remains largely unexplained by standard models like CAPM, FF3, Carhart4, FF5, and DHS, with the only exception being the TF retail factor model. Notably, none of the factor models can explain the TF factor. This finding underscores the unique explanatory power of our retail-related factors.

Next, we evaluate whether the proposed retail-related two factor models can price long-horizon anomaly portfolio returns, especially in the R4 and R5 groups of anomalies that retail investors trade against in Table 2. In total, 104 testing anomaly portfolios are used. Table 4 Panel B reports the average absolute alphas, average t-statistics, and number of anomalies with significant alphas (t-statistics >= 1.96) from each factor model under consideration. The results show that pricing errors decrease monotonically as more factors are included: CAPM (0.74%), FF3 (0.53%), Carhart (0.46%), and FF5 (0.38%). While the pricing error for the DHS model is the same as that of FF5, the number of significant anomalies is more than halved compared to FF5. Panel B highlights the superior performance of our two retail-related factor models, evident in the lowest pricing errors, average t-statistics, and the fewest anomalies with significant alphas.

In comparison, the pricing errors for FF5 and DHS are nearly 50% higher than those for the retail-related models. The analysis further reveals that incorporating R5F or TF into the DHS model—originally based on firm trading—markedly improves its performance. This is evidenced in the last two rows, where the addition of our retail trading-based factors results in a significant reduction of the pricing error, from 0.38% to 0.19%. This substantial improvement underscores a key insight: factors derived from retail trading hold independent distinct information from long-horizon factors based on firm trading. The three-factor model, which combines these diverse sources of trading information, demonstrates robust performance in the recent sample period, outperforming all other factor models.

5. Persistent Retail Trading and the Long-Horizon Risk

Why do the long-horizon alphas of the anomalies that retail investors trade against thrive? Based on our findings so far, we hypothesize that retail investors have stubborn preferences/beliefs, inflicting long-horizon risks on arbitrageurs countering them. Retail investors' stubbornness may stem from preferences, for example, for skewness, as suggested by Bali, Hirshleifer, Peng, and Tang (2021) and Barberis and Huang (2008), or from serially correlated misperceptions as noted by De Long, Shleifer, Summers, and Waldman (1998). Such tendencies can lead to significant deviations in stock prices from their fundamental values. In this section, we investigate whether retail investors are inclined to persist in their investment errors by analyzing their holding patterns 24 months before and after anomaly formation. If retail investors tend to maintain their positions despite market signals, this behavior can sustain the risk of capital loss for arbitragers and delay the reversion of prices to fundamental values. To further understand the risks faced by arbitragers, we also examine short-seller holdings during the same periods. We focus on their positions in anomalies that are traded against by retail investors. Our analysis reveals that arbitragers' positions in these anomalies remain elevated even 24 months after their formation, highlighting the persistent nature of the risk they face in such scenarios.

5.1 Stubborn Retail Anomaly Holding

Figure 3, Panel A plots retail anomaly holding intensity for the 60% of anomalies that retail investors trade against over the sample period from 2009 to 2020, defined as anomaly ranking date. In the 24-month periods leading up to the anomaly formation, there is a noticeable buildup of negative AHI values. This decrease in negative position underscores a continuing acquisition of stocks in the short leg by retail investors. After the anomaly formation, the AHI values slowly ascend to 0, a gradual divestment from the previously accumulated positions. Even 24 months later, retail AHI is still half the size of the peak value around anomaly formation. The protracted nature of the unwinding process post-anomaly formation emphasizes retail investors' reluctance to swiftly rectify their positions, highlighting a potential hesitancy to acknowledge and act upon signals that contradict their initial trading decisions.

We further explore the notion of retail investors' misguided persistence and examine the direct correlation between their pre-formation ATI and the adherence to their erroneous positions following anomaly formation. Specifically, we investigate the relationship between cross-sectional standardized retail ATI a quarter before anomaly formation (ATI_{t-2tot}) and the post-formation retail holding persistence. Using an event window of 24 months post-formation, we construct anomaly retail holding persistence measures as the AR(1) coefficient by regressing subsequent event time AHI on current event time AHI.

Table 5 Column (1) shows across anomalies that retail investors trade against, the more negative the ATI_{t-2tot} is, the more persistent the after-formation holding intensity becomes. For these anomalies, the average AR(1) coefficient of holding is 0.92, corresponding to a half-life of 8.3 months[13]. A one standard deviation decrease in ATI is associated with a 0.008 increase in holding persistence, equivalent to an additional 1 month in half-life. Column (2), focusing on anomalies that retail investors trade with, shows no significant result.

---

[13] the half-life calculation is from half-life = $-\frac{\ln(2)}{\ln(|\emptyset|)}$, where $\emptyset$ is the AR(1) coefficient.

Overall, the persistent holding patterns of retail investors, despite potential losses or contradictory public signals, can lead to sustained mispricing in the market, elevating the risks for arbitragers targeting these anomalies. Arbitragers not only face fundamental risks but also prolonged noise trader risks, as they navigate the protracted price correction delayed by retail stubbornness.

In contrast, Figure IA2 in the Internet Appendix illustrates the anomaly trading intensity (ATI) for the remaining 40% of anomalies that retail investors trade with. Retail investors tend to trade most positively before anomaly formation and gradually decrease the buying to nonexistence after less than two quarters. Since retail investors trade in the same direction as arbitragers' anomaly trading, the long-horizon risk is effectively absent. Consequently, the performance of these anomalies is short-lived.

5.2 Short Seller Anomaly Holding and the Long-Horizon Risk

Alternatively, Figure 3, Panel B shows short seller anomaly holding in response to anomalies that are traded against by retail investors. We find that short sellers' holding counterforces that of retail investors. Prior to the anomaly formation, short sellers progressively accumulate positions, as retail investors push up asset prices above fundamental values in the short-leg stocks. This build-up suggests anticipatory actions by short sellers in expectation of future price corrections. Following anomaly formation, while AHI values exhibit a decline, short seller positions remain markedly elevated, even two years post-formation. This sustained high holding accentuates the long-horizon risks undertaken by short sellers, as fundamental risk, noise trader risk, and transaction costs accumulate as holding horizon lengthens.

Complementing our descriptive analysis of long-horizon risk, we introduce a more tangible metric—Value at Risk (VaR)—to gauge the potential risk arbitrageurs may encounter when their trading strategies counter those of retail investors, especially over a 24-month holding period. VaR quantifies the potential size of loss at a given probability, with the magnitude of such losses escalating as the holding period extends. For professional fund managers, particularly sensitive to the risk of outflows stemming from poor performance, VaR is a critical metric, informing their decision-making processes. We compute 1% and 0.1% VaR for the long-term alpha ($\alpha_{t+2tot+24}$), using parametric method. This approach delineates VaR as the difference between the time-series mean of alphas and the product of z-score for the designated confidence level and standard deviation of anomaly portfolio alphas. The formula is $VaR = \overline{\alpha} - z - score \times \sigma_{\alpha}$, with z-scores of 2.33 for 1% and 3.09 for 0.1% VaR. Table 6 columns (4) and (5) show that the estimated 1% and 0.1% VaR for the long-term alpha of an average anomaly within the R5 group can reach staggering -42% and -62% respectively.

These findings illuminate the stark reality that the risk-reward tradeoff inherent in the R5 anomaly strategies may not be as compelling as the long-term alpha values suggest. This substantial VaR not only underscores the severe risk arbitrageurs assume when opposing retail investors over extended investment horizons, but also casts doubt on the viability of long-term anomaly-based strategies.

6. Three Event Studies on Retail Anomaly Trading

We provide two exogenous shocks to retail anomaly trading to shed light on the causal inference of retail trading on anomalies. The first pertains to the COVID-19-induced stay-at-home mandates, which precipitated an unparalleled surge in retail trading. This surge is largely attributable to the mass transition to working from home and a significant increase in media consumption during the pandemic period. The second pertains to the 1st economic impact payment. Following the methodology of Divakaruni and Zimmerman (2021), we implement a regression discontinuity design to capture the effects of this sudden wealth increase on retail anomaly trading.

Throughout this section, we classify anomalies into treatment and control groups based on pre-pandemic retail $\mathrm{ATI}_{\mathrm{t-2tot}}$. The treatment group, with $\mathrm{ATI}_{\mathrm{t-2tot}}$ t-statistics (Table 2 Panel A) below the 40th percentile, contains anomalies facing consistent negative retail trading from January 2015 to February 2020. The control group consists of anomalies with $\mathrm{ATI}_{\mathrm{t-2tot}}$ t-statistics in the 50th to 80th percentiles, which reflects a neutral retail stance—typically trading against these anomalies 40% to 50% of the time. Anomalies in the 40th to 50th percentile are withheld for robustness tests. The study considers 104 anomalies in the treatment group and 78 in the control group.

6.1 The Pandemic Induced Stay-at-Home Shock

6.1.1 Diff-in-Diff Design

For the analysis of the first exogenous shock to retail investor trading, induced by the pandemic stay-at-home orders, we narrowed our sample period to January 2015 through December 2020. We define 'during the pandemic' as the period from March 2020 to December 2020, and 'before the pandemic' as from January 2015 to February 2020. Figure 5 shows that the retail monthly volume dramatically increases during the pandemic. This finding is quantified by the regression results shown in Table 7, Panel A. Following the onset of stay-at-home orders, there is a significant upsurge in retail trading volume. It effectively doubled from a pre-pandemic average of 2.13% to an average of 4.62% during the pandemic period.

6.1.2 Empirical Results

Table 7, Panel B illustrates the differences in Anomaly Trading Intensity (ATI) between treatment and control group anomalies, both a quarter before and after anomaly formation. Column (1) details the regression outcomes for retail ATI prior to anomaly formation. The findings indicate that, on average, the retail ATI in the treatment group is 0.12% lower than in the control group before the pandemic. This disparity is accentuated during the pandemic, as evidenced by the significantly negative coefficient of -0.37% for the interaction term *Treat×Post*. This signifies that the difference in ATI between the treatment and control groups quadruples in comparison to the pre-pandemic period[14]. Columns (3) and (4) reveal that this heightened negative difference in ATI between treatment and control groups persists beyond the anomaly ranking month during the pandemic posing a credible threat to arbitrageurs countering retail investors.

Previously, we documented that pre-pandemic, for the anomalies traded against by retail investors, the risk-adjusted return of the anomaly portfolio a quarter before formation could be as low as nearly -4%, with a quick reversal occurring after ranking. During the pandemic, retail investors traded even more strongly against anomalies in the treatment group and for extended periods. Consequently, we hypothesized that

[14] Table IA3 of the Internet Appendix further indicates that the change in the ATI difference between treatment and control during the pandemic is mainly attributed to the ATI change in the short leg.

during the pandemic, the negative cross-sectional impact of retail trading on anomaly portfolio returns could extend beyond the anomaly ranking. Table 8 confirms this hypothesis. Column (1) reveals that a quarter before formation and before the pandemic, the treatment group anomalies' portfolio returns were 5.27% lower than those of the control group anomalies. During the pandemic, this difference more than tripled, with treatment group anomaly portfolio returns being 17.56% lower than those of the control group. Column (3) examines the formation month alpha. Before the pandemic, treatment group anomalies outperformed the control group, with the "Treat variable coefficient for the value-weighted anomaly returns being significantly positive at 0.63%. However, during the pandemic, the alpha for the treatment group anomalies, on average, was 2.16% lower than that of the control group anomalies. [15] [16] This is consistent with more negative Anomaly Trading Intensity (ATI) following the ranking month, suggesting that the intensified negative retail trading indeed caused arbitrageurs to suffer greater losses.

Figure 6 presents detailed plots of the return differences between treatment and control group anomalies, both before and during the pandemic, extending from a quarter before anomaly formation up to 24 months afterward. For this analysis, we employed a regression model of $\alpha = \beta_1 Treat + \beta_2 Treat \times Post$ with double clustered standard errors. The light green bars show the estimate of β_1 together with a 5% confidence interval, i.e., the difference in anomaly portfolio performance between the treatment and control groups before the pandemic. The dark green shaded bars show the estimate of $\beta_1 + \beta_2$ and 5% confidence interval, i.e., the performance difference between the treatment and the control group anomalies during the pandemic. Before the pandemic, anomaly portfolio returns of treatment anomalies are significantly lower than those of the control group anomalies. Such difference disappears right after the anomaly ranking month. However, during the pandemic and before anomaly formation, the anomaly portfolio returns for the treatment group anomalies are even more negative compared to those of the control

[15] Table IA4, Panel A of the Internet Appendix shows 72% of the during pandemic formation-month alpha difference between treatment and control is due to increased alpha in the short leg.

[16] Table IA4, Panel C of the Internet Appendix further reveals that as the ATI of the treatment group becomes more negative (from the bottom 50th to bottom 30th percentile), the formation-month alpha difference between treatment and control group anomalies during the pandemic becomes more negative.

group, and this trend continues even after the anomaly signal formation. The difference in anomaly portfolio returns remains significantly negative up to one quarter after anomaly formation, followed by a more pronounced reversal from the 7th to the 24th month. This outcome is consistent with the retail anomaly trading intensity difference remaining negative for a quarter after the anomaly signal formation, posing a credible risk to arbitrageurs.

Finally, using the panel regression of $\alpha_{i,t+1} = \sum_{t=2015}^{2020} \beta_t Treat_i \times Year_t + \mu_i + \gamma_t + \epsilon_{it}$, Figure 7 shows parallel trend interaction term coefficients β_t for the formation-month alpha. $Year_t$ is a pseudo-post-event dummy, assigned a value of 1 for the sample range from March in year t to March in the subsequent year, t+1. Both firm and date-fixed effects are incorporated in the analysis. The year 2015 serves as the reference period and its interaction term is omitted to avoid multicollinearity. The parallel trend test suggests that using the sample period of 6 years, for α_{t+1} , it is only in 2020 that the difference between treatment and control group anomaly portfolio returns is significantly lower than that in 2015. This finding highlights a distinct shift in the relationship between treatment and control group anomalies specifically during the surge in retail trading.

6.2 The First Stimulus Check

6.2.1 Regression Discontinuity Design

The first Economic Impact Payment (EIP) in the U.S., part of the response to the COVID-19 pandemic, was announced on March 30, 2020. It provided up to $1,200 for individuals. The payments, which began distribution shortly after the announcement, were aimed at offering financial relief to households and stimulating economic activity. Starting from April 2020, the US government sent these payments directly to households. Figure 8 from Divakaruni and Zimmerman (2021) shows the accumulative distribution in terms of the number of paychecks and the total amount paid. The first paycheck starts to distribute on April

9 and the distribution continues beyond June 8[17]. Thus, we use daily retail anomaly trading intensity from Jan 2, 2020, to Jun 12, 2020, and define the post-disbursement period from April 9 to June 12.

The key idea of regression discontinuity design is that we examine retail ATI within a \$100 window centered on the \$1,200 EIP—\$50 above and \$50 below this threshold. With the wealth constraint relaxed post-disbursement, a change in retail trading activity should only be observed in the lower cost range. To do so, we form daily retail ATI at the pre-formation month (ranking month) by dollar amount traded. Specifically, we calculate two stock level ATI based on whether the dollar amount of a retail order falls into the range of [1150,1200] or (1200, 1250]. Then, we detrend the stock level trading intensity using 2019's average daily trading intensity in each dollar range. Consistent with monthly retail stock level trading methodology, daily trading is winsorized at 0.5% and 99.5% levels by different dollar ranges. The resultant daily anomaly trading intensity was derived from the average stock level trading intensity in the long leg subtracted by that of the short leg. A negative retail ATI indicates that retail investors traded in a manner that exacerbated anomaly mispricing- buying more stocks in the short leg.

Figure 9 shows a descriptive analysis of daily trading volume and the number of trades by dollar amount trade, using the pre-EIP sample from Jan 2 to Apr 9, 2020. As the dollar costs increase, the number of trades and volume placed by retail investors decrease. Within the [1150,1200) and (1200, 1250] cost brackets, an average of approximately 15 trades per stock per day was made by retail investors.

We first run diff-and-diff regressions for treatment and control group anomalies separately:

$$ATI_{it} = \beta_1 \mathrm{I}(< 1200)_i + \beta_2 \mathrm{I}(< 1200)_\mathrm{i} \times \mathrm{Post}_\mathrm{t} + \gamma_t + \epsilon_{it}$$

Here, ATI_{it} is the anomaly retail trading intensity at one month prior to formation month, $\mathrm{I_i}(< 1200)$ is an indicator and takes on the value of 1 if the retail order falls within the lower cost range [1150,1200), and 0 otherwise, and γ_t is the time-fixed effect. We hypothesize that before disbursement, we should observe no difference in ATI in the two dollar cost ranges both for treatment and control group anomalies,

[17] CNBC reports on June 8, 2020: 35 million checks are yet to be disbursed *https://www.cnbc.com/2020/06/08/35-million-stimulus-checks-havent-been-sent-out-who-is-waiting-for-money.html*.

i.e., β_1 is insignificant from 0. Post-disbursement, retail investors place more buy orders in the range [1150, 1200) for short-leg stocks in the treatment group anomalies, resulting in a more negative ATI_{it} , thus, a significant negative β_2. Control group anomalies, where retail investors are expected to maintain neutrality, should exhibit an insignificant β_2. Moreover, since the EIP acts as an exogenous shock to individuals' wealth, we expect the observed differences to arise predominantly from changes in the buy-side anomaly trading intensity. Therefore, we decompose the retail ATI into buy-side and sell-side components: the buy-side ATI is the difference in the average number of shares bought in the long and short legs of an anomaly, while the sell-side ATI is the corresponding difference in shares sold.

To facilitate the incorporation of additional variables, such as an announcement and crash dummies, we have developed a triple interaction regression model. The model is specified as:

$$ATI_{it} = \beta_1 Treat_i + \beta_2 I\ (< 1200)_i + \beta_3 Treat_i \times Post_t$$
$$+\boldsymbol{\beta_4} Treat_i \times I\ (< 1200)_i + \beta_5 I\ (< 1200)_i \times Post_t$$
$$+\boldsymbol{\beta_6} Treat_i \times Post_t \times I\ (< 1200)_i + \gamma_t + \epsilon_{it}$$

In line with the difference-in-differences design, we hypothesize that, prior to the disbursement, there is a distinguishable difference in trading intensities between treatment and control group anomalies within each dollar cost range. Specifically, $\beta_1 + \beta_4$ should be significantly negative for the lower cost range, while β_1 alone should be significantly negative for the higher cost range. However, the difference between the treatment and control groups should remain consistent across both dollar cost ranges, implying that $\boldsymbol{\beta_4}$ should not be significantly different from zero. After the disbursement, we anticipate retail investors are more likely to increase trading in the dollar range of [1150,1200), and the difference between treatment and control group trading is more pronounced in the lower cost than the higher cost. As a result, $\boldsymbol{\beta_6}$ should be significantly negative.

6.2.2 Empirical Results

Table 9, Panel A represents results for diff-and-diff design. Columns (1) and (2) show treatment group anomalies. In Column (1), focusing on the period before the paycheck disbursement, we observe no significant difference in retail trading between the low-cost and high-cost dollar ranges for the buy-side ATI. After the disbursement, retail investors buy more of the short-leg stock within the lower cost range, leading to a negative difference between the buy-side ATI in the lower cost range and the high cost range. Column (2), in contrast, indicates that there are no significant findings from the sell-side ATI, confirming our hypothesis that the sell-side trading behavior should not exhibit noticeable changes in response to the disbursement. Furthermore, columns (3) and (4) reveal that there is no change in the retail ATI in control group anomalies. This observation further confirms our hypothesis that control group anomalies are those where retail investors take a neutral stance.

After identifying the retail anomaly trading intensity change in the treatment group anomalies, we further decompose retail trading within treatment group anomalies into short-leg and long-leg trading. Column (2) of Table IA5 of the Internet Appendix suggests that, after the disbursement, retail investors tend to buy stocks in the short leg of treatment group anomalies, and it’s the primary driver behind the observed changes in the treatment group's anomaly trading intensity.

Table 9, Panel B represents the results of the triple difference analysis. Column (1) reveals that retail investors generally exhibit more negative trading towards treatment group anomalies compared to control group anomalies, with an average difference of approximately 0.057 basis points (bp) per day within each dollar cost range. Furthermore, the trading difference between treatment and control group anomalies across the [1150,1200) and (1200, 1250] ranges is initially insignificant at 0.002bp before the disbursement, but becomes more negative, reaching -0.010bp, after the disbursement. Consistent with the findings in Panel A, there are no significant results from the sell-side ATI.

To account for potential influences beyond the exogenous shock to household wealth, we introduce an announcement dummy into the regression. This dummy, denoted as 'Ann', takes the value of 1 for ATI during the first stimulus check announcement week, specifically from -9 to -6 days relative to the event day. Columns (2) and (4) demonstrate that our results remain robust after including the announcement dummy. Additionally, we consider the possibility that the results could be influenced by retail investors providing liquidity during the market crash. To address this, we include a 'crash' dummy in our analysis. The S&P 500 index experienced a significant drop of 29% from March 2 to March 16, 2020. We define 'Crash=1' for trading intensities falling within the event days of -26 to -13. Columns (3) and (6) confirm the robustness of our results.

6.3 The Tech Bubble: An Out-Of-Sample Test of Anomaly Performance

The latter half of the 1990s witnessed a remarkable escalation in technology stock values, leading to the phenomenon known as the 'tech bubble.' This era was characterized by robust investment activities by both institutional and retail investors, as noted by Griffin, Harris, Shu, and Topaloglu (2011). Our analysis focuses on the treatment and control groups' anomaly performance during the formation month within the tech bubble, using a sample timeframe from January 1993 to January 2000. The commencement of the tech bubble, as identified by Baker and Wurgler's (2006) initial public offering (IPO) returns, extends from September 1998 to January 2000. Figure 10 elucidates the comparative performance trajectories of the treatment and control group anomalies during two significant market upheavals: the tech bubble and the COVID-19 pandemic. In both instances, with an immediate correction occurring at the formation month, the anomalies within the treatment group initially outperform those in the control group. Amidst the crises, the treatment group's anomalies tend to underperform, attributed to retail investors' increased purchase of short-leg stocks within the treatment group anomalies. This activity challenges arbitrageurs' capacity to correct mispricing, resulting in the persistence of mispricing beyond the anomaly ranking month, followed by a subsequent significant larger reversal.

The empirical robustness of our findings is further solidified when viewed through the lens of the technology bubble of the late 1990s—an out-of-sample event when considering our initial selection of treatment and control group anomalies based on retail trading data from the 2010s. Despite the temporal and contextual disparities, Figure 10 demonstrates a strikingly similar anomaly performance pattern during the tech bubble, mirroring the trends observed amidst the COVID-19 pandemic. This congruence suggests that the retail trading behaviors influencing anomaly persistence are not confined to recent market conditions but also extend to earlier periods of significant market exuberance.

7. Conclusion

Our examination of 260 anomalies challenges the prevailing notion that market efficiency erodes anomaly-based profits, these anomalies continue to thrive, especially over longer timeframes. We demonstrate that retail investors play a pivotal role in the persistence of these anomalies. Their stubborn trading patterns, especially against anomalies, not only contribute to initial mispricing but also lead to delayed price corrections. This behavior imposes long-horizon risks on arbitrageurs, complicating the market's return to equilibrium and underscoring the complexity of market dynamics in the presence of retail trading. Our study makes significant contributions to the understanding of long-horizon returns through the two novel retail-related factor models. Additionally, our analysis of the pandemic-induced stay-at-home orders, EIP, and the tech bubble era reveals the intensified impact of retail trading on anomaly performance, exacerbating mispricing for an extended periods and presenting a substantial challenge to arbitrage strategies.

References

Akbas, F., Armstrong, W.J., Sorescu, S., & Subrahmanyam, A., 2015, Smart money, dumb money, and capital market anomalies. Journal of Financial Economics 118, 355-382.

Baba-Yara, F., Boons, M. and Tamoni, A., 2020. New and old sorts: Implications for asset pricing.

Baker, M. and Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. The journal of Finance, 61, 1645-1680.

Bali, T.G., Hirshleifer, D., Peng, L. and Tang, Y., 2021. Attention, social interaction, and investor attraction to lottery stocks (No. w29543). National Bureau of Economic Research.

Barber, B.M., Lin, S. and Odean, T., 2021. Resolving a paradox: Retail trades positively predict returns but are not profitable. Journal of Financial and Quantitative Analysis, 1-67.

Barber, B.M., Huang, X., Odean, T. and Schwarz, C., 2022. Attention-Induced Trading and Returns: Evidence from Robinhood Users. The Journal of Finance, 77, 3141-3190.

Barber, B.M., Odean, T. and Zhu, N., 2008. Do retail trades move markets?. The Review of Financial Studies, 22, 151-186.

Barberis, N. and Huang, M., 2008. Stocks as lotteries: The implications of probability weighting for security prices. American Economic Review, 98, 2066-2100.

Ben-David, I., Li, J., Rossi, A. and Song, Y., 2021. Discontinued Positive Feedback Trading and the Decline of Return Predictability (No. w28624). National Bureau of Economic Research.

Boehmer, E., Jones, C.M., Zhang, X. and Zhang, X., 2021. Tracking retail investor activity. The Journal of Finance, 76, 2249-2305.

Bouchaud, J.P., Krueger, P., Landier, A. and Thesmar, D., 2019. Sticky expectations and the profitability anomaly. The Journal of Finance, 74, 639-674.

Bowles, B., Reed, A.V., Ringgenberg, M.C. and Thornock, J.R., 2020. Anomaly time. Available at SSRN 3069026.

Birru, J., 2018. Day of the week and the cross-section of returns. Journal of financial economics, 130, 182-214.

Bryzgalova, S., Pavlova, A. and Sikorskaya, T., 2023. Retail trading in options and the rise of the big three wholesalers. The Journal of Finance, 78, 3465-3514.

Calluzzo, P., Moneta, F. and Topaloglu, S., 2019. When anomalies are publicized broadly, do institutions trade accordingly? Management Science, 65, 4555-4574.

Carhart, M.M., 1997. On persistence in mutual fund performance. The Journal of Finance, 52, 57-82.

Chernov, M., Lochstoer, L.A. and Lundeby, S.R., 2022. Conditional dynamics and the multihorizon risk-return trade-off. The Review of Financial Studies, 35, 1310-1347.

Cho, T. and Polk, C., 2019. Putting the price in asset pricing. Available at SSRN 3499681.

Chordia, T., Subrahmanyam, A. and Tong, Q., 2014. Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? Journal of Accounting and Economics, 58, 41-58.

Cohen, R.B., Polk, C. and Vuolteenaho, T., 2009. The price is (almost) right. The Journal of Finance, 64, 2739-2782.

Daniel, K., Hirshleifer, D., & Sun, L., 2020. Short- and long-horizon behavioral factors. Review of Financial Studies 33, 1673-1736.

De Long, J.B., Shleifer, A., Summers, L.H. and Waldmann, R.J., 1990. Noise trader risk in financial markets. Journal of Political Economy, 98, 703-738.

Divakaruni, A. and Zimmerman, P., 2021. Uncovering retail trading in bitcoin: The impact of COVID-19 stimulus checks (No. 21-13).

Dong, X., Feng, S., & Sadka, R., 2019, Liquidity risk and mutual fund performance. Management Science 65, 1020-1041.

Dong, X., Li, Y., Rapach, D.E. and Zhou, G., 2022. Anomalies and the expected market return. The Journal of Finance, 77, 639-681.

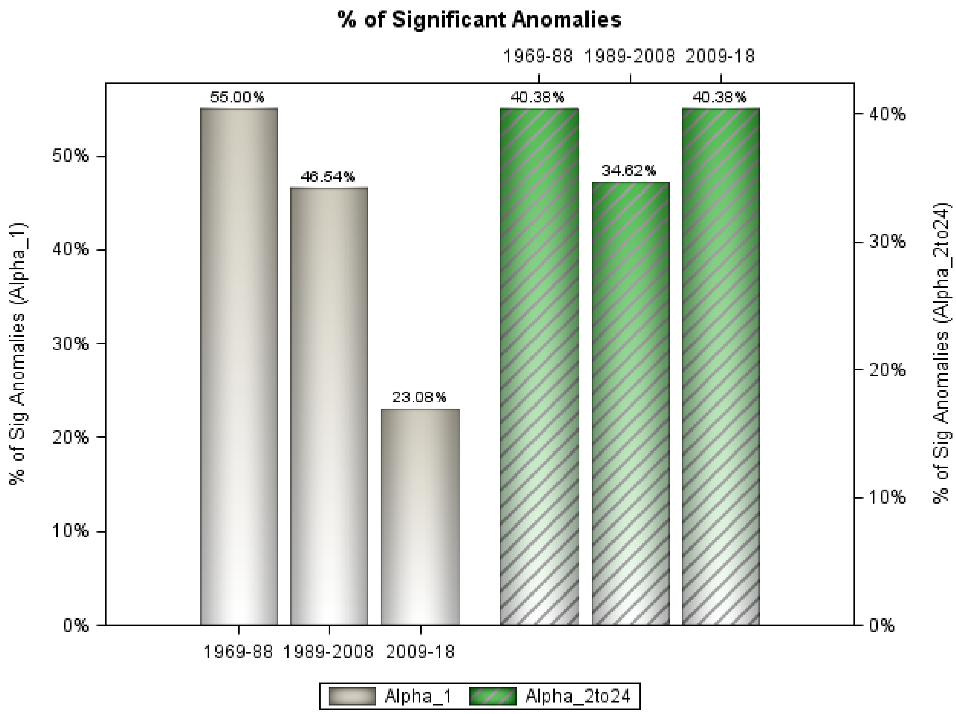
Dong, X., Kang, N. and Peress, J., 2020. Fast and Slow Arbitrage: Fund Flows and Mispricing in the Frequency Domain.
Edelen, R.M., Ince, O.S. and Kadlec, G.B., 2016. Institutional investors and stock return anomalies. Journal of Financial Economics, 119, 472-488.
Engelberg, J.E., Reed, A.V. and Ringgenberg, M.C., 2018. Short-selling risk. The Journal of Finance, 73, 755-786.
Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33, 3-56.
Fama, E.F. and French, K.R., 2015. A five-factor asset pricing model. Journal of financial economics, 116, 1-22.
Fama, E.F. and French, K.R., 2021. The value premium. The Review of Asset Pricing Studies, 11, 105-121.
Gormsen, N.J. and Lazarus, E., 2023. Duration-Driven Returns. The Journal of Finance, 78, 1393-1447.
Green, J., Hand, J.R.M, & Zhang, X.F., 2017. The characteristics that provide independent information about average US monthly stock returns. Review of Financial Studies 30, 4389-4436.
Greenwood, R., Laarits, T. and Wurgler, J., 2022. Stock market stimulus. National Bureau of Economic Research.
Griffin, J.M., Harris, J.H., Shu, T. and Topaloglu, S., 2011. Who drove and burst the tech bubble?. The Journal of Finance, 66, 1251-1290.
Haddad, V., Kozak, S. and Santosh, S., 2020. Factor timing. Review of Financial Studies 33, 1980-2018.
Han, B. and Kumar, A., 2013. Speculative retail trading and asset prices. Journal of Financial and Quantitative Analysis, 48, 377-404.
Harvey, C.R., Liu, Y. and Zhu, H., 2016. … and the cross-section of expected returns. The Review of Financial Studies, 29, 5-68.
Hou, K., Xue, C. and Zhang, L., 2017. Replicating anomalies. The Review of Financial Studies
Jensen, T.I., Kelly, B.T. and Pedersen, L.H., 2022. Is there a replication crisis in finance? Journal of Finance.
Kelley, E.K. and Tetlock, P.C., 2013. How wise are crowds? Insights from retail orders and stock returns. The Journal of Finance, 68, 1229-1265.
Kaniel, R., Saar, G. and Titman, S., 2008. Individual investor trading and stock returns. The Journal of Finance, 63, pp.273-310.
Kaniel, R., Liu, S., Saar, G. and Titman, S., 2012. Individual investor trading and return patterns around earnings announcements. The Journal of Finance, 67, 639-680.
Kim, Y.H., Ivkovich, Z. and Muravyev, D., 2021. Causal Effect of Information Costs on Asset Pricing Anomalies. Available at SSRN 3921785.
Kumar, A., 2009. Who gambles in the stock market? The Journal of Finance, 64, 1889-1933.
Kumar, A. and Lee, C.M., 2006. Retail investor sentiment and return comovements. The Journal of Finance, 61, 2451-2486.
McLean, R.D., & Pontiff, J., 2016, Does academic research destroy stock return predictability? Journal of Finance 71, 5-32.
McLean, R.D., Pontiff, J. and Reilly, C., 2020. Taking sides on return predictability. Available at SSRN 3637649.
Muravyev, D., Pearson, N.D. and Pollet, J.M., 2022. Anomalies and their short-sale costs. Available at SSRN 4266059.
Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics, 108, 1-28.
Ozik, G., Sadka, R. and Shen, S., 2021. Flattening the illiquidity curve: Retail trading during the COVID-19 lockdown. Journal of Financial and Quantitative Analysis, 56, 2356-2388.
Pástor, Ľ. and Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. Journal of Political Economy, 111, pp.642-685.

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics, 42, 35-52.
Shleifer, A. and Vishny, R.W., 1997. The limits of arbitrage. The Journal of Finance, 52, 35-55.
Stambaugh, R.F., Yu, J., & Yuan, Y., 2012, The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288-302.
Stambaugh, R.F., Yu, J., & Yuan, Y., 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle. Journal of Finance 70, 1903-1948.
Stambaugh, R.F., & Yuan, Y., 2017, Mispricing factors. Review of Financial Studies 30, 1270-1315.
Smith, S.C. and Timmermann, A., 2022. Have risk premia vanished? Journal of Financial Economics, 145, 553-576.
Teo, M., 2011, The liquidity risk of liquid hedge funds. Journal of Financial Economics 100, 24-44.
Titman, S., Wei, K.C.J., & Xie, F., 2004, Capital investment and stock returns. Journal of Financial & Quantitative Analysis 39, 677-700.
Van Binsbergen, J.H., Boons, M., Opp, C.C. and Tamoni, A., 2023. Dynamic asset (mis) pricing: Build-up versus resolution anomalies. Journal of Financial Economics, 147, 406-431.
Van Binsbergen, J.H. and Opp, C.C., 2019. Real anomalies. The Journal of finance, 74, 1659-1706.
Welch, I., 2022. The wisdom of the Robinhood crowd. The Journal of Finance, 77, 148.

Figure 1 Short and Long-term Anomaly Portfolios

Figure 1 panel A and panel B show the percentage of significant anomaly portfolios and return magnitude using CAPM-alpha over three distinct intervals spanning the past half-century. Alphas are calculated using in-sample method. For example, the formation-month α_{t+1} (Alpha_1) from 1969 to 1988, is calculated using the sample period from Jan 1969 to Dec 1988, defined as anomaly ranking months, by regressing anomaly portfolio return on market premium, and excluding market risk premium. The long-term $\alpha_{t+2tot+24}$ is the cumulative CAPM-adjusted return over the event periods from 2 to 24 months after anomaly formation. T-stats are adjusted for serial correlation using Newey-West method. $\alpha_{t+2tot+24}$ is adjusted for 23 lags. An anomaly portfolio return is significant when t-stats are higher than 1.96.

Panel A: Short and Long-term Percentage of Significant Anomalies



Panel B: Short and Long-term Anomaly Performance

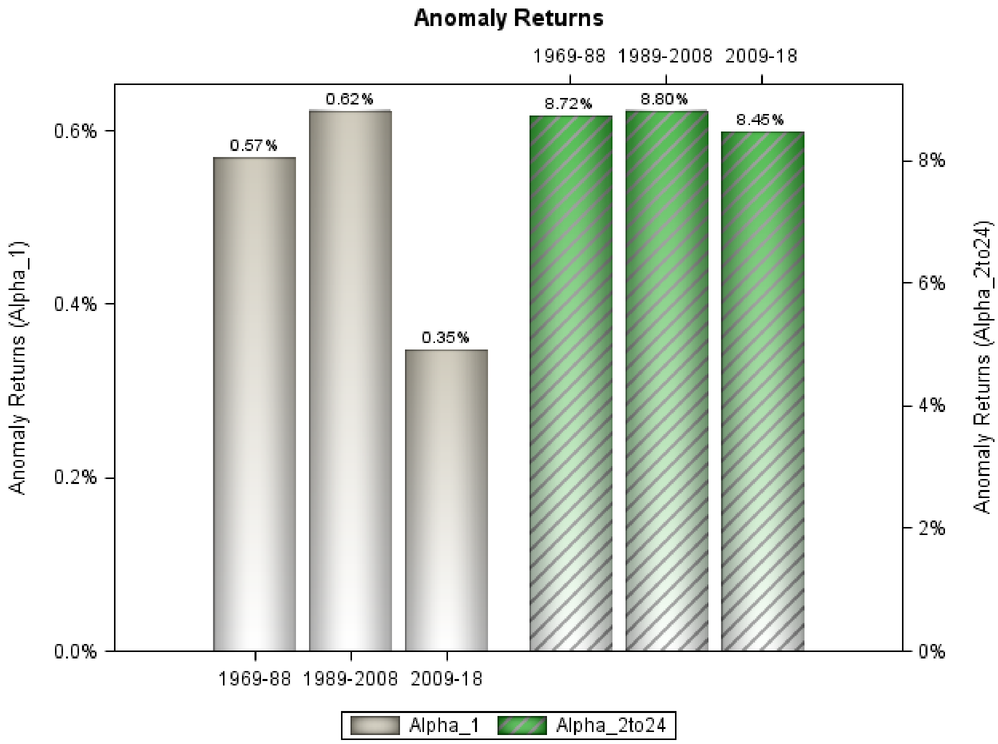
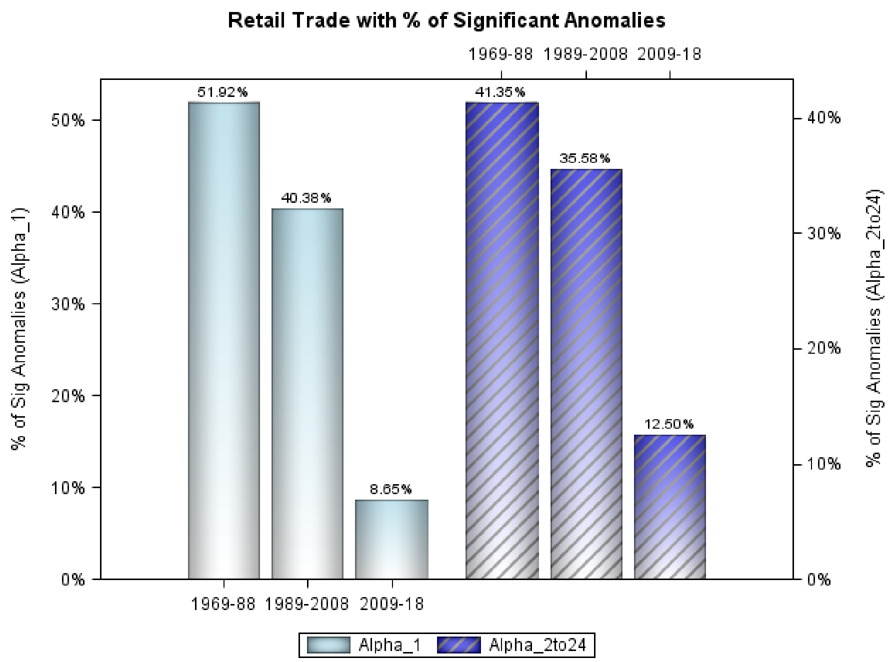


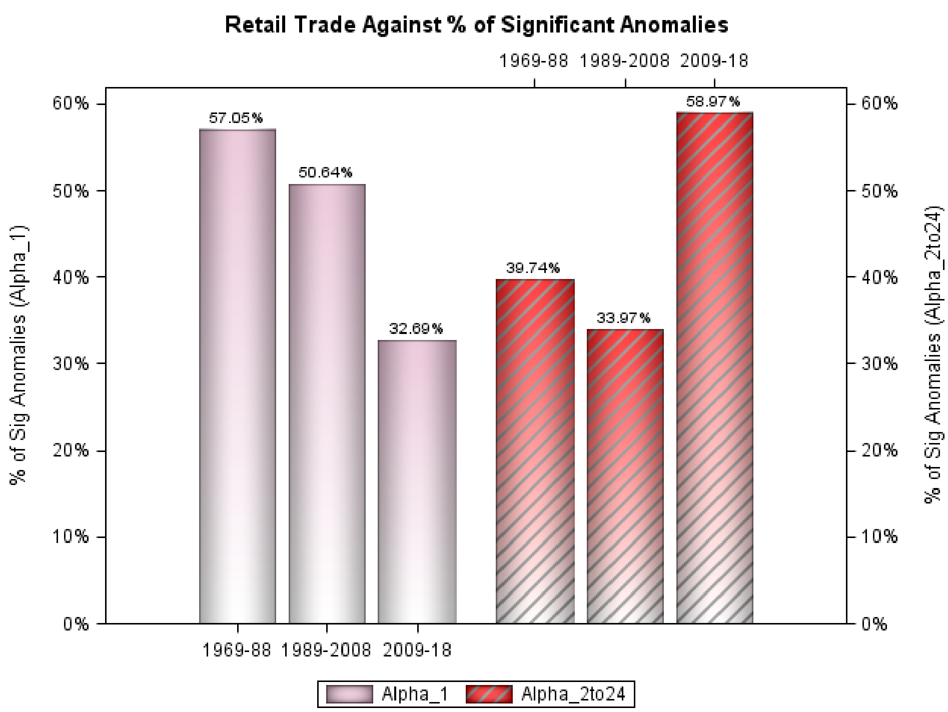
Figure 2 Short and Long-term Anomaly Portfolio by Retail Trading Direction

Figure 2 further divides anomalies into retail investors trade with versus trade against and examines short and long-term anomaly portfolio alphas. If retail investors buy more (less) stocks in the short leg than in the long leg of an anomaly, the anomaly is classified into the group where retail investors trade against (with). Panel A1 and A2 show the percentage of significant anomaly portfolios. Panel B1 and B2 show the alpha magnitude.

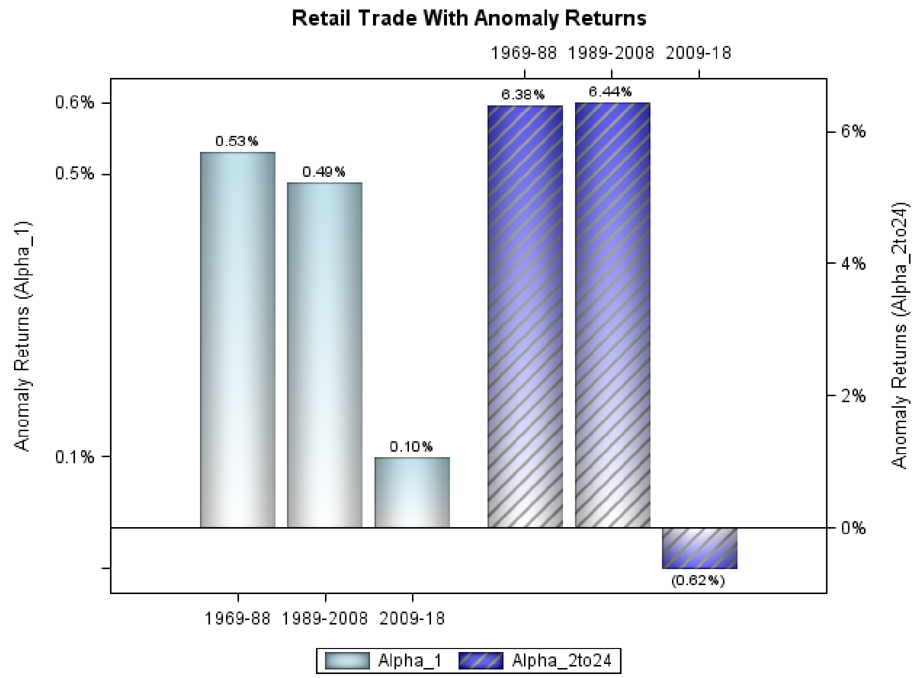
Panel A1: Retail Trade with Percentage of Significant Anomalies



Panel A2: Retail Trade Against Percentage of Significant Anomalies



Panel B1: Retail Trade with Anomaly Alphas



Panel B2: Retail Trade Against Anomaly Alphas

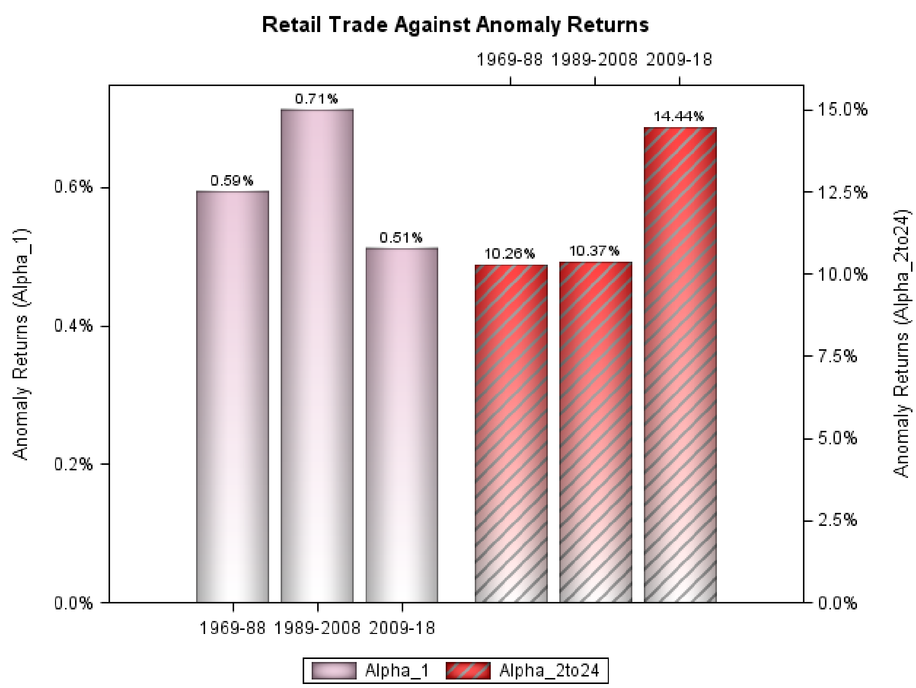
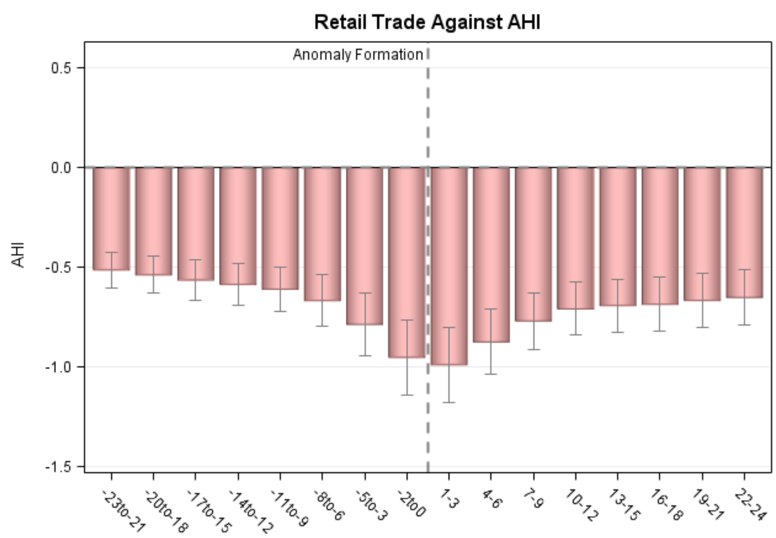


Figure 3 Retail and Short-Seller Anomaly Holding Intensity

Figure 3 shows retail and short seller anomaly holding intensity (AHI) in percent for the anomalies that retail investor trade against. Figure 3 plots AHI 24 event-month before anomaly formation and 24-month after. Sample period is from 2009 to 2020, defined as anomaly ranking months. Panel A shows retail AHI. Panel B shows short seller AHI.

Panel A: Retail AHI



Panel B: Short-Seller AHI

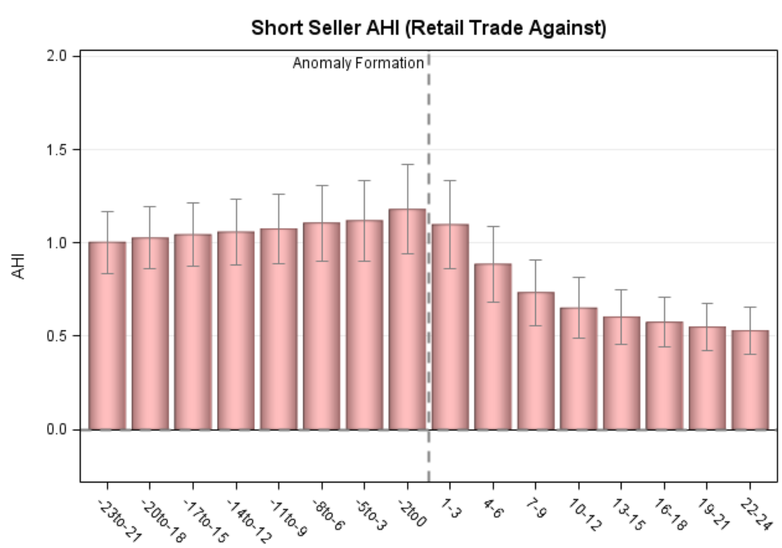


Figure 4 Monthly Average Anomaly Alpha T-stats

Figure 4 shows average anomaly portfolio alpha t-statistics within each R1 to R5 group using the sample period from 2009 to 2020, defined as anomaly ranking months. Retail investors trade with R1 and R2 group of anomalies, and trade against R3 to R5 group of anomalies.

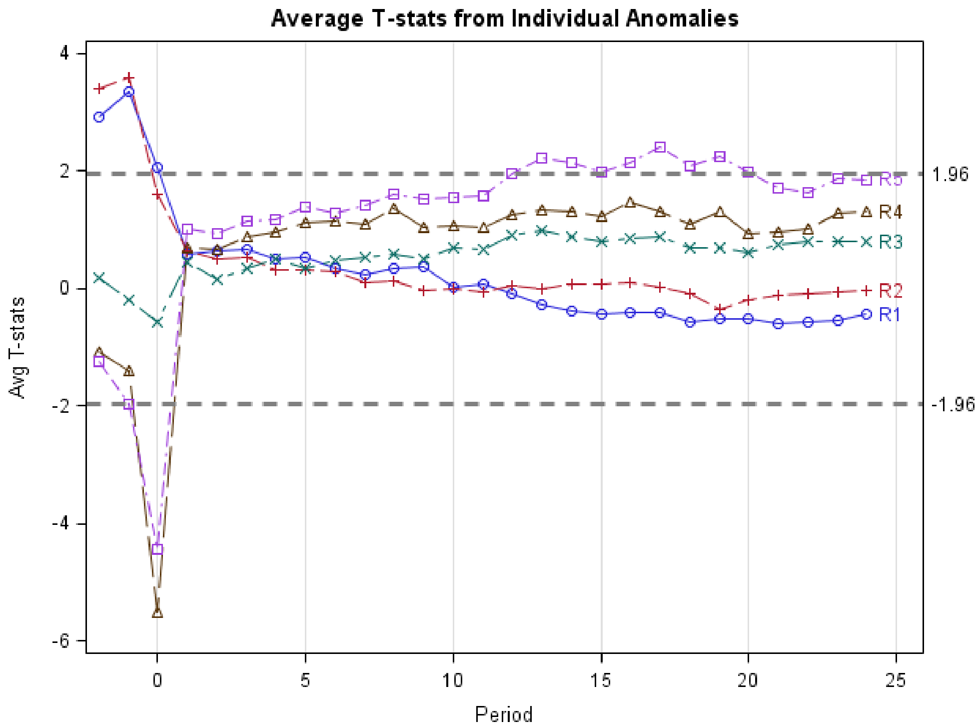


Figure 5 Retail Trading Volume

Figure 5 plots the retail time series monthly average stock level trading volume (in percent). Monthly average retail trading is calculated as the average retail trading of both shares bought and sold scaled by shares outstanding across all companies that are used in anomaly portfolio construction.

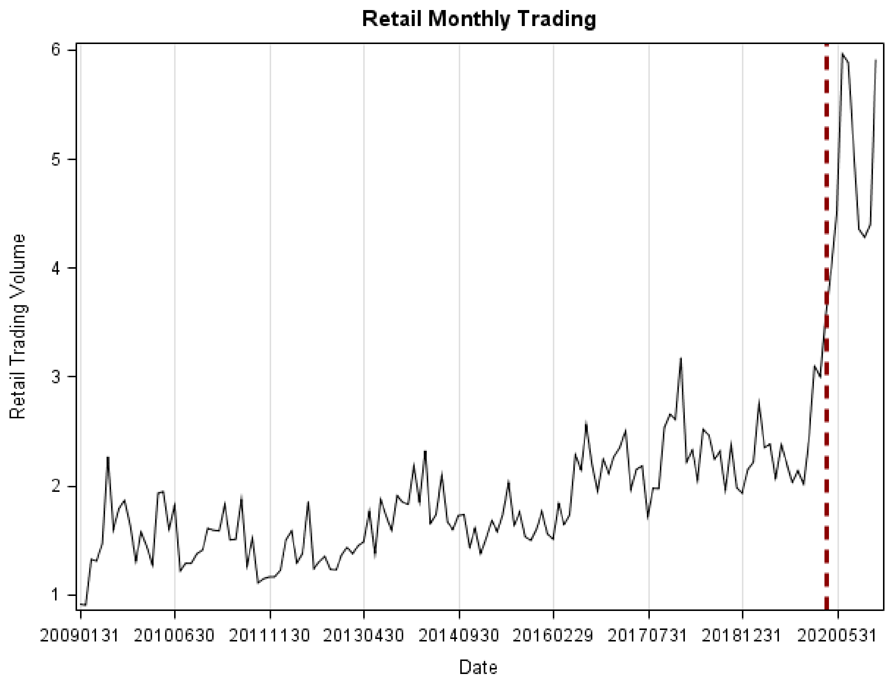


Figure 6 Anomaly Performance Plot

Figure 6 plots the anomaly portfolio alpha difference between treatment and control group anomalies five years before and during the pandemic, extending from one quarter before anomaly formation to several quarters after. The light green bars show the difference before the pandemic and the dark green shaded bars show the difference during the pandemic. The lines indicate 95% confidence interval. Sample period is from 2015 to 2020, defined as anomaly ranking months. Treatment and control group anomalies are defined using pre-pandemic t-stats of ATI a quarter before anomaly formation. The treatment anomalies are those in the lower 40th percentile (104 anomalies). Control group anomalies are those in the 50^{th} to 80^{th} percentile and there are 78 anomalies in the control group. Standard errors are double clustered by date and anomaly.

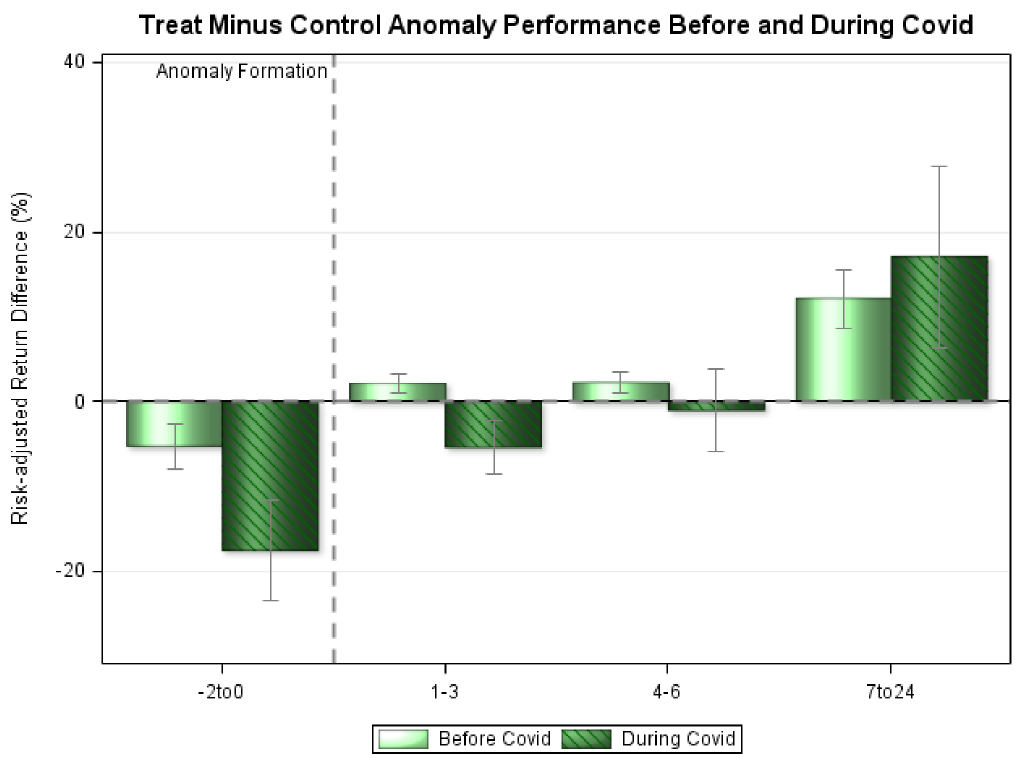


Figure 7 Parallel Trend

Figure 7 plots the estimated coefficient β_t for the interaction terms from the parallel trends analysis using the panel regression. The dependent variables is anomaly portfolio alphas at the formation month α_{t+1} . Independent variables include a series of interaction terms between the treatment indicator (*Treat*) and pseudo-post-event yearly dummies ($Year$) from 2015 to 2020. The year 2015 serves as the reference period and its interaction term is omitted to avoid multicollinearity. The plotted β_t coefficients indicate the change in the difference in α_{t+1} between the treatment and control groups over time compared to that of the base year 2015.

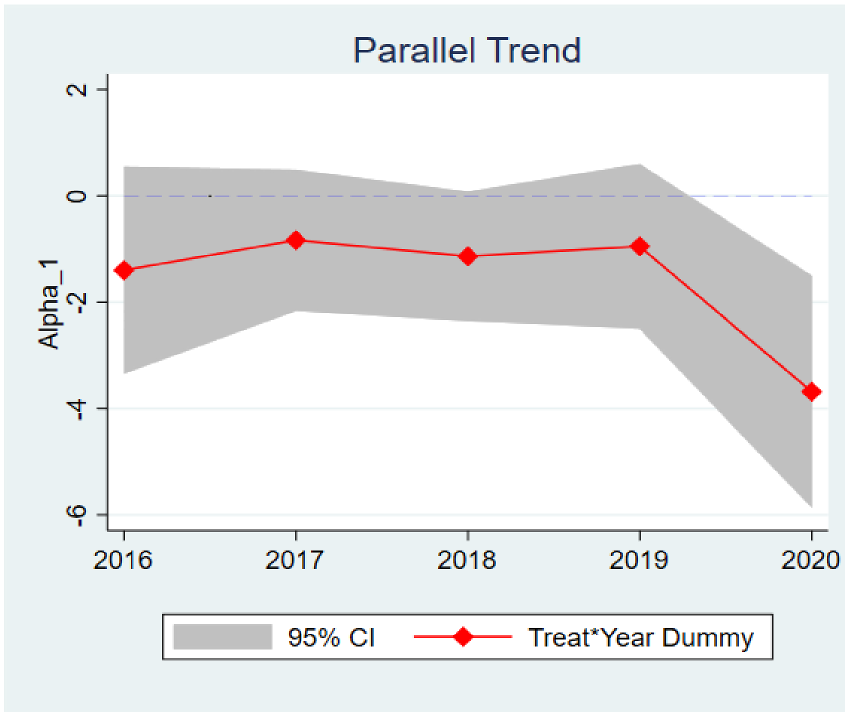


Figure 8 EIP Distribution Schedule

Figure 8 shows the Economic Impact Payment distribution Schedule (Divakaruni and Zimmerman (2021)). The dashed line shows the cumulative number of EIPs and the solid line shows the cumulative amount paid.

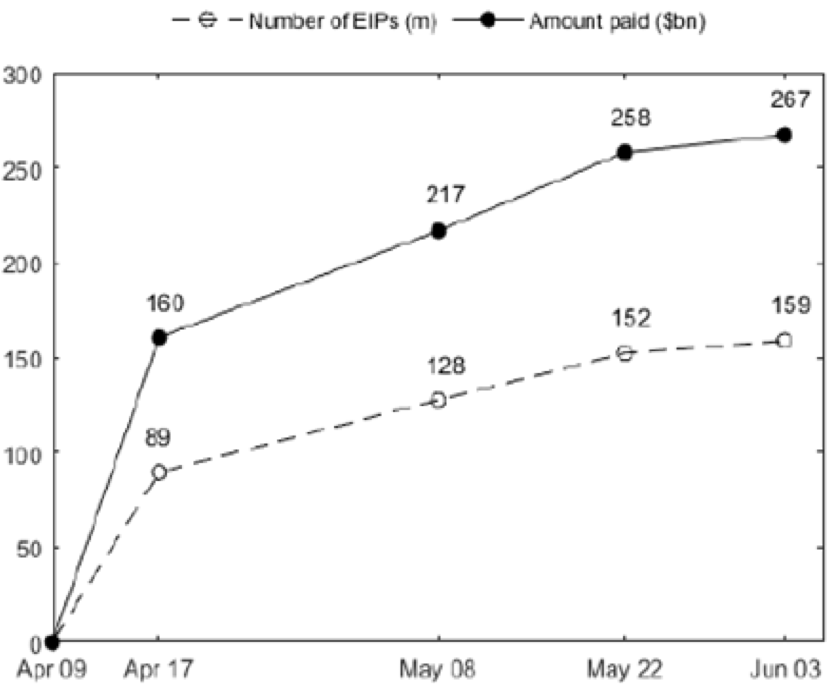


Figure 9 RDD Design

Figure 9 shows daily retail trading volume and the number of trades by dollar amount traded. Retail trading by dollar amount is identified using the algorithm proposed by Boehmer, Jones, Zhang, and Zhang (2021). We calculate retail trading volume as shares bought and sold scaled by shares outstanding across all companies that are used in anomaly portfolio construction.

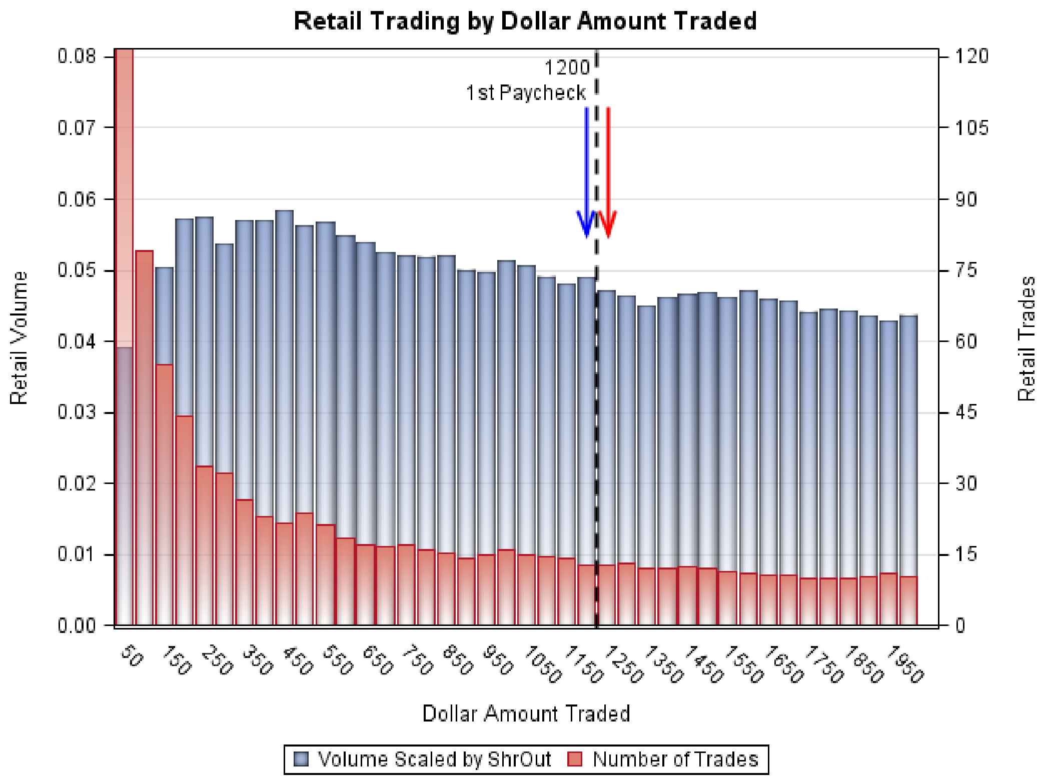


Figure 10 During the Tech Bubble and COVID-19 Anomaly Performance Comparison

Figure 10 contrasts the anomaly performance at formation months across two distinct market events characterized by heightened retail investor trading: the Tech Bubble and the COVID-19 pandemic. The red line represents the treatment group alpha, while the blue line depicts the control group alpha.

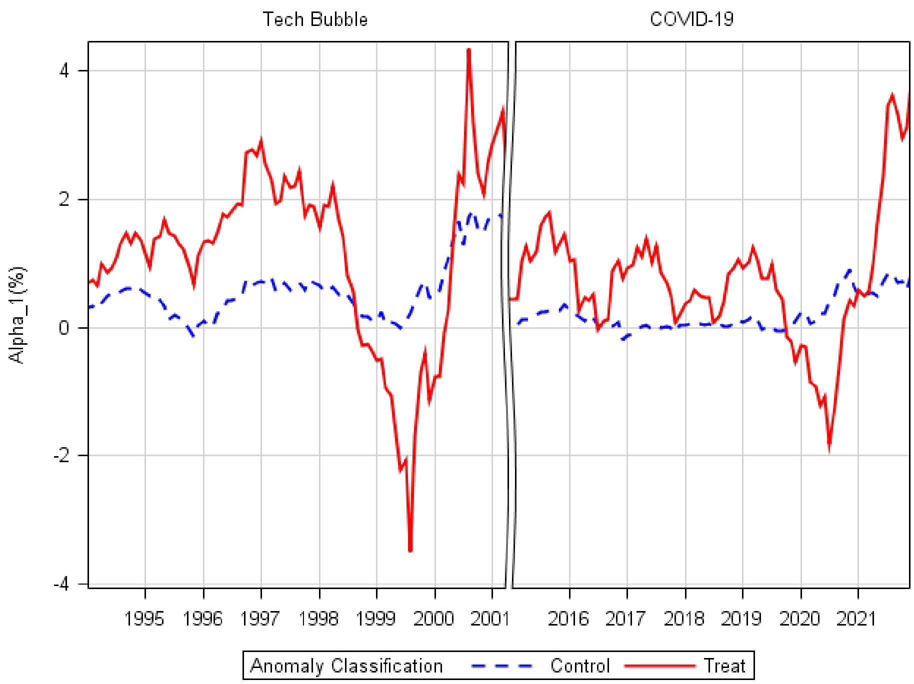


Table 1 Summary Statistics

Panel A shows summary statistics for the 260 anomaly portfolios CAPM-adjusted returns and raw returns. Anomaly portfolio alphas are calculated by regressing anomaly portfolio raw returns on market premium and excluding the anomaly's market risk premium $\alpha = Ret - \beta \times Mktrf$. The sample period is from January 2009 to December 2020, defined as anomaly ranking months. $\alpha_{t-2tot0}$ is the cumulative CAPM-adjusted anomaly return from period t-2 to t (one quarter before anomaly formation). α_{t+1} is the formation month anomaly portfolio alpha. $\alpha_{t+1tot+12}$ is the cumulative anomaly portfolio alphas one year after anomaly formation. *Panel B* shows *stock level* monthly trading and holding intensity summary statistics for retail investors and short sellers. We invert the sign of short interest directly from Compustat. A more negative trading intensity is interpreted as high short interest. *Panel C* shows summary statistics for anomaly trading and holding intensity for retail investors and short sellers one quarter before anomaly formation.

Panel A: Anomaly Portfolio Risk-adjusted Returns and Raw Returns

	N	Mean	Std.Dev.	Median	N	Mean	Std.Dev.	Median
				Value Weight (%)				
	Jan 2009 to Feb 2020				*March 2020 to Dec 2020*			
α_{t-2tot}	34,668	0.07	4.26	0.02	2,583	-2.75	23.67	-3.12
α_{t+1}	34,668	0.29	10.99	0.27	2,583	-0.40	6.89	-0.25
$\alpha_{t+1tot+6}$	34,662	1.69	15.99	1.26	2,583	-0.19	16.81	-0.17
$\alpha_{t+1tot+12}$	34,658	3.30	20.31	2.29	2,583	3.29	20.71	1.58
$\alpha_{t+1tot+24}$	34,658	6.74	25.91	3.90	2,583	16.05	41.10	9.88
				Value Weight (%)				
	Jan 2009 to Feb 2020				*March 2020 to Dec 2020*			
Ret_{t-2tot}	34,668	-0.84	15.17	-0.62	2,583	-4.21	25.91	-3.96
Ret_{t+1}	34,668	0.05	4.65	0.07	2,583	-1.41	7.52	-0.81
$Ret_{t+1tot+6}$	34,662	0.07	11.42	-0.24	2,583	-4.25	19.11	-3.47
$Ret_{t+1tot+12}$	34,658	0.16	15.86	-0.23	2,583	-3.13	23.94	-3.10
$Ret_{t+1tot+24}$	34,658	0.49	22.35	-0.19	2,583	9.75	39.53	5.00

Panel B: Stock-Level Monthly Trading and Holding Intensity (in percent)

	N	Mean	Std.Dev.	25^{tho}%ile	Median	75^{tho}%ile
Retail Trading (Detrended)	500,539	0.003	0.447	-0.070	-0.002	0.062
Retail Holding (Detrended)	496,145	0.098	1.253	-0.210	-0.014	0.191
Short Seller Holding (Detrended)	492,999	-0.141	2.340	-0.662	0.015	0.652

Panel C: Anomaly-Level Monthly Trading and Holding intensity (in percent)

	N	Mean	Std.Dev.	25^{tho}%ile	Median	75^{tho}%ile
			Jan 2009 to Feb 2020			
$Retail\ ATI_{t-2tot}$	34,668	-0.017	0.181	-0.087	-0.007	0.063
$Retail\ AHI_{t-2tot}$	34,668	-0.471	1.199	-0.927	-0.303	0.142
$Short\ Seller\ AHI_{t-2tot}$	34,668	0.908	1.755	-0.207	1.397	1.911
			Mar 2020 to Dec 2020			
$Retail\ ATI_{t-2tot}$	2,575	-0.145	0.571	-0.352	-0.067	0.063
$Retail\ AHI_{t-2tot}$	2,583	-1.114	3.648	-3.428	-0.477	1.220
$Short\ Seller\ AHI_{t-2tot}$	2,583	1.765	3.219	-0.207	1.397	3.166

Table 2 Retail Anomaly Trading Intensity and Anomaly Performance

Panel A of Table 2 shows cross-sectional evidence of retail anomaly trading and anomaly performance. This analysis uses the sample period from 2009 to Feb 2018, defined as the anomaly ranking months, omitting the COVID-19 period for a dedicated event study. Anomaly portfolios are sorted into 5 groups according to average retail ATI (in percent) a quarter before anomaly formation. Column (1) shows the average percentage of the time that retail investors trade against anomalies within each group. Column (2) shows the average ATI a quarter before anomaly formation. Column (3) to column (8) shows cumulative anomaly portfolio alphas. T-stats are adjusted for serial correlation using Newey-West approach. α_{1to6}, α_{1to12} , and α_{1to24} are adjusted for 6, 12, and 24 lags. Panel B shows the result using the Jagadeesh-and-Titman-approach to address the overlapping issue. Monthly returns are compounded to get the holding period return. T-statistics in panel B are on monthly alphas. The alphas of the anomaly portfolios that retail investors trade against are highlighted in bold.

Panel A: Newey West Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Before Ranking (%)		**After Ranking** (%)				
	% Time with Neg Trade	Trading	Return	Short-term Return		Long-term Return		
		ATI_{t-2tot}	α_{t-2tot}	α_{t+1}	$\alpha_{t+1tot+6}$	$\alpha_{t+1tot+12}$	$\alpha_{t+1tot+24}$	$\alpha_{t+2tot+24}$
R1 (Pos)	28	0.063***	4.03***	0.07	0.26	0.17	-1.31	-1.38
		(6.59)	*(8.29)*	*(0.66)*	*(0.45)*	*(0.16)*	*(-0.85)*	*(-1.36)*
R2	43	0.025***	3.93***	0.13	0.41	0.55	0.32	0.14
		(5.36)	*(8.29)*	*(1.51)*	*(1.10)*	*(1.00)*	*(0.48)*	*(0.27)*
R3	52	-0.009	0.19	**0.28*** **	**1.46*** **	**3.33*** **	**8.35*** **	**7.95*** **
		(-1.61)	*(0.62)*	*(2.76)*	*(4.21)*	*(5.37)*	*(8.99)*	*(9.99)*
R4	61	-0.030***	-3.54***	**0.43*** **	**2.69*** **	**5.86*** **	**13.23*** **	**12.69*** **
		(-2.90)	*(-6.97)*	*(2.84)*	*(4.47)*	*(7.09)*	*(12.67)*	*(13.57)*
R5 (Neg)	77	-0.096***	-3.68***	**0.83*** **	**4.57*** **	**10.17*** **	**23.49*** **	**22.48*** **
		(-5.43)	*(-3.99)*	*(2.80)*	*(3.88)*	*(5.01)*	*(10.25)*	*(10.95)*

Panel B: Jagadeesh and Titman Approach

	(1)	(2)	(3)	(4)	(5)
	After Ranking (%)				
	Short-term Return (%)		Long-term Return (%)		
	α_{t+1}	$\alpha_{t+1tot+6}$	$\alpha_{t+1tot+12}$	$\alpha_{t+1tot+24}$	$\alpha_{t+2tot+24}$
R1 (Pos)	0.00	-0.17	-0.93	-2.25	-2.25
	(0.01)	*(-0.31)*	*(-0.87)*	*(-1.23)*	*(-1.28)*
R2	0.11	0.38	0.76	1.03	0.92
	(1.45)	*(1.11)*	*(1.34)*	*(1.03)*	*(0.98)*
R3	**0.43*** **	**2.37*** **	**4.89*** **	**10.45*** **	**9.98*** **
	(3.31)	***(3.17)***	***(3.32)***	***(3.57)***	***(3.57)***
R4	**0.59*** **	**3.63*** **	**7.31*** **	**14.95*** **	**14.28*** **
	(2.84)	***(2.97)***	***(3.05)***	***(3.16)***	***(3.17)***
R5 (Neg)	**1.06*** **	**5.31*** **	**10.69*** **	**23.05*** **	**21.76*** **
	(4.04)	***(3.48)***	**(3.50)**	**(3.80)**	***(3.78)***

Table 3 Retail Anomaly Trading Intensity Groups

Table 3 groups retail anomaly portfolio trading intensity a quarter before anomaly formation according to group classification by Hou, Xue, and Zhang. (2020), Daniel, Hirshleifer, and Sun (2020), Akbas, Armstrong, Sorescu, and Subrahmanyam (2015), Kumar (2009), and Birru (2018) and report average retail ATI within a group (in percent). Column (1) lists related literature. Column (2) lists anomaly group names in the corresponding literature. Column (3) shows the retail ATI a quarter before anomaly formation within the group and column (4) shows t-stats calculated from time series grouped retail ATI. Column (5) shows the average percentage of the time that retail investors trade against each group classification.

(1) Literature	(2) Group Name	(3) ATI_{t-2tot}	(4) T-Stats	(5) % Time with Neg Trade
HXZ(2020)	Intangible	0.007***	(4.80)	36
	Investment	0.023***	(6.29)	24
	Momentum	0.020**	(2.32)	51
	Profitability	-0.049***	(-6.78)	76
	TradingFriction	-0.056***	(-8.69)	77
	VvsG	-0.025***	(-3.27)	65
DHS(2020)	Long-Horizon	-0.017	(-1.37)	47
	Short-Horizon	0.016**	(2.19)	46
AASS(2015)	Investment	0.058***	(7.28)	27
	Non-Investment	-0.035***	(-4.99)	70
Kumar (2009)	Lottery	-0.083***	(-8.94)	78
	Non-Lottery	-0.013***	(-4.25)	60
Birru(2018)	Long Leg	0.007**	(2.01)	46
	Short Leg	-0.080***	(-8.36)	78

Table 4 Retail-Related Factor Models to Explain Long-horizon Anomaly Performance

Table 4 Panel A reports alphas from the regression of FIN factor (Daniel et al. 2020), the retail factor formed from the group of anomalies that retail investor trade the mostly negative, denoted R5F, and the retail factor directly related to trading, denoted TF, on other set of factor models: capital asset pricing model, denoted CAPM; the three- and five-factor model of Fama and French (1993) and Fama and French (2015), denoted FF3 and FF5; the four-factor model of Carhart (1997); the behavioral model introduced by Daniel, Hershleifer, and Sun (2020); and the two retail-related models introduced in this study. Panel B shows the average absolute alphas, average t-statistics, and number of anomalies with significant alphas (t-statistics >= 1.96) from the time-series regression of R4 and R5 group anomaly Jagadeesh-Titman type 24-month monthly portfolio returns on different factor models. The sample period is from 2011 to 2022, defined as anomaly return dates. Standard errors are adjusted for serial correlation with Newey-West method of 6 lags.

Panel A: Spanning Test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α	CAPM	FF3	Carhart4	FF5	DHS	MKT + R5F	MKT + TF
FIN	0.43	0.36**	0.38**	0.12		-0.19	-0.19
	(1.38)	(2.41)	(2.46)	(1.15)		(-1.05)	(-0.80)
R5F	1.14***	0.90***	0.82***	0.63***	0.50**		0.24
	(2.84)	(3.24)	(2.89)	(3.50)	(2.27)		(0.86)
TF	-0.61***	-0.54***	-0.57***	-0.46***	-0.50***	-0.35***	
	(-4.08)	(-4.90)	(-4.60)	(-5.69)	(-6.69)	(-3.51)	

Panel B: Model Performance Comparison

(1)	(2)	(3)	(4)
Model	Average Alpha	Average T-stats	# Anomalies with α**
CAPM	0.74%	1.87	47
FF3	0.53%	1.82	41
Carhart4	0.46%	1.58	32
FF5	0.38%	1.64	37
DHS	0.38%	1.16	16
MKT + R5F	0.21%	0.75	6
MKT + TF	0.23%	0.78	6
DHS + R5F	0.18%	0.69	6
DHS + TF	0.19%	0.66	6

Table 5 Retail Trading Persistence

Table 5 shows the results of retail ATI magnitude and persistence correlation. Using an event window of 24 months after anomaly ranking month, anomaly retail holding persistence measure is calculated by regressing current event time AHI on lagged event time AHI and taking the AR(1) coefficient. Retail ATI is standardized between 0 and 1 using the cross-sectional retail ATI at each month. The sample period is from 2009 to 2020, defined as anomaly ranking months. Standard errors are clustered by anomaly and date.

	(1) Holding Persistence (Against)	(2) Holding Persistence (With)
$\mathrm{ATI}_{\mathrm{t-2tot}}$ (%)	-0.01***	0.00
	(-3.78)	(1.18)
Date FE	YES	YES
Observations	22,556	14,544
Adjusted R-squared	0.095	0.050

Table 6 Risk Management Measures

Table 6 shows average standard deviation and 1% Value-at-Risk for anomaly portfolios within each R1 to R5 group. Retail investors trade with R1 and R2 group of anomalies, and trade against R3 to R5 group of anomalies. For an anomaly portfolio, 1% VaR is calculated as the difference between the times-series mean of alphas and the product of 1% z-score (2.33) and standard deviation of anomaly portfolio alphas, that is, $1\%VaR = \overline{\alpha} - 2.33 \times \sigma_{\alpha}$. Similarly, $0.1\%VaR = \overline{\alpha} - 3.09 \times \sigma_{\alpha}$.

	(1)	(2)	(3)	(4)	(5)
			Std	1%VaR(%)	0.1%VaR(%)
Short-term	R1 (Pos)	α_{t+1}	4.28	-9.85	-12.73
	R2	α_{t+1}	4.01	-9.1	-11.98
	R3	α_{t+1}	4.33	-9.79	-12.90
	R4	α_{t+1}	4.62	-10.33	-13.64
	R5 (Neg)	α_{t+1}	5.46	-11.89	-15.76
Long-term	R1 (Pos)	$\alpha_{t+2tot+24}$	18.99	-44.1	-58.53
	R2	$\alpha_{t+2tot+24}$	18.36	-42.14	-56.10
	R3	$\alpha_{t+2tot+24}$	21.64	-43.65	-60.10
	R4	$\alpha_{t+2tot+24}$	21.31	-38.39	-54.59
	R5 (Neg)	$\alpha_{t+2tot+24}$	26.11	-42.02	-61.86

Table 7 Retail Trading Before and During the Pandemic

Table 7 shows retail trading volume and ATI results using the sample period from 2015 to 2020, defined as anomaly ranking months. Panel A shows the regression result of retail trading volume (in percent) at the stock level to a *Post* dummy, where the post dummy is equal to 1 if the sample is from March 2020 to December 2020, and is equal to 0 if otherwise. Panel B reports regression results of retail ATI (in percent) on *Treat* and *Treat*×*Post* interaction term. Standard errors are clustered by date and anomaly.

Panel A: Retail Trading Volume

	(1) Retail Volume (%)
Post	2.49***
	(9.16)
Constant	2.13***
	(46.42)
Firm FE	YES
Observations	250,113
Adjusted R-squared	0.342

Panel B: Retail ATI

	(1)	(2)	(3)	(4)
	Before Ranking		After Ranking	
	ATI_{t-2tot}	ATI_t	ATI_{t+1}	$ATI_{t+1tot+3}$
Post x Treat	-0.37***	-0.13***	-0.07***	-0.16***
	(-3.57)	(-3.54)	(-2.82)	(-2.86)
Treat	-0.12***	-0.05***	-0.00	0.02
	(-6.73)	(-6.20)	(-0.67)	(1.03)
Post	0.01	-0.00	-0.00	-0.02
	(0.17)	(-0.12)	(-0.16)	(-0.48)
Anomaly FE	NO	NO	NO	NO
Observations	12,717	13,079	12,898	12,717
Adjusted R-squared	0.153	0.130	0.025	0.027

Table 8 Diff-in-Diff Results for Anomalies Portfolio Returns

Table 8 shows diff-in-diff regression results of anomaly portfolio alphas (in percent) a quarter before and after formation on the treat dummy and Treat×Post interaction term. Standard errors are clustered by date and anomaly.

	(1)	(2)	(3)	(4)
	Before Ranking		After Ranking	
	α_{t-2ot}	α_t	α_{t+1}	$\alpha_{t+1ot+3}$
Treat x Post	-12.29***	-4.40***	-2.79**	-7.54***
	(-5.49)	(-3.34)	(-2.39)	(-4.53)
Treat	-5.27***	-3.53***	0.63**	2.14***
	(-3.89)	(-4.19)	(2.07)	(4.02)
Date FE	YES	YES	YES	YES
Observations	12,717	13,079	12,898	12,536
Adjusted R-squared	0.109	0.078	0.114	0.122

Table 9 Retail Anomaly Trading During the 1st Stimulus Check

Table 9 panel A show DID results for treatment and control group anomaly portfolios separately using the 1st stimulus check payment schedule. The sample period is Jan 2, 2020, to Jun 12, 2020. The post-disbursement period is from April 9 to June 12. The daily abnormal retail ATI are formed by dollar amount traded, [1150,1200) and (1200, 1250]. $\mathrm{I}(< 1200)$ is an indicator and takes on the value of 1 if the retail order falls within the lower cost range [1150,1200), and 0 otherwise. Panel B shows triple difference results by adding an announcement and crash dummy. Congress passed the CARES Act on Mar 27, 2020. Ann is a dummy that takes the value of 1 for ATI during the first stimulus check announcement week, specifically from -9 to -6 days relative to the event day. The S&P 500 index dropped by 29% from March 2 to March 16, 2020. We define *Crash=1* if trading intensity falls into the event day of -26 to -13. Standard errors are clustered by date and anomaly.

Panel A: Diff-in-Diff

	(1)	(2)	(3)	(4)
	Treat		Control	
	$BuyATI_t$	$SellATI_t$	$BuyATI_t$	$SellATI_t$
I (<$1200) x Post	**-0.009*** **	**-0.001**	0.001	0.001
	(-2.76)	(-0.32)	(0.95)	(0.42)
I (<$1200)	0.001	-0.001	0.000	0.000
	(0.73)	(-0.27)	(-0.60)	(-0.67)
Date FE	YES	YES	YES	YES
Observations	22,826	22,826	17,402	17,402
Adjusted R-squared	0.24	0.239	0.006	0.005

Panel B: Triple Difference

	(1)	(2)	(3)	(4)	(5)	(6)
		$BuyATI_t$			$SellATI_t$	
Treat	-0.057***	-0.057***	-0.059***	-0.056***	-0.056***	-0.057***
	(-9.54)	(-9.70)	(-10.32)	(-9.54)	(-9.64)	(-9.96)
I (<$1200)	-0.000	-0.000	-0.000	-0.000	-0.001	-0.000
	(-0.61)	(-0.82)	(-0.43)	(-0.67)	(-1.06)	(-0.39)
Treat x Post	-0.074***	-0.074***	-0.072***	-0.076***	-0.076***	-0.075***
	(-9.00)	(-8.90)	(-8.60)	(-8.87)	(-8.79)	(-8.54)
Treat x I (<$1200)	**0.002**	**0.002**	**0.003**	**-0.000**	**0.000**	**0.000**
	(0.90)	(1.09)	(1.58)	(-0.07)	(0.16)	(0.10)
I (<$1200) x Post	0.001	0.001	0.001	0.001	0.001	0.000
	(0.96)	(1.02)	(0.80)	(0.42)	(0.60)	(0.31)
Treat x Ann		0.000			-0.001	
		(0.01)			(-0.19)	
I (<$1200) x Ann		0.002			0.004*	
		(0.71)			(1.78)	
Treat x I (<$1200) x Post	**-0.010*** **	**-0.010*** **	**-0.011*** **	**-0.002**	**-0.002**	**-0.002**
	(-2.75)	(-2.82)	(-3.09)	(-0.45)	(-0.57)	(-0.52)
Treat x I (<$1200) x Ann		-0.008*			-0.008	
		(-1.90)			(-1.15)	
Treat x Crash			0.009			0.005
			(1.09)			(0.76)
Doldum x Crash			-0.000			-0.000
			(-0.32)			(-0.27)
Treat x Doldum x Crash			-0.007			-0.002
			(-1.22)			(-0.36)
			-0.059***			-0.057***
Date FE	YES	YES	YES	YES	YES	YES
Observations	40,228	40,228	40,228	40,228	40,228	40,228
Adjusted R-squared	0.342	0.342	0.343	0.345	0.345	0.345

Internet Appendix

Figure IA1 Retail Before Ranking ATI Time-Series

Figure IA1 plots retail ATI_{t-2tot} from R1(retail investors trade the most positively) to R5(retail investors trade the most negatively) over time.

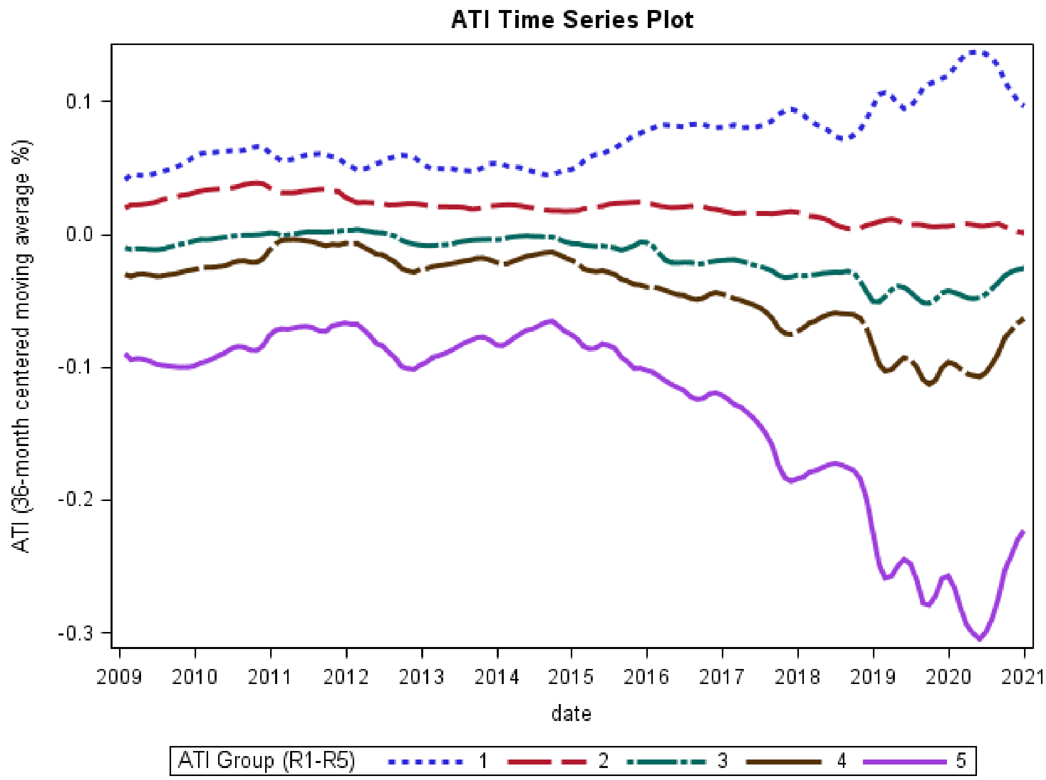
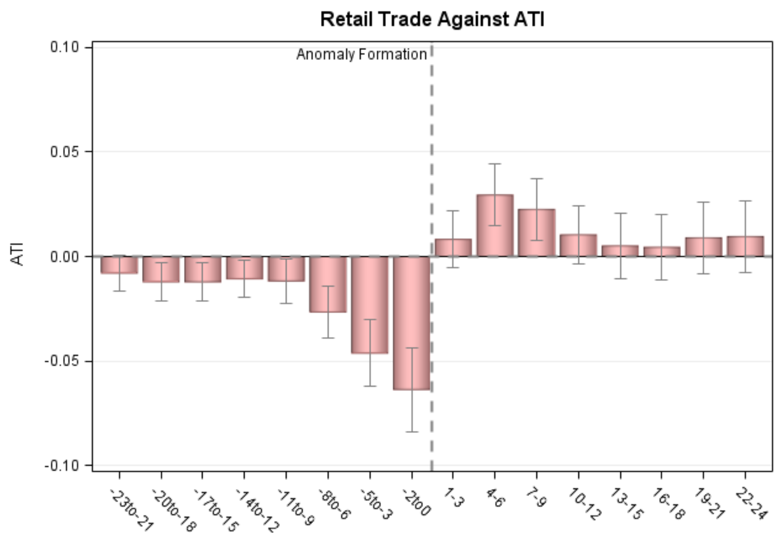


Figure IA2 Retail ATI

Figure IA2 shows retail ATI in percent for the anomalies that retail investor trade against. Sample period is from 2009 to 2020, defined as anomaly ranking months.

Panel A: Retail Trade Against ATI



Panel B: Retail Trade with ATI

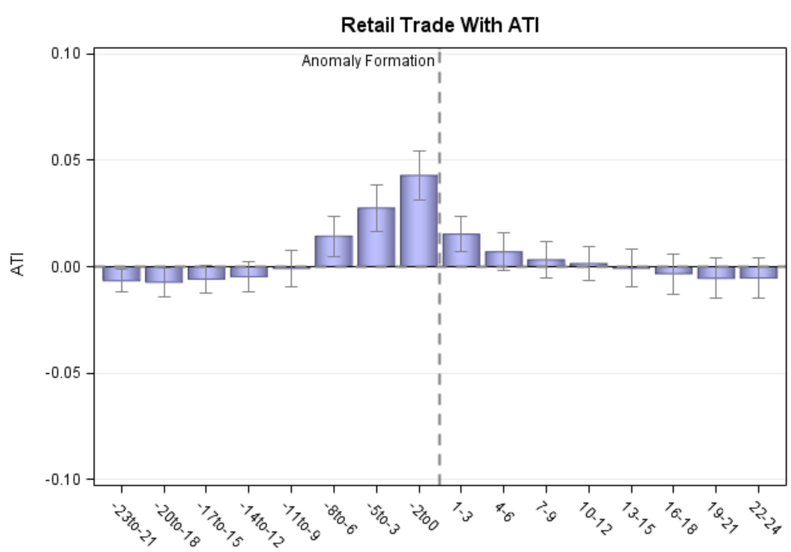


Table IA1 Retail Anomaly Trading Intensity and Anomaly Performance–With COVID-19

Table IA1 shows cross-sectional evidence of retail anomaly trading and anomaly performance, using the sample period from January 2009 to December 2020, defined as anomaly ranking. Anomaly portfolios are sorted into 5 groups according to average retail ATI (in percent) a quarter before anomaly formation. Column (1) to column (6) shows cumulative anomaly portfolio alphas. T-stats are adjusted for serial correlation using Newey-West. α_{1to6}, α_{1to12} , and α_{1to24} are adjusted for 6, 12, and 24 lags. The alphas of the anomaly portfolios that retail investors trade against are highlighted in bold.

	Before Ranking	**After Ranking**				
	(1)	(2)	(3)	(4)	(5)	(6)
	Return (%)	Short-term Return (%)		Long-term Return (%)		
	$\alpha_{\mathrm{t-2tot}}$	α_{t+1}	$\alpha_{t+1tot+6}$	$\alpha_{t+1tot+12}$	$\alpha_{t+1tot+24}$	$\alpha_{t+2tot+24}$
R1 (MostPos)	5.04***	0.15	0.89	1.09	-0.44	-0.58
	(8.38)	*(1.47)*	*(1.53)*	*(1.06)*	*(-0.25)*	*(-0.53)*
R2	4.21***	0.20***	0.77**	0.86*	0.86	0.65
	(21.12)	*(2.59)*	*(2.50)*	*(1.66)*	*(1.26)*	*(1.21)*
R3	**-0.19**	**0.14**	**0.84****	**2.31*****	**6.62*****	**6.42*****
	(-0.65)	*(1.49)*	*(1.97)*	*(2.81)*	*(5.92)*	*(6.00)*
R4	**-4.25*****	**0.24***	**2.02*****	**4.70*****	**11.17*****	**10.94*****
	(-8.22)	*(1.49)*	*(2.91)*	*(4.19)*	*(7.78)*	*(9.00)*
R5 (MostNeg)	**-5.35*****	**0.48***	**3.23*****	**7.46*****	**18.56*****	**18.14*****
	(-5.39)	*(1.72)*	*(2.65)*	*(3.28)*	*(6.13)*	*(7.22)*

Table IA2 Retail Anomaly Trading Intensity and Anomaly Performance–Drop High Fee Stocks

Table IA2 shows cross-sectional evidence of retail anomaly trading and anomaly performance by excluding high fee stocks. Following Muravyev, Pearson, and Pollet (2023), we use the 'Indicative Fee' from the Makit short-selling dataset for short-selling costs, and filter out stocks whose fee exceeds 1% per annum as of the end of the previous month. Anomaly portfolios are sorted into 5 groups according to average retail anomaly trading intensity (ATI) in percent a quarter before anomaly formation. Column (1) to column (6) shows cumulative anomaly portfolio alphas. T-stats are adjusted for serial correlation using Newey-West. α_{1to6}, α_{1to12} , and α_{1to24} are adjusted for 6, 12, and 24 lags. The alphas of the anomaly portfolios that retail investors trade against are highlighted in bold.

	Before Ranking	**After Ranking**				
	(1)	(2)	(3)	(4)	(5)	(6)
	Return (%)	Short-term Return (%)		Long-term Return (%)		
	$\alpha_{\mathrm{t-2tot}}$	α_{t+1}	$\alpha_{t+1tot+6}$	$\alpha_{t+1tot+12}$	$\alpha_{t+1tot+24}$	$\alpha_{t+2tot+24}$
R1 (MostPos)	3.41***	0.14	0.55	0.65	-0.43	-0.58
	(7.90)	*(1.46)*	*(0.98)*	*(0.62)*	*(-0.28)*	*(-0.39)*
R2	3.82***	0.11	0.42	0.58	0.18	0.02
	(21.57)	*(1.22)*	*(1.13)*	*(1.07)*	*(0.28)*	*(0.03)*
R3	**0.19**	**0.22****	**1.17*****	**2.70*****	**6.84*****	**6.56*****
	(0.65)	*(2.47)*	*(3.25)*	*(4.46)*	*(9.46)*	*(9.11)*
R4	**-3.38*****	**0.30****	**2.09*****	**4.71*****	**10.81*****	**10.43*****
	(-7.09)	*(2.08)*	*(3.59)*	*(6.17)*	*(13.08)*	*(12.61)*
R5 (MostNeg)	**-3.56*****	**0.53***	**2.90****	**7.17*****	**17.29*****	**16.72*****
	(-4.17)	*(1.87)*	*(2.55)*	*(3.81)*	*(7.86)*	*(7.63)*

Table IA3 DID Retail Anomaly Trading by Long and Short Leg

Table IA3 shows retail anomaly trading results using the diff-in-diff sample period from January 2015. The table reports regression results of retail anomaly long-short $\mathrm{ATI}_{\mathrm{t-2tot}}$ (in %), together with short and long leg $\mathrm{ATI}_{\mathrm{t-2tot}}$ on *Treat* and *Treat×Post* interaction term, where the post dummy is equal to 1 if in the pandemic, from March 2020 and onwards, and is equal to 0 if otherwise. Standard errors are clustered by date and anomaly.

	(1) Long-Short Leg $\mathrm{ATI}_{\mathrm{t-2tot}}$	(2) Short Leg $\mathrm{ATI}_{\mathrm{t-2tot}}$	(3) Long Leg $\mathrm{ATI}_{\mathrm{t-2tot}}$
Treat x Post	-0.37***	0.29***	-0.09***
	(-3.57)	(3.46)	(-3.57)
Treat	-0.12***	0.09***	-0.03***
	(-6.73)	(6.29)	(-6.27)
Post	0.01	0.12*	0.12**
	(0.17)	(1.82)	(2.35)
constant	0.01*	0.02**	0.02***
	(1.68)	(2.16)	(3.17)
Anomaly FE	NO	NO	NO
Adj R-squared	0.153	0.153	0.089
Observations	12,717	12,717	12,718

Table IA4 DID Anomaly Performance by Long and Short Leg

Panel A: Anomaly Performance

	(1) Long-Short α_{t+1}	(2) Short-Leg α_{t+1}	(3) Long-Leg α_{t+1}
Treat x Post	**-2.79****	**2.00****	**-0.79****
	(-2.39)	(2.15)	(-2.44)
Treat	0.63**	-0.45*	0.18**
	(2.07)	(-1.83)	(2.25)
Date FE	YES	YES	YES
Observations	12,898	12,898	12,899
Adjusted R-squared	0.114	0.212	0.028

Panel B: Parallel Trend

	(1) Long-Short α_{t+1}	(2) Short-Leg α_{t+1}	(3) Long-Leg α_{t+1}
Treat x Post(2016)	-1.40	1.26	-0.14
	(-1.20)	(1.36)	(-0.51)
Treat x Post(2017)	-0.84	1.02	0.19
	(-1.05)	(1.62)	(0.82)
Treat x Post(2018)	-1.14	0.97	-0.16
	(-1.56)	(1.53)	(-1.23)
Treat x Post(2019)	-0.95	1.03	0.08
	(-1.02)	(1.33)	(0.42)
Treat x Post(2020)	**-3.68*****	**2.86*****	**-0.82****
	(-2.82)	(2.72)	(-2.31)
Date FE	YES	YES	YES
Anomaly FE	YES	YES	YES
Observations	12,534	12,534	12,535
Adjusted R-squared	0.120	0.216	0.039

Panel C Anomalies that Retail Investors Trade More Against

	(4)	(5)	(6)
	<50%	<40%	<30%
	α_{t+1}	α_{t+1}	α_{t+1}
Treat x Post	**-2.60****	**-2.79****	**-3.11****
	(-2.45)	(-2.39)	(-2.32)
Treat	0.57**	0.63**	0.72**
	(2.12)	(2.07)	(2.13)
Date FE	YES	YES	YES
Observations	14,743	12,898	11,073
Adjusted R-squared	0.114	0.114	0.111

Table IA5 RD Retail Anomaly Trading Intensity within Treatment Group by Long and Short Leg

	(1)	(2)	(3)	(4)	(5)	(6)
		$BuyATI_t$			$SellATI_t$	
	Long-Short	Short-Leg	Long-Leg	Long-Short	Short-Leg	Long-Leg
I (<$1200) x Post	-0.009***	0.010***	0.002**	-0.001	0.001	-0.000
	(-2.76)	(2.89)	(2.29)	(-0.32)	(0.21)	(-0.38)
I (<$1200)	0.001	-0.002	-0.001	-0.001	-0.000	-0.001*
	(0.73)	(-0.90)	(-1.16)	(-0.27)	(-0.13)	(-1.91)
Date FE	YES	YES	YES	YES	YES	YES
Observations	22,826	22,826	22,826	22,826	22,826	22,826
AdjustedR-squared	0.240	0.411	0.179	0.239	0.407	0.178