Buying the Dip: Return-Contingent Retail Trades

Xiao Yin and Dongchen Zou *

March 26, 2026

Abstract

Using a novel dataset of actual retail order flows for 2017–2024, we document an asymmetric relation between retail activity and past returns: inflows surge following large price declines but barely respond to gains. This "buying the dip" pattern is stable across return horizons and persists throughout the sample period. We show the behavior is not driven by clientele effects but instead reflects institutional accommodation. Retail purchases are strongest when mutual fund redemptions, under-performance pressure, and tightening risk constraints force other institutions to sell. A quasi-natural experiment using the Melvin Capital liquidation confirms the mechanism. Retail flows in distressed states predict return reversals and lower bid-ask spreads, consistent with state-contingent liquidity provision. However, while the average retail dip-buying trade is profitable, the dollar-weighted return is negative, consistent with retail investors holding a rational expectation of mean reversion but lacking stock-level selection skill.

JEL codes: G11, G12, G14, G23

Keywords: retail trading, buying the dip, institutional fire sales, liquidity provision, return reversal, order flow

*Yin: University College London, xiao.yin@ucl.ac.uk; Zou: Indiana University, dongzou@iu.edu. We thank seminar participants at Deakin University and Australian National University for helpful comments and suggestions. All errors are our own.

1 Introduction

Retail trading activity has surged to unprecedented levels. Prior to the COVID-19 pandemic, retail investors accounted for less than 10% of U.S. equity market order flow; by late April 2025, that share had climbed to 36%.[1] Driven by the proliferation of zero-commission brokerages and the gamification of mobile trading platforms, retail investors have become a major and growing force in U.S. equity markets. The profound impact of this demographic was most vividly illustrated during the "meme stock" episodes, which demonstrated how concentrated retail attention can overwhelm institutional short sellers and temporarily detach prices from fundamental values. However, beyond these episodic anomalies, the systematic trading behavior of modern retail investors remains debated (Barber et al., 2022; Kaniel et al., 2008). Specifically, how do retail investors allocate capital in response to market movements, and what are the ultimate market outcomes of their trades?

This paper addresses these questions by introducing a novel, publicly accessible S&P Retail Flow dataset constructed from actual retail transactions between 2017 and 2024. Because this dataset relies on ground-truth trades rather than inferential algorithms applied to public tape data—which recent literature has shown to be increasingly misclassified (Battalio et al., 2023; Barber et al., 2024)—it provides an exceptionally clear window into retail behavior. The data is also broadly representative: according to S&P Global, the sample accounts for approximately 70–80% of total U.S. retail trading activity.

Using this data, we document a pervasive stylized fact: retail investors systematically "buy the dip." We uncover a robust, return-contingent asymmetry in retail order flows. Retail traders supply massive inflows following negative stock returns but rarely exhibit a symmetric propensity to sell following positive returns. This behavior generates a distinct nonlinear asymmetry in the relationship between retail order flows and lagged returns, distinguishing retail traders from institutional investors, who generally exhibit momentum-like trading patterns in distressed states.

The striking magnitude of this asymmetric behavior raises a fundamental question

[1] J.P. Morgan Asset Management (2025).

about its economic driver. We broadly evaluate two mechanisms. The first is a *Retail Clientele* hypothesis, which suggests that retail investors inherently prefer distressed stocks due to behavioral biases, such as heightened attention to "biggest losers" and contrarian sentiment, or mechanical shifts in firm characteristics, such as an inflated book-to-market ratio. However, we find that retail flows exhibit remarkably weak persistence at the stock level, and the inflows cannot be fully explained by retail attention metrics, sentiment proxies, or changing firm fundamentals. Furthermore, stocks that receive heavy retail inflows subsequently experience positive next-period returns, inconsistent with the negative risk premium typically associated with retail investors having preferences over these stocks (Pástor et al., 2021).

Instead, our evidence strongly supports an *Institutional Accommodation* hypothesis. Financial institutions frequently face strict mandates, fund redemptions, or Value-at-Risk (VaR) constraints that force them to liquidate positions mechanically following negative returns. This institutional imperative to sell the dip creates an excess supply of distressed equity. Because retail investors are largely unconstrained by such mandates, they step into this newly created habitat, supplying liquidity conditional on negative returns and actively accommodating the institutional fire sales.

We test this mechanism using the January 2021 GameStop (GME) short squeeze and the subsequent near-collapse of Melvin Capital as a quasi-natural experiment. Facing unprecedented margin calls on its short book, Melvin Capital was forced to rapidly liquidate its existing portfolio to raise cash. We show that retail investors aggressively bought the specific loser stocks that Melvin was forced to sell, perfectly mirroring the exogenous institutional outflow. We confirm this mechanism broadly in the cross-section by documenting that retail inflows surge when active mutual funds face severe outflow-driven selling pressure Coval and Stafford (2007) and under-performance "alpha" pressure. Conversely, during highly anticipated and passive index rebalancing events (e.g., S&P 500 or Russell 1000/2000 transitions) where no "sell-the-loser" mandate exists, the retail response is notably muted.

Finally, we examine the market outcomes and profitability of this state-contingent liquidity provision. We find that retail order flow into distressed stocks predicts higher

next-month returns, and dynamic event-time specifications confirm that retail inflows help stabilize prices and normalize deteriorating bid-ask spreads. However, an analysis of retail profitability reveals a critical capacity constraint. While the simple average retail trade in a distressed stock is profitable, the dollar-weighted profit is distinctly negative. Retail investors, while consistent with a rational expectation of mean reversion following negative returns, do not exhibit sufficient skills in stock selection to profit from the trades.

Related Literature. This paper relates to three strands of literature. The first concerns the measurement of retail trading. Early work proxies retail activity using small trade size (Hvidkjaer, 2008a). Boehmer et al. (2021) introduce the widely used BJZZ algorithm, which identifies retail trades from sub-penny price improvements granted by wholesalers, and show that retail order imbalances predict short-term returns. Recent work raises concerns about inferential approaches: Battalio et al. (2023) document growing difficulties in cleanly identifying retail trades from public data, and Barber et al. (2024) show that BJZZ captures only ap-proximately 30% of a known retail sample and propose a modified algorithm. We contribute to this literature by providing and validating a ground-truth retail flow measure built from actual retail transactions, and by using it to reveal systematic behavioral patterns that inferential proxies obscure.

The second strand concerns retail investor behavior and clientele effects. The noise-trader framework of De Long et al. (1990) characterizes retail investors as uninformed agents whose sentiment-driven demand creates mispricing. Consistent with behavioral biases, Barber and Odean (2008) show that retail investors are net buyers of attention-grabbing stocks experiencing extreme price moves or heavy media coverage. Barber et al. (2022) document attention-induced herding among Robinhood users that generates return predictability, consistent with retail investors chasing salient price movements. Kaniel et al. (2008) document that individual investor trading is contrarian at short horizons and predicts positive subsequent returns, suggesting informed contrarian behavior. On the clientele side, Dorn and Huberman (2010) find that retail investors cluster in stocks with particular risk characteristics, while Laarits and Sammon (2025) show that retail flows exhibit strong stock-level persistence consistent with stable demand habitats. Luo et al.

(2025) further document contrarian retail behavior around news events and its relation to the momentum anomaly. We extend this literature by showing that the contrarian response is sharply asymmetric and state-dependent—concentrated in negative-return states—and is not driven by stable stock-level clienteles but by episodic institutional selling pressure.

The third strand concerns institutional sales and retail liquidity. Coval and Stafford (2007) show that mutual fund redemptions generate fire-sale price pressure in commonly held stocks with predictable subsequent reversals. Barrot et al. (2016) investigate whether retail traders are compensated for providing liquidity, finding positive average returns to retail contrarian trading in certain market conditions. We connect these studies by documenting that retail liquidity provision in distressed states arises specifically as a response to constrained institutional selling, and by showing that while this provision stabilizes prices on average, it exposes retail investors to uncompensated risk in the largest fire-sale episodes.

2 Data

2.1 The S&P Retail Flow Dataset

The primary data source for our empirical analysis is a novel, publicly accessible retail order-flow dataset provided by S&P, covering the period from 2017 to 2024 at a daily frequency. Unlike traditional measures of retail trading that rely on proxies or inferential algorithms applied to public tape data, the S&P Retail Flow dataset is constructed from actual retail trades. This dataset provides a ground-truth look into the trading behavior of retail investors. As shown in Figure 1, these actual retail trades represent a substantial portion of market activity, accounting for approximately 70% to 80% of total gross trading volume for the majority of our sample period.

2.2 Comparison with TAQ-Inferred Measures

Historically, the literature has relied on Trade and Quote (TAQ) data to approximate retail trading activity. Early methods utilized trade size, classifying small trades as retail Hvidkjaer (2008b). More recently, Boehmer et al. (2021) (BJZZ) introduced a widely used algorithm that identifies retail trades based on sub-penny price improvements typically

granted by wholesalers.

However, recent evidence highlights the limitations of these inferential approaches. Battalio et al. (2023) document the increasing difficulty of cleanly identifying retail trades from public tape data. Furthermore, Barber et al. (2024) (BHJOS) show that the BJZZ algorithm only identifies approximately 30% of trades within a known retail sample, prompting them to modify the algorithm by signing trades using quoted spread midpoints. Because our dataset is built from actual retail transactions, it serves as a reliable benchmark to evaluate the efficacy of these proxies.

2.3 Time-Series and Cross-Sectional Validations

We benchmark our S&P Retail Flow measure against several established proxies: the BJZZ sub-penny measure, the BHJOS modified measure, total market flow, and two institutional flow proxies. Specifically, we proxy for institutional activity using large trades inferred from orders exceeding $50,000 (Hvidkjaer, 2008b), denoted as *inst50k*, and institutional trades inferred from securities lending supply (Barardehi et al., 2024), denoted as *instsec*.

Table 1 reports the aggregate time-series correlations. In the aggregate time series, the actual retail flow exhibits a correlation of roughly zero (-0.01) with the BJZZ-identified flows, suggesting that the sub-penny algorithm struggles to capture macro-level shifts in retail sentiment over time. The BHJOS measure performs modestly better, exhibiting a positive time-series correlation of 0.33 with our actual retail flow. Notably, actual retail flow is negatively correlated with total market flow (-0.17) and large institutional trades (-0.23).

In the cross-section, the proxies align slightly better with actual retail flows but remain imperfect, as detailed in Table 2 and illustrated in Figure 2. The BJZZ measure exhibits a low cross-sectional correlation of 0.19 with actual retail flow, while the BHJOS modification improves this correlation to 0.34. Total order flow remains negatively correlated with retail flow (-0.32) in the cross-section.

To further illustrate the precision of our dataset, we examine the GameStop (GME) short squeeze in early 2021. As shown in Figure 3, during this heavily retail-driven meme-stock episode, our S&P Retail Flow measure and the BHJOS order imbalances

show massive spikes. In stark contrast, the BJZZ order flow fails to exhibit the same pattern, underscoring the risk of using sub-penny price improvements as a proxy during periods of extreme market volatility and heightened retail attention.

3 Buying the Dip

3.1 Documenting the Nonlinearity

A central premise of this paper is that retail trading behavior is highly state-dependent. Using our trade-identified S&P Retail Flow dataset, we document a robust return-contingent asymmetry in retail order flows. Specifically, retail investors systematically and aggressively "buy the dip": They supply heavy inflows following negative stock returns but do not exhibit a symmetric propensity to sell following positive returns.

This behavior generates a distinct nonlinear asymmetry in the relationship between retail order flows and lagged returns, as shown in Figure 4. The relationship is not linear: retail inflows accelerate sharply as returns move into the left tail, indicating that the buying response intensifies disproportionately as losses deepen. This convex pattern is inconsistent with simple mechanical rebalancing toward target portfolio weights, which would predict a proportional and symmetric response to price moves in both directions. Instead, retail demand surges precisely in the deepest distress states.

Crucially, the response is one-sided. Stocks with positive past returns generate little systematic retail selling. Net outflows for recent winners are small and often statistically close to zero. The contrarian pattern is therefore concentrated almost entirely in down states, distinguishing buying-the-dip from standard symmetric reversal strategies. While total market flows and non-retail flows generally exhibit a momentum-like pattern (selling the dip), retail flows act as a strong contrarian force specifically in distressed states.

This asymmetry is robust across multiple return horizons. Whether measured over the prior month, the prior 12-to-1 months, or the 24-to-12 month window, the same left-tail pattern emerges. The consistency across horizons suggests a persistent feature of retail trading behavior rather than a transient microstructure artifact. Appendix Figures A.1 and A.2 confirm robustness to log-return specification and alternative flow denominators.

The asymmetric flow-return pattern is also visible at the level of individual securities.

Figure A.4 examines two illustrative cases: Snapchat Inc. (SNAP) and Duke Energy Corp. (DUK). Both cases reveal massive retail inflows precisely coinciding with sharp price drops, followed by a normalization of retail flows as prices recover. Specifically, the correlations between retail order flows and returns are negative and significant in both cases (-0.46 and -0.37), while the total flow-return relations are modestly positive (0.12 and 0.31). The timing illustrates the reactive and episodic nature of the behavior.

3.2 The Pattern Is Stable Over Time

To assess whether buying-the-dip is concentrated in specific market episodes, Figure 5 plots average monthly retail flows for stocks sorted into momentum terciles from 2017 to 2024. The lowest-momentum stocks consistently attract the strongest retail inflows throughout the entire sample—both before the COVID-19 pandemic and after the 2020–2021 meme-stock surge. Notably, while aggregate retail trading volume and momentum-chasing activity spiked during 2020–2021 alongside the meme-stock episode, the dip-buying pattern persisted before, during, and after this period without interruption. Return-chasing was a sporadically prominent feature of the meme-stock era; buying-the-dip was not—it was present throughout. This stability across distinct market regimes indicates that buying-the-dip is a structural feature of retail trading rather than an artifact tied to any particular episode or cohort of new retail entrants.

3.3 Competing Mechanisms: Ruling Out Retail Clientele Effects

Given the striking asymmetry in retail flows toward distressed stocks, the natural question is: What drives this behavior? We broadly evaluate two competing mechanisms. The first is the Retail Clientele hypothesis, which posits that retail investors inherently prefer these stocks in down states due to behavioral biases or shifting firm fundamentals. The second is the Institutional Accommodation hypothesis, which argues that institutions are forced to sell these stocks due to mandates or constraints, and retail investors step in to provide liquidity conditional on negative returns.

Before testing the institutional accommodation mechanism, we first rule out the primary components of the retail clientele hypothesis:

Attention and Sentiment. One behavioral explanation is that under-performing stocks disproportionately draw retail attention or trigger contrarian sentiment. However, utilizing data from StockTwits to proxy for retail attention and sentiment (Figure 6), we find that neither variable sufficiently explains the magnitude or the nonlinear asymmetry of the observed retail flows. Both high- and low-attention stocks, and both high- and low-sentiment stocks, exhibit the same sharp left-tail buying pattern. The dip-buying asymmetry is therefore not primarily a consequence of differential visibility or emotional reactions to losses.

Firm Characteristics. Alternatively, the preference might be driven by mechanical changes in a stock's profile following a price drop. For instance, a declining stock automatically appears more "value-like" as its book-to-market ratio rises or its size rank falls. As shown in Figure 7, splitting the sample by book-to-market, market capitalization, gross profitability, and net income, we find that the asymmetric response to negative returns persists across all characteristic subgroups. The dip-buying pattern is not subsumed by shifts in observable firm fundamentals.

Retail Flow Persistence. If the retail clientele hypothesis holds, order flows should exhibit meaningful persistence: stocks favored by retail investors should continue to receive inflows over time, reflecting stable demand. To test this, we estimate the autoregressive persistence of retail flows at the stock level:

$$Flow_{i,t} = \alpha_i + \phi\, Flow_{i,t-1} + \varepsilon_{i,t} \tag{1}$$

As reported in Table 3, retail flows exhibit remarkably weak persistence. The estimated AR(1) coefficient ϕ is just 0.005 at the daily frequency and 0.133 at the monthly frequency, similar to the persistence typical of standard stock characteristics such as momentum or valuation ratios. The limited persistence implies that retail flows are largely episodic and reactive rather than reflecting the steady accumulation that would characterize stable stock-level clienteles. Retail investors do not appear to maintain sustained directional positions in particular securities over time.

This finding is fully consistent with the time-series evidence in Figure 5, which shows

persistent inflows into the lowest-momentum portfolio. That portfolio-level persistence simply reflects the fact that stocks with negative returns tend to receive inflows *because* they keep experiencing negative returns, not because retail investors form durable habitats in specific firms or characteristics. The two findings together clarify the nature of dip-buying: it is a response to price states, not a preference for particular securities.

Preferences and Next-Period Returns. Finally, if retail inflows were driven by behavioral biases that push prices away from fundamentals, we would expect subsequent returns to be negative following heavy retail buying, as clientele demand generates a negative risk premium (Hong and Kacperczyk, 2009; Pástor et al., 2021). Instead, as shown in Figure 8, next-period returns are positive for low-momentum stocks that receive retail inflows. This is inconsistent with retail investors overpaying for stocks they irrationally prefer; it is instead consistent with retail investors purchasing stocks that are temporarily depressed by selling pressure from other market participants, and earning a reversal premium as that pressure subsides.

The evidence against the clientele hypothesis points naturally toward an accommodation story. As shown in Figure 9, the contrast with non-retail investors is striking: when retail investors are net buyers following negative returns, total market flows and institutional proxies are net sellers. This near-mirror-image relationship suggests that retail buying is not an isolated demand shift but a direct counterpart to institutional selling, as the Institutional Accommodation hypothesis predicts. We explore this mechanism formally in the next section.

4 Mechanism: Accommodating Institutional Selling

4.1 Hypothesis: Institutional Selling Creates the Retail "Habitat"

Having ruled out pure clientele preferences and behavioral biases as the primary drivers of retail flows into distressed stocks, we propose an alternative mechanism: the Institutional Accommodation hypothesis. As shown in Figure 9, while retail investors "buy the dip," non-retail investors systematically "sell the dip."

Financial institutions frequently face strict mandates, redemptions, or Value-at-Risk (VaR) constraints that force them to liquidate positions following negative returns,

independent of the stock's fundamental value. Because retail investors are not subject to these same institutional constraints, they are uniquely positioned to step in and absorb this supply. Thus, the institutional imperative to sell creates a habitat where retail investors act as de facto liquidity providers, conditional on negative returns. The core prediction of this hypothesis is that retail buying should be strongest precisely when mandated institutional selling pressure is most intense.

4.2 Quasi-Natural Experiment: The Melvin Capital Short Squeeze

To cleanly test this mechanism, we require an exogenous shock to institutional selling pressure that is unrelated to the fundamental value of the stocks being sold. The January 2021 GameStop (GME) short squeeze and the subsequent near-collapse of Melvin Capital provides an ideal quasi-natural experiment (see Figure 10).

Melvin Capital held massive short positions in heavily shorted names. As retail coordinated a massive squeeze, Melvin faced unprecedented margin calls. To raise cash and meet these liquidity requirements, the fund was forced into a rapid fire sale of its existing long portfolio. This liquidation was entirely driven by margin requirements on their short book, making the sale of their long positions fundamentally exogenous.

We establish January 2021 as the event date ($t = 0$). The treatment group consists of long stocks that Melvin Capital was forced to liquidate (sold), while the control group comprises long stocks they retained (kept). We further partition the sample cross-sectionally into past Losers ($R_{t-1} < 0$) and past Winners ($R_{t-1} \geq 0$). If retail investors are accommodating institutional fire sales, we should observe a sharp increase in retail inflows specifically directed at the treated loser stocks.

Figure 11 shows a clean differential pattern: retail inflows surge specifically in the treated loser stocks—those that Melvin was simultaneously forced to liquidate and that had experienced negative prior returns—but not in the control stocks that Melvin retained, nor in the treated winner stocks. This targeted response confirms that retail accommodation is directed at forced institutional selling in down states, rather than at all losing stocks or all stocks associated with Melvin. Figure A.6 plots the corresponding cumulative returns: the treated loser stocks exhibit a modest subsequent recovery, consistent with the initial price decline having been partially driven by temporary

institutional selling pressure rather than deteriorating fundamentals, and with retail buyers earning a reversal as that pressure dissipates.

4.3 Institutional Selling Pressure: Mutual Fund Distress

Moving beyond episodic events, we evaluate broad institutional selling pressure using two distinct proxies for mutual fund distress: forced selling driven by fund outflows, and alpha pressure from under-performing funds.

Fund Outflows. Following Coval and Stafford (2007), we calculate mutual fund selling pressure driven by severe investor redemptions. For stock i at month t, the pressure induced by domestic equity funds j is defined as:

$$Pressure_{i,t} = \frac{\sum_j (Sell_{j,i,t} \mid flow_{j,t} < -5\%) - \sum_j (Buy_{j,i,t} \mid flow_{j,t} > 5\%)}{\sum_j Own_{j,i,t-1}}. \quad (2)$$

We regress both next-month returns and concurrent retail flows on this pressure measure, interacting it with an indicator for negative past returns ($\mathbb{I}[R_{i,t} < 0]$). As reported in Table 4, the interaction term is highly significant and positive ($\beta_3 = 0.0477$, t-stat $= 3.35$), indicating that distressed stocks facing severe institutional outflow pressure receive significantly more retail inflows.

Fund Under-performance (Alpha Pressure). As a secondary proxy, we utilize alpha pressure, which captures stocks held by poorly performing active equity funds that are likely to face future redemptions. To avoid mechanical correlation, we compute the fund's return excluding stock i's contribution:

$$ExStockRet_{j,i,t} = \frac{PortRet_{j,t} - (w_{j,i,t} \times Ret_{i,t})}{1 - w_{j,i,t}} \quad (3)$$

We aggregate rolling 24-month net-of-fee alphas, as well as expense ratios, to the stock level. Consistent with our hypothesis, Table 5 shows that negative past returns interacted with under-performance proxies predict significant retail inflows, reflecting state-contingent liquidity provision to struggling active managers, concentrated in stocks with negative past returns.

4.4 Passive Rebalancing and Institutional Constraints

A natural validation test is to contrast active fire sales with mechanical, passive index rebalancing. When stocks are added to or excluded from major indices such as the S&P 500 or the Russell 1000/2000, index funds face known, pre-announced, and predictable demand shocks. Unlike mutual fund outflow-driven sales, these reconstitutions carry no distress signal and do not reflect deteriorating fundamentals or forced liquidation under binding constraints.

The key prediction of the Institutional Accommodation hypothesis is as follows. Retail investors do not need to be sophisticated enough to "see through" flows and distinguish mispricing from mechanical rebalancing. Rather, the hypothesis predicts that retail accommodation arises specifically when the broader institutional sector is constrained and unable to absorb selling pressure. In the case of index reconstitutions, because the flows are predictable and non-distress-driven, active institutional investors face no constraints on providing liquidity, as there is no "missing buyer" problem. Accordingly, we find little to no asymmetric retail response to index-driven demand shocks. This null result validates the mechanism: retail absorbs institutional selling not because it can identify mispricing, but because it is the last available buyer when the broader institutional sector is simultaneously constrained.

Finally, Table 6 examines two institutional frictions that amplify the retail accommodation effect: realized volatility as a proxy for binding Value-at-Risk (VaR) constraints (VOL), and the Herfindahl-Hirschman index of institutional ownership concentration (HHI^{IO}) as a proxy for limited risk-sharing capacity. Both variables capture states in which the broader institutional sector is most uniformly unable to absorb selling pressure. In high-volatility environments, risk-managed institutions face tighter constraints and are more likely to reduce positions simultaneously, leaving fewer natural counterparties. When ownership is concentrated among a small number of institutional holders, correlated liquidity shocks have larger price effects with fewer buyers available to absorb them. Consistent with these predictions, the interaction term $\mathbb{I}[R_{i,t} < 0] \times X_{i,t}$ is positive and highly significant in the retail flow regression for both proxies: retail inflows are amplified precisely in the states where other institutions are most constrained. This

reinforces the view that retail accommodation is a function of institutional capacity, not retail sophistication.

5 Market Outcomes and Retail Profitability

Having established that retail investors actively accommodate institutional selling pressure, we now examine the market outcomes of this liquidity provision. Specifically, we investigate whether retail order flow predicts future returns, how prices and liquidity respond dynamically, and whether retail investors ultimately profit from "buying the dip."

5.1 Future Returns and Flow Intensity

We begin by testing the predictive power of retail flows for future returns in the cross-section. We double-sort stocks into 5×5 portfolios based on their past one-month return (distress) and their concurrent retail flow intensity.

As illustrated in Figure 12, we observe high next-month returns primarily concentrated in the most distressed stocks (Quintile 1). Interestingly, within this distressed quintile, the *magnitude* of the retail flow does not generate a monotonically increasing premium. The return premium appears to be fundamentally driven by the distressed state itself—the underlying institutional fire sale—rather than the intensity of the retail buying. Retail investors are correctly identifying and participating in these high-expected-return states, but their flow intensity alone does not drive the reversal.

This pattern carries an important implication for the nature of retail trading. The liquidity premium in distressed stocks is not created by retail demand; it pre-exists retail entry and reflects the institutional selling pressure that temporarily depresses prices. Retail inflows are a symptom of the same distress states that generate the premium. Consistent with this interpretation, adding controls for retail flow intensity does not materially reduce the return predictability of the momentum-quintile sort.

5.2 Dynamic Responses: Event-Time Specification

To understand the precise timing of market stabilization, we estimate dynamic responses using an event-time specification for horizons $h \in \{-5, \ldots, 10\}$:

$$Y_{i,t-1\to t+h} = \alpha_i + \beta_F \, Flow_{i,t} + \Gamma' X_{i,t-1} + \epsilon_{i,t+h} \tag{4}$$

where $h < 0$ captures placebo pre-trends and $h \geq 0$ measures the post-event effects. Standard errors are clustered by firm and date. We evaluate multiple outcomes, focusing on cumulative returns ($r_{i,t+h}$) and liquidity, proxied by bid-ask spreads ($(ask-bid)/close$).

Figure 13 plots the dynamic response of returns, volatility, and bid-ask spreads. Returns exhibit a sharp stabilization and reversal following the retail inflow event, consistent with the absorption of a temporary liquidity shock: the pre-event price decline is partially reversed over the subsequent months, with no evidence of a pre-trend in the placebo horizons ($h < 0$). Bid-ask spreads deteriorate sharply during the institutional sell-off—reflecting the elevated adverse-selection and inventory risk that accompanies forced selling—and subsequently normalize as retail investors supply liquidity and absorb the inventory imbalance. This dynamic pattern confirms that retail accommodation has measurable market-quality benefits: it shortens the period of price dislocation and restores normal transactions costs more quickly than would occur absent the retail inflow.

5.3 Retail Investor Profits and Capacity Constraints

Do retail investors actually profit from accommodating these institutional fire sales? To answer this, we decompose total mark-to-market (MTM) profits into two components: (1) Execution (the return from the cost basis to the month-end) and (2) Trend (the return from month-end to the next month).

When we calculate the *simple average* ROI of these trades, the results suggest highly profitable outcomes. As shown in Figure 14, the average retail trade in extreme loser stocks (Quintile 1) generates a significantly positive return. This suggests that the typical retail investor who buys the dip successfully captures the liquidity premium associated with the reversal.

However, a starkly different picture emerges when we calculate the *dollar-weighted* profit. As plotted in Figure 14, when aggregating total dollar profits over total dollars deployed, the net profit becomes deeply negative.

This dichotomy reveals an important distinction between return expectations and stock-selection skill. Retail investors appear to hold a rational expectation that momentum reversal occurs on average, and the positive simple-average ROI confirms this expectation is correct as a statistical matter. However, retail investors do not exhibit

meaningful stock-level selection or timing skill: they cannot identify which particular distressed stocks will reverse, nor do they calibrate the magnitude of institutional pressure in individual cases. Rather than picking the right stocks, they buy broadly across the distressed cross-section and earn the average reversal premium.

The negative dollar-weighted return reflects what we call "catching the falling knife." Because retail investors cannot distinguish routine institutional selling from catastrophic fire sales at the individual stock level, the largest capital deployments tend to coincide with the deepest distress events, precisely the episodes where institutional selling is too large and too persistent for retail capital to absorb profitably. In these cases, retail investors enter early and aggressively before the institutional selling has run its course, and the subsequent price decline overwhelms the eventual reversal gain.

This finding does not imply that retail buyers are uninformed about the direction of mean reversion: the positive simple-average ROI confirms they are correct on average. It instead reveals a capacity constraint: retail capital is sufficient to earn the liquidity premium in routine distress events, but insufficient to stabilize prices during the largest fire sales, where continuing institutional outflows ultimately reverse any early retail gains.

6 Conclusion

This paper documents a robust and structural feature of retail trading: retail investors systematically increase net purchases following large price declines, and this response accelerates nonlinearly in the left tail of the return distribution. Using a novel dataset of actual retail transactions from 2017 to 2024, we show that this "buying the dip" pattern is stable across market regimes. It predates the pandemic, persists through the 2020–2021 meme-stock surge, and holds across return horizons, flow denominators, and individual stock examples. The behavior is one-sided: positive returns generate little symmetric selling, distinguishing dip-buying from standard contrarian rebalancing.

We rule out a retail clientele explanation. Retail flows exhibit negligible stock-level persistence, are not subsumed by firm characteristics or attention and sentiment proxies, and are followed by positive rather than negative subsequent returns. Instead, the evidence strongly supports an institutional accommodation mechanism. When institutional investors face binding constraints, such as mutual fund redemptions, fund

under-performance pressure, elevated VaR, or concentrated ownership, they are forced to sell. In these same states, retail investors step in as the marginal buyers. This is not because retail investors can identify mispricing or "see through" institutional flows; rather, retail fills the gap precisely when other institutional investors are simultaneously constrained and unable to provide liquidity. The contrast with passive index rebalancing is instructive: when mechanical index-driven selling occurs, active institutions face no constraints and absorb the supply themselves, leaving no accommodation role for retail.

Examining market outcomes, we find that retail inflows into distressed stocks predict positive subsequent returns and are associated with faster normalization of bid-ask spreads following institutional sell-offs. The return premium, however, is generated by the distressed state itself, i.e., by the institutional selling pressure that temporarily depresses prices, and not by the intensity of retail buying. In this sense, retail investors in distressed states act as if they are providing liquidity to the market, though this role is state-contingent rather than a general feature of retail trading.

The profitability analysis, however, offers a cautionary perspective. While the simple-average return on retail dip-buying trades is positive—confirming a rational expectation of mean reversion—the dollar-weighted return is negative. Retail investors deploy the most capital during the deepest distress events, precisely where institutional selling pressure is too large to be profitably absorbed. Rather than reflecting stock-picking skill, retail accommodation is a diffuse bet on mean reversion that earns the liquidity premium in routine episodes but loses capital in the largest fire sales. More broadly, the state-contingent nature of retail accommodation raises an open question about its welfare implications: because retail demand is concentrated in distress states, it may introduce additional fragility rather than unconditionally stabilize prices.

Several questions remain open for future research. Most notably, the analysis focuses on the entry side of retail trades. Whether retail investors also exit at the right time, i.e., capturing the full reversal premium rather than holding through subsequent price movements, is an important open question that we leave for future work.

References

Barardehi, Y. H., Z. Da, P. Dixon, and J. L. Wang (2024, November). You can only lend what you own: Inferring daily institutional trading from lendable equity inventory. Working paper, SSRN. Date written: November 15, 2024; posted November 25, 2024; last revised October 30, 2025.

Barber, B. M., X. Huang, P. Jorion, T. Odean, and C. Schwarz (2024). A (sub)penny for your thoughts: Can analyzing order submissions improve our understanding of retail investor behavior? *The Journal of Finance 79*(4), 2311–2356.

Barber, B. M., X. Huang, T. Odean, and C. Schwarz (2022). Attention-induced trading and returns: Evidence from robinhood users. *The Journal of Finance 77*(6), 3141–3190.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies 21*(2), 785–818.

Barrot, J.-N., R. Kaniel, and D. Sraer (2016). Are retail traders compensated for providing liquidity? *Journal of Financial Economics 120*(1), 146–168.

Battalio, R. H., R. H. Jennings, M. Saglam, and J. Wu (2023). Difficulties in obtaining a representative sample of retail trades from public data sources. *Available at SSRN 4579159*.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang (2021). Tracking retail investor activity. *The Journal of Finance 76*(5), 2249–2300.

Coval, J. and E. Stafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*(2), 479–512.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Noise trader risk in financial markets. *Journal of Political Economy 98*(4), 703–738.

Dorn, D. and G. Huberman (2010). Preferred risk habitat of individual investors. *Journal of Financial Economics 97*(1), 155–173.

Hong, H. and M. Kacperczyk (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics 93*(1), 15–36.

Hvidkjaer, S. (2008a). Small trades and the cross-section of stock returns. *Review of Financial Studies 21*(3), 1123–1151.

Hvidkjaer, S. (2008b). Small trades and the cross-section of stock returns. *The Review of Financial Studies 21*(3), 1123–1151.

J.P. Morgan Asset Management (2025). Who is buying U.S. equities? J.P. Morgan Asset Management Market Insights. Accessed March 2026.

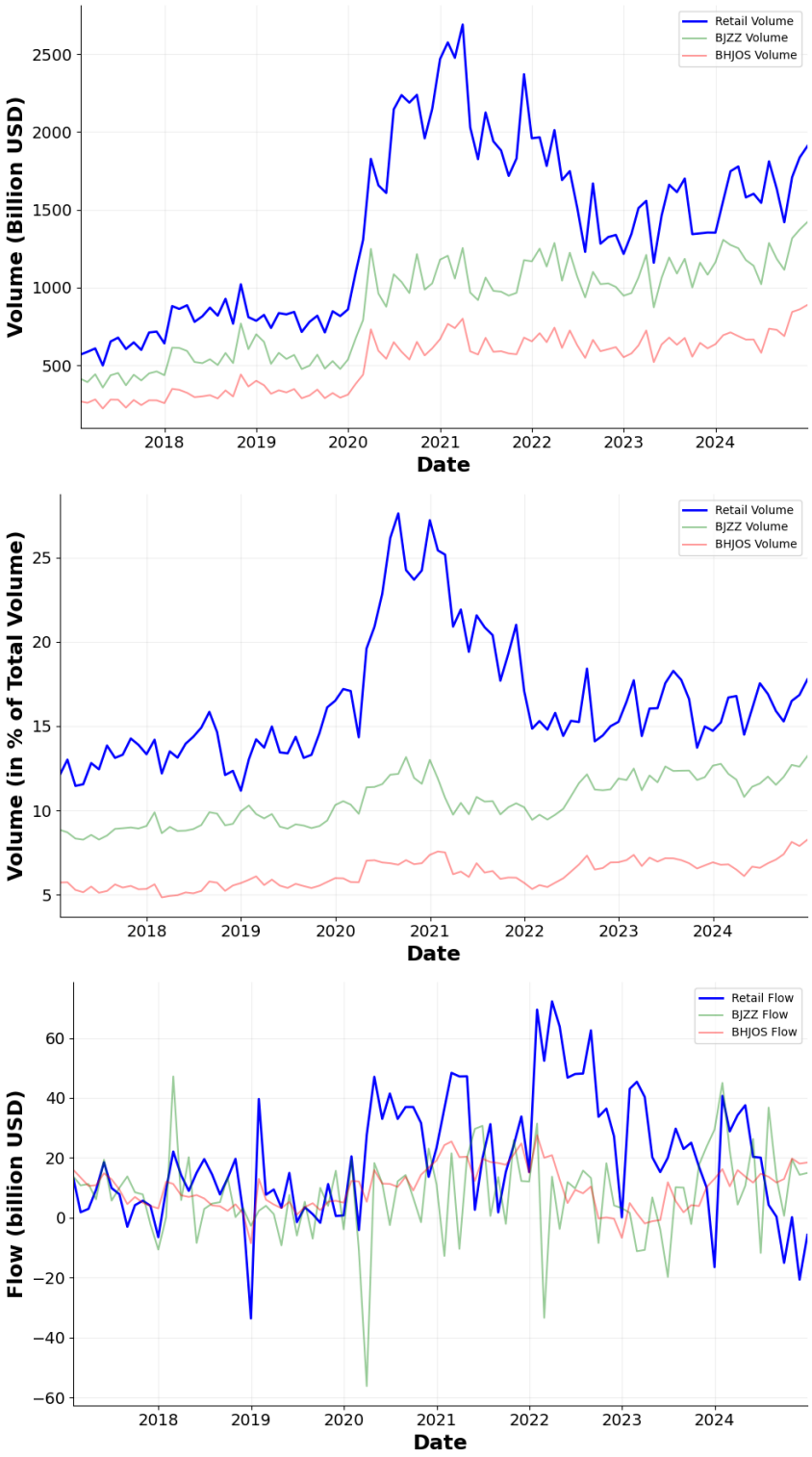
Kaniel, R., G. Saar, and S. Titman (2008). Individual investor trading and stock returns. *Journal of Finance 63*(1), 273–310.

Laarits, T. and M. Sammon (2025). The retail habitat. *Journal of Financial Economics 172*, 104144.

Luo, B., E. Ravina, M. Sammon, and L. M. Viceira (2025). Retail investors' contrarian behavior around news, attention, and the momentum effect. Working Paper 34086, National Bureau of Economic Research.

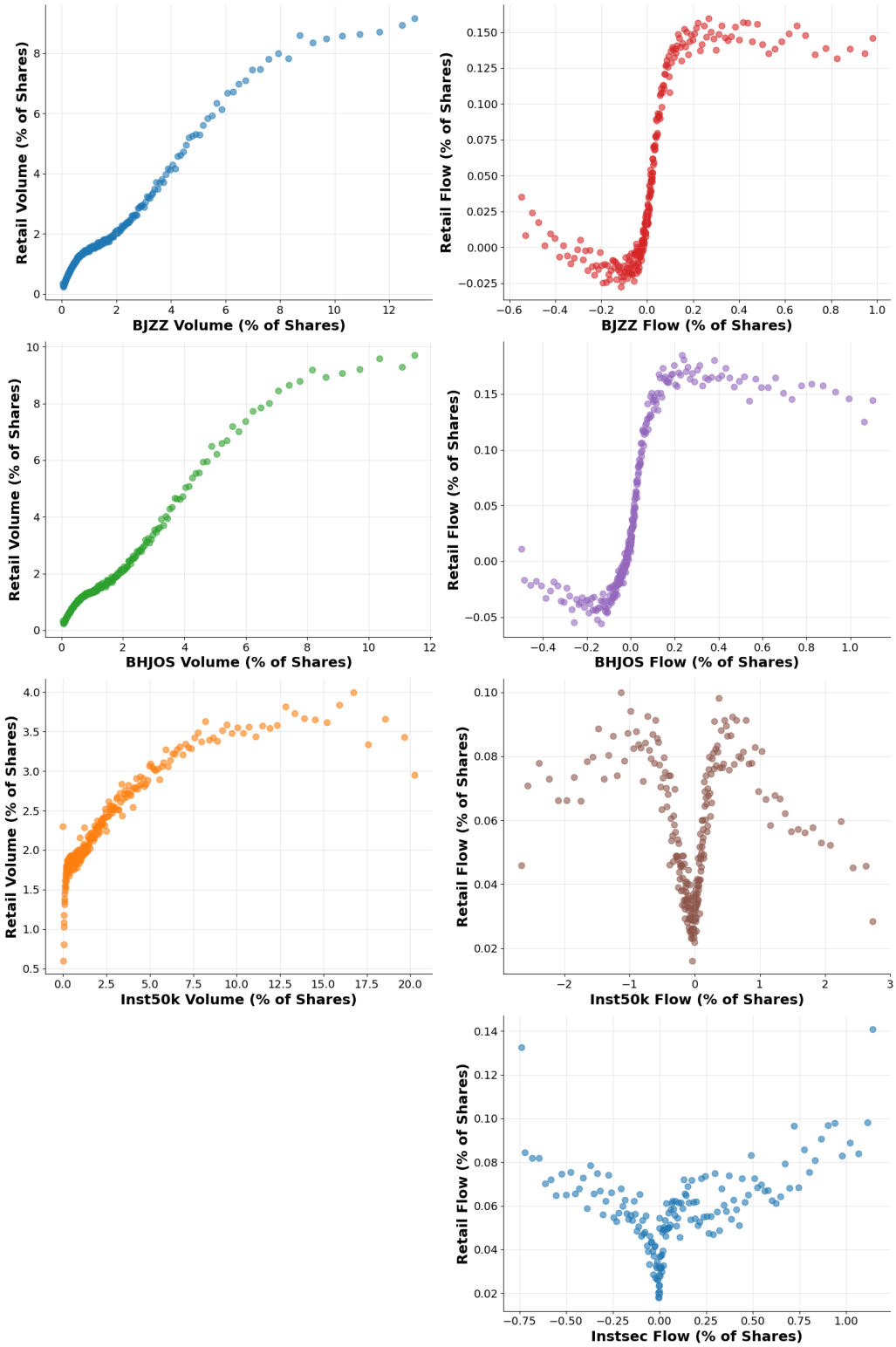
Pástor, Ľ., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. *Journal of Financial Economics 142*(2), 550–571.

Figure 1. Aggregate Retail Volumes and Flows



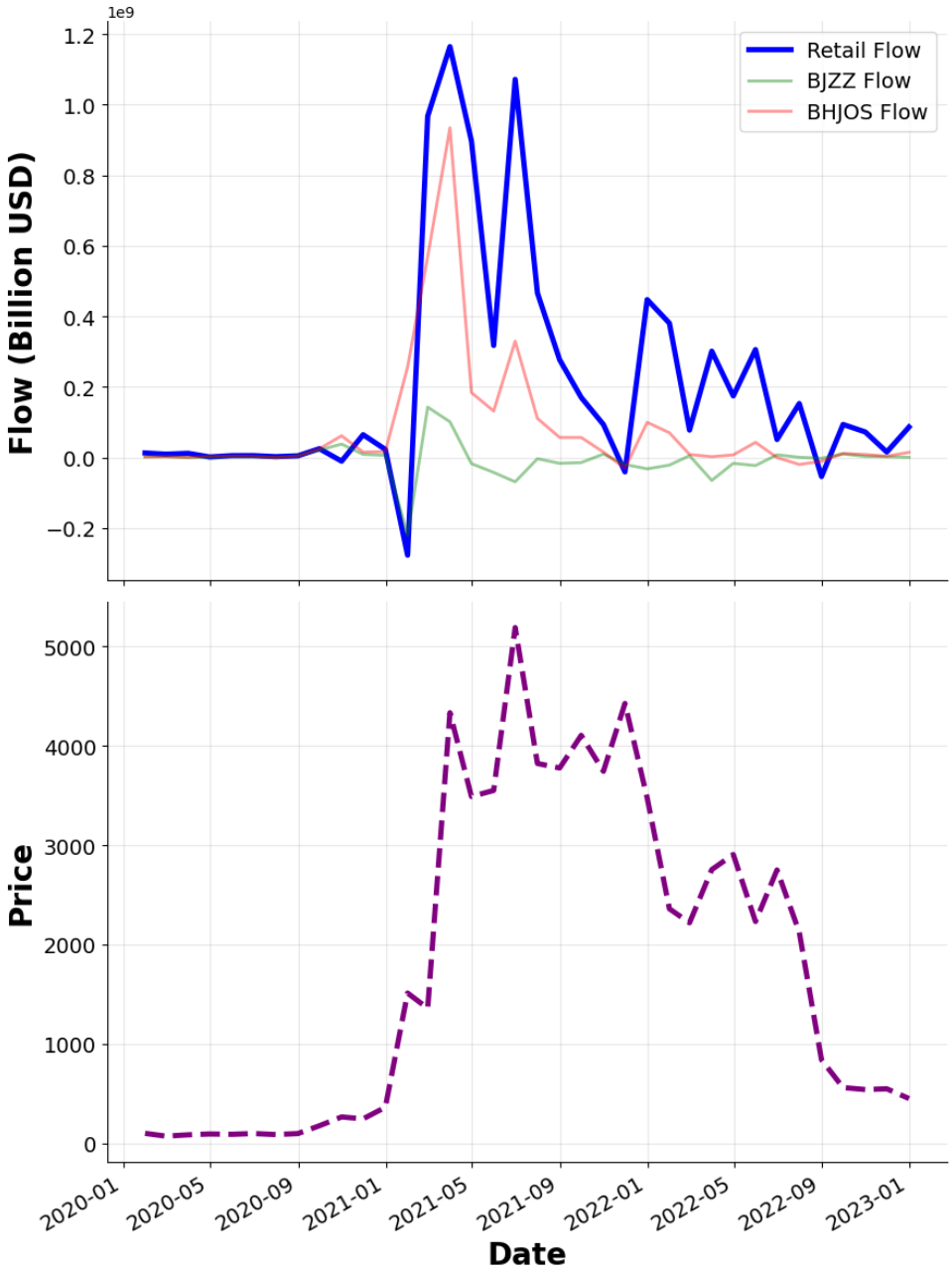
Notes: This figure displays aggregate monthly trading activity from 2017 to 2024. The top panel plots total retail dollar volume (in billions USD). The middle panel shows retail volume as a percentage of total market volume. The bottom panel plots aggregate retail, BJZZ, and BHJOS net order flows over time (in billions USD).

Figure 2. Cross-Validation of Retail Volumes and Flows: Pairwise Comparison



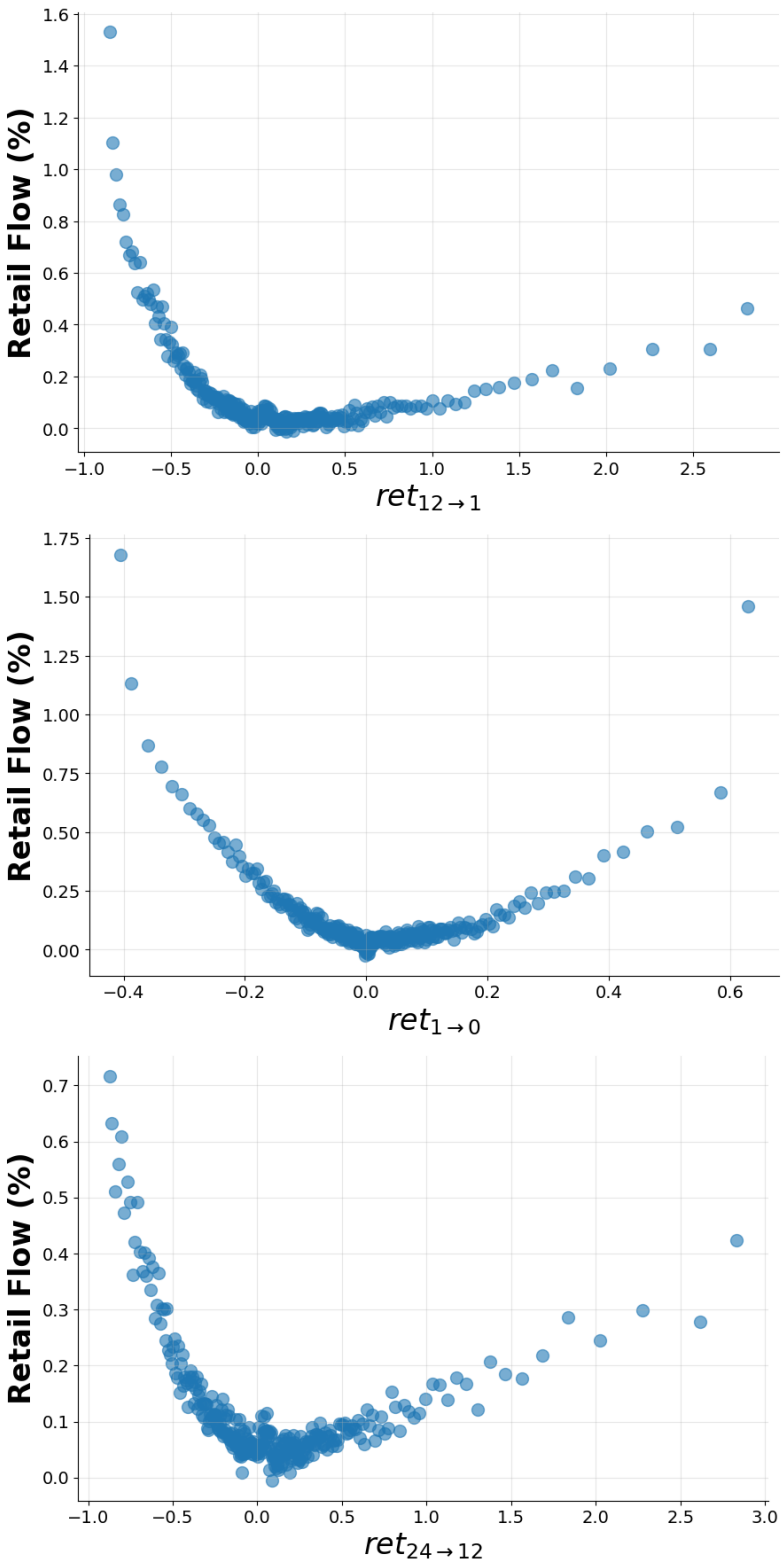
Notes: This figure cross-validates the S&P Retail Flow measure against four alternative flow proxies: BJZZ (row 1), BHJOS (row 2), Inst50k (row 3), and Instsec (row 4). The left column plots each proxy's volume against retail volume; the right column plots each proxy's flow against retail flow. All variables are expressed as a percentage of shares outstanding.

Figure 3. Event Study: Order Flows during the GameStop (GME) Short Squeeze



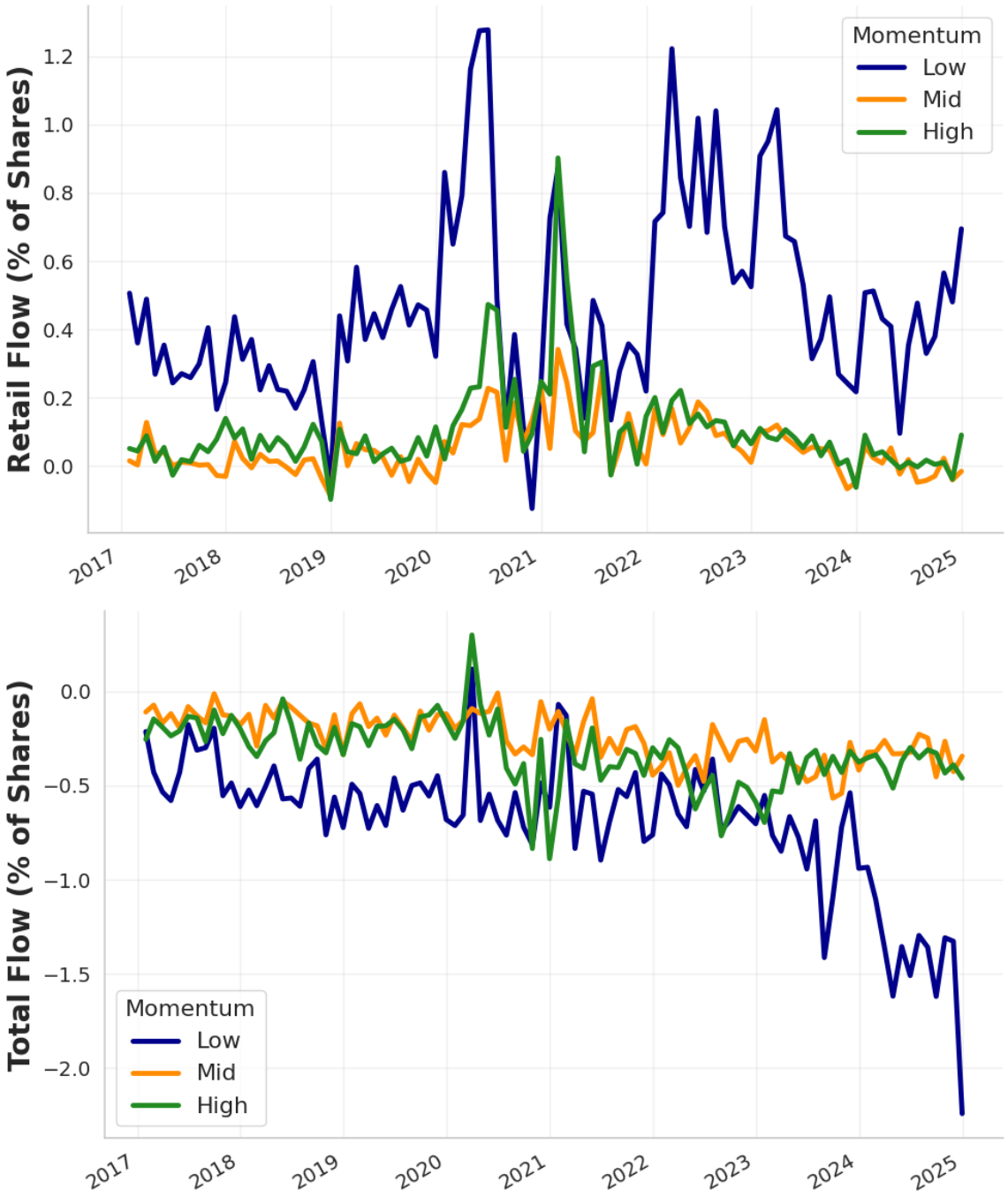
Notes: This figure compares the order imbalances of actual retail flow, BJZZ inferred flows, and BHJOS modified flows during the GameStop meme-stock episode from 2020 to 2022. The top panel plots the retail flows in terms of total notional values (in billion USD). The bottom panel plots the stock price of GameStop during the period.

Figure 4. Buying the Dip in the Retail Flow



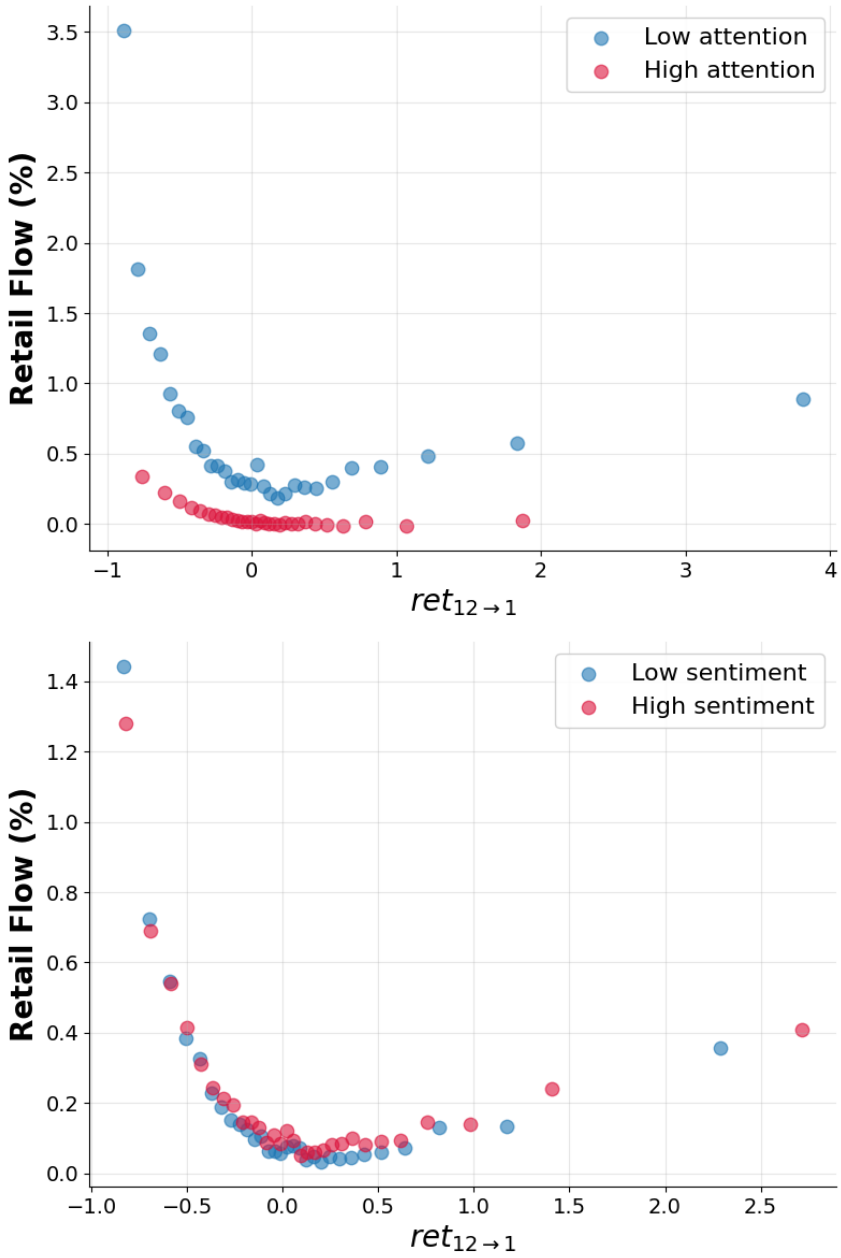
Notes: This figure plots binscatter estimates of retail order flow (as % of shares outstanding) against lagged stock returns across three horizons: the 12-to-1 month return ($ret_{12\to1}$, top panel), the most recent one-month return ($ret_{1\to0}$, middle), and the 24-to-12 month return ($ret_{24\to12}$, bottom).

Figure 5. Buying the Dip Over Time



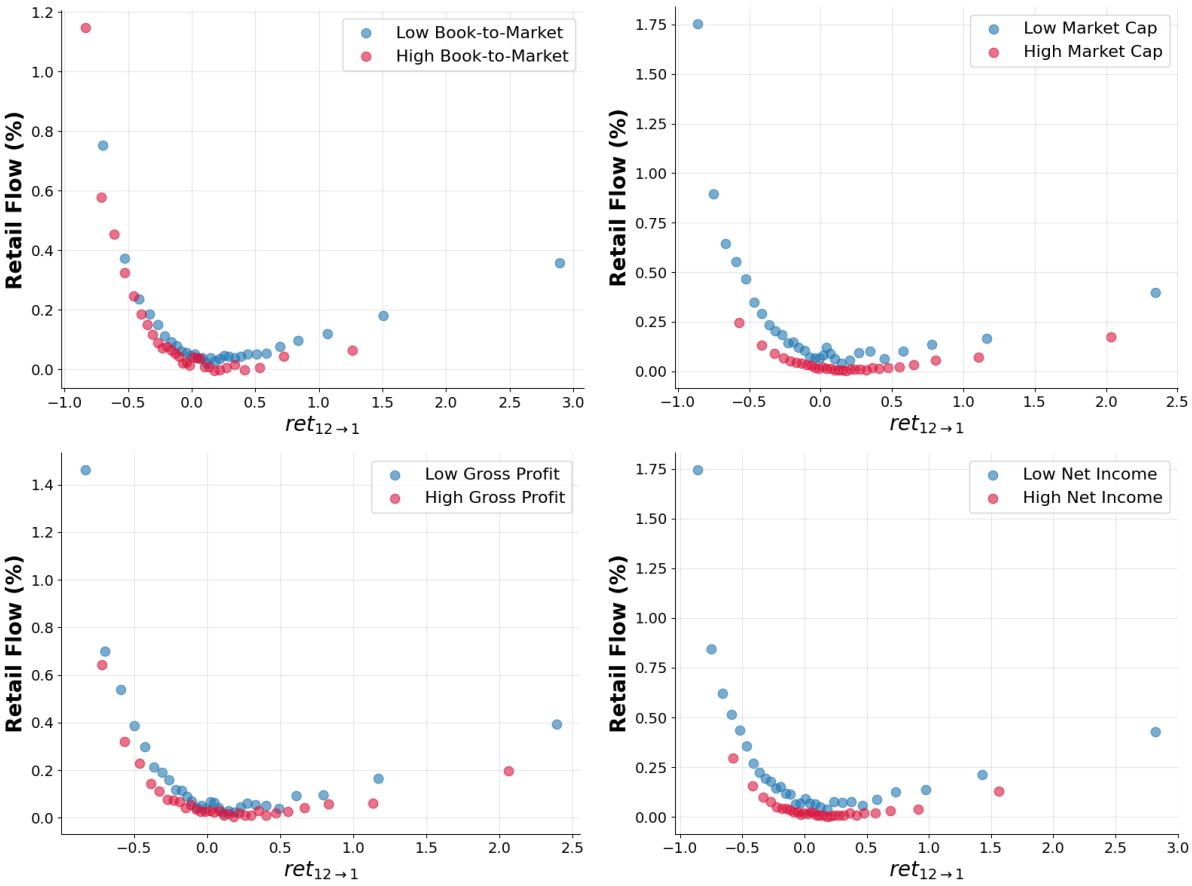
Notes: This figure plots the average monthly retail flow (top) and total flow (bottom) by past-return momentum tercile (Low, Mid, High) from 2017 to 2024.

Figure 6. Retail Clientele: Attention and Sentiment



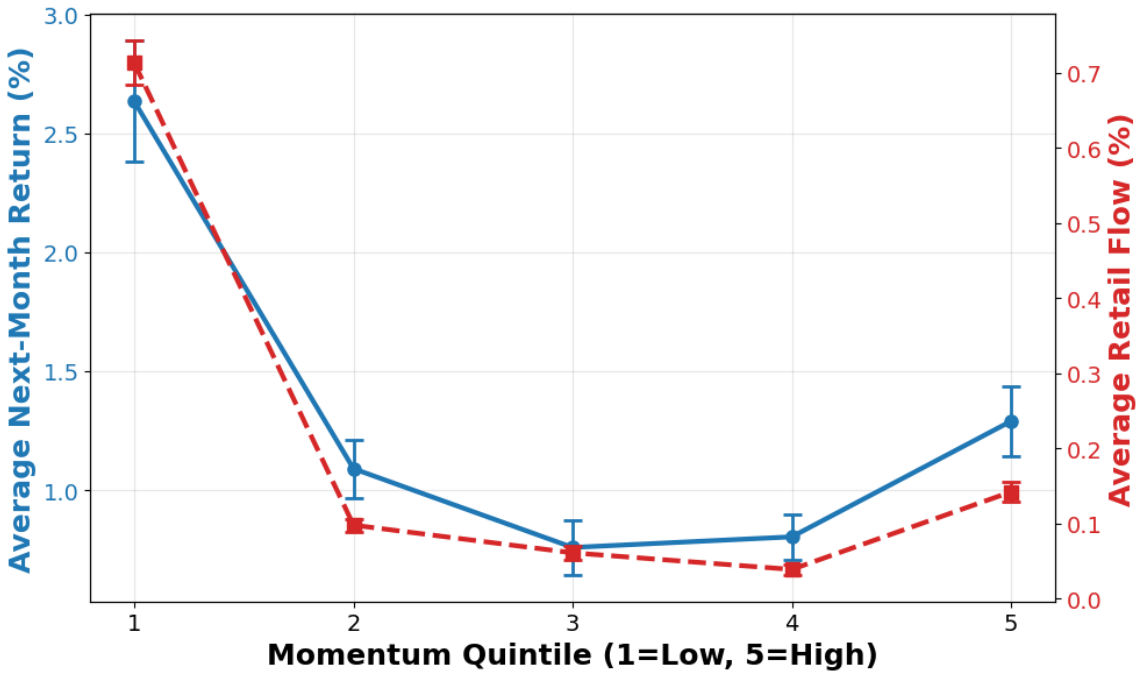
Notes: This figure tests the Retail Clientele hypothesis using StockTwits data. Each panel plots a binscatter of retail flow against $ret_{12\to1}$, split by stocks in the top versus bottom half of StockTwits attention (top) and sentiment (bottom).

Figure 7. Retail Clientele: Firm Characteristics



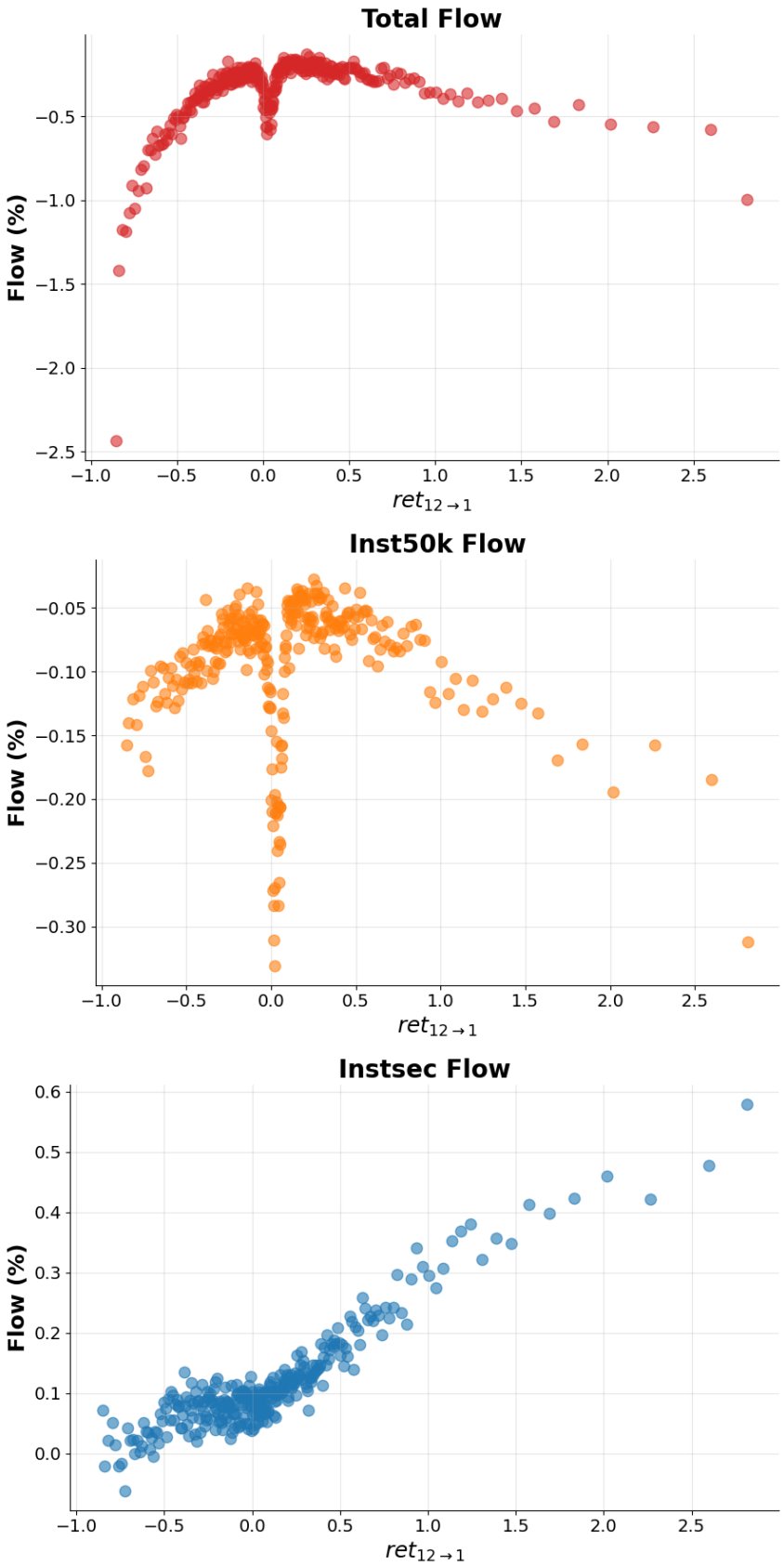
Notes: This figure evaluates whether the "buy the dip" pattern is explained by firm-level characteristics. Each panel shows a binscatter of retail flow against $ret_{12\to1}$, split by stocks in the top versus bottom half of: book-to-market (top left), market capitalization (top right), gross profitability (bottom left), and net income (bottom right).

Figure 8. Buying the Dip and Next-Month Returns by Momentum Quintile



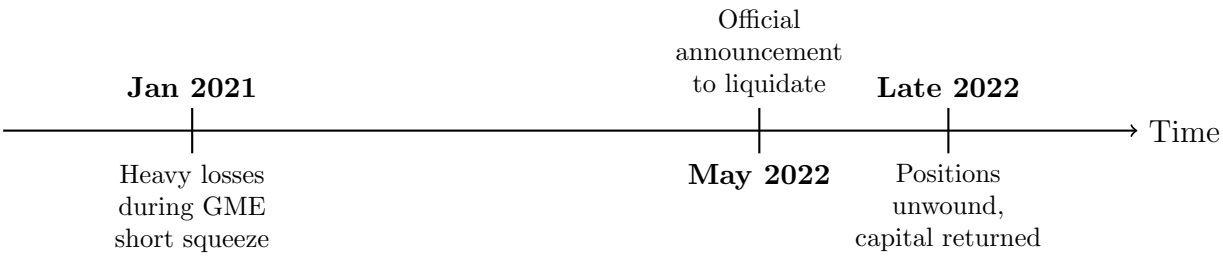
Notes: This figure plots average next-month returns (left axis, blue) and average retail flow (right axis, red) across five momentum quintiles (1 = lowest past return, 5 = highest). Error bars denote 95% confidence intervals.

Figure 9. Selling the Dip from Other Investors



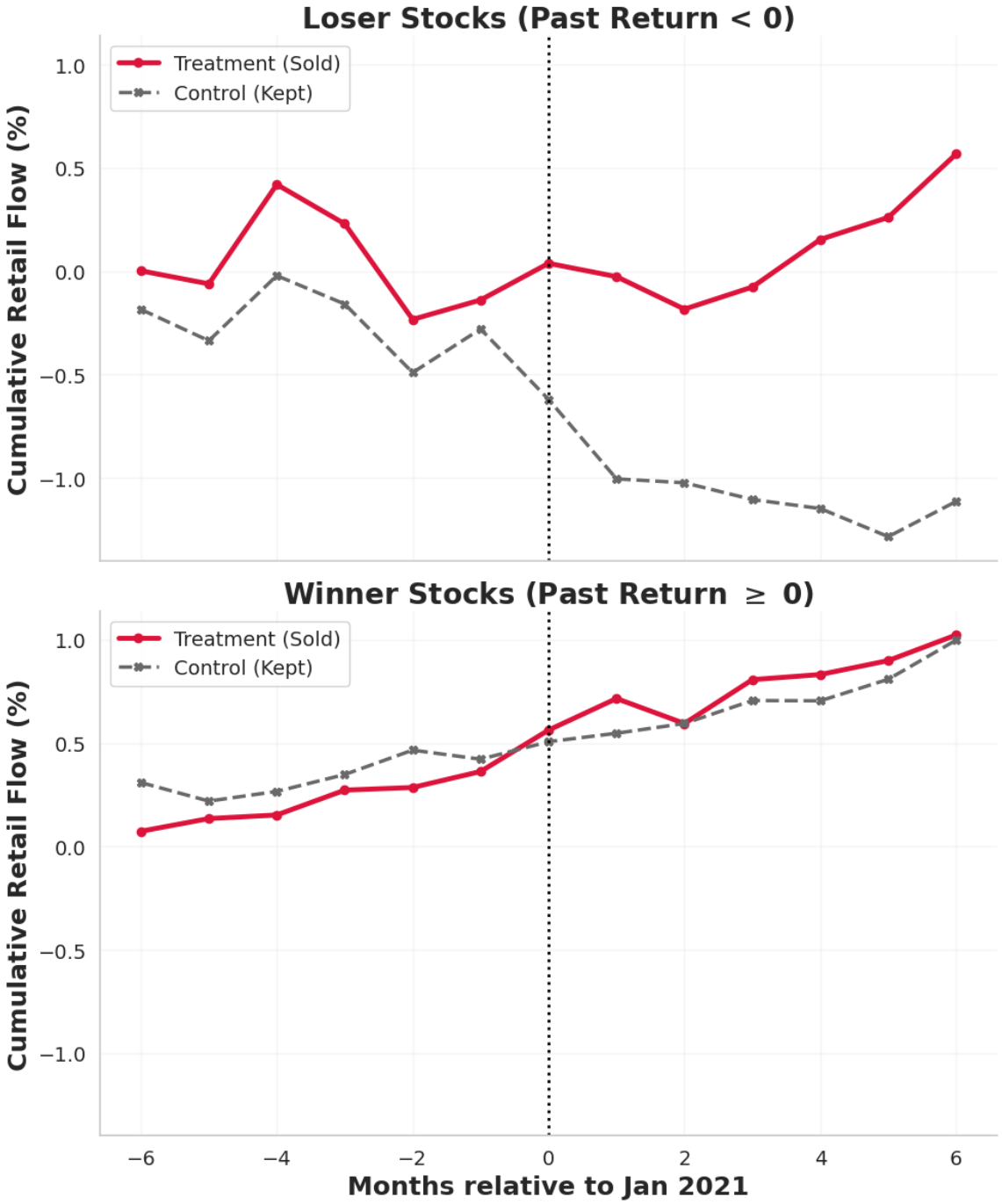
Notes: This figure plots binscatter estimates of net order flow against lagged $ret_{12\to1}$ for three non-retail measures: total market flow (top), large-trade institutional flow Inst50k (middle), and securities-lending-based institutional flow Instsec (bottom).

Figure 10. Timeline of Melvin Capital's Liquidation



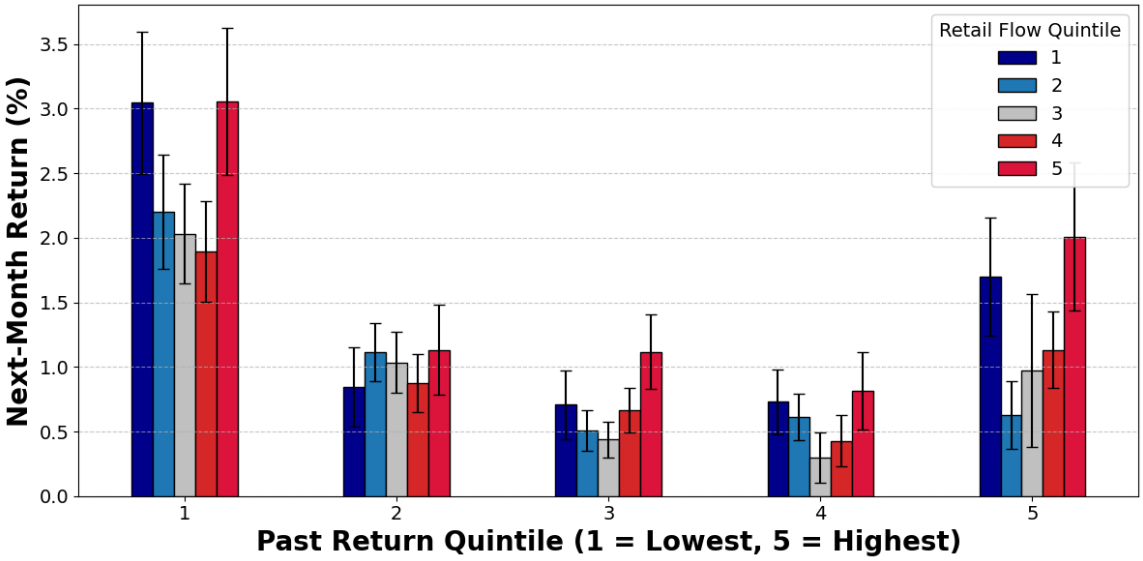
Notes: This diagram outlines the timeline and mechanics of the Melvin Capital short squeeze event in January 2021.

Figure 11. Event Study of Melvin Capital: Cumulative Retail Flows



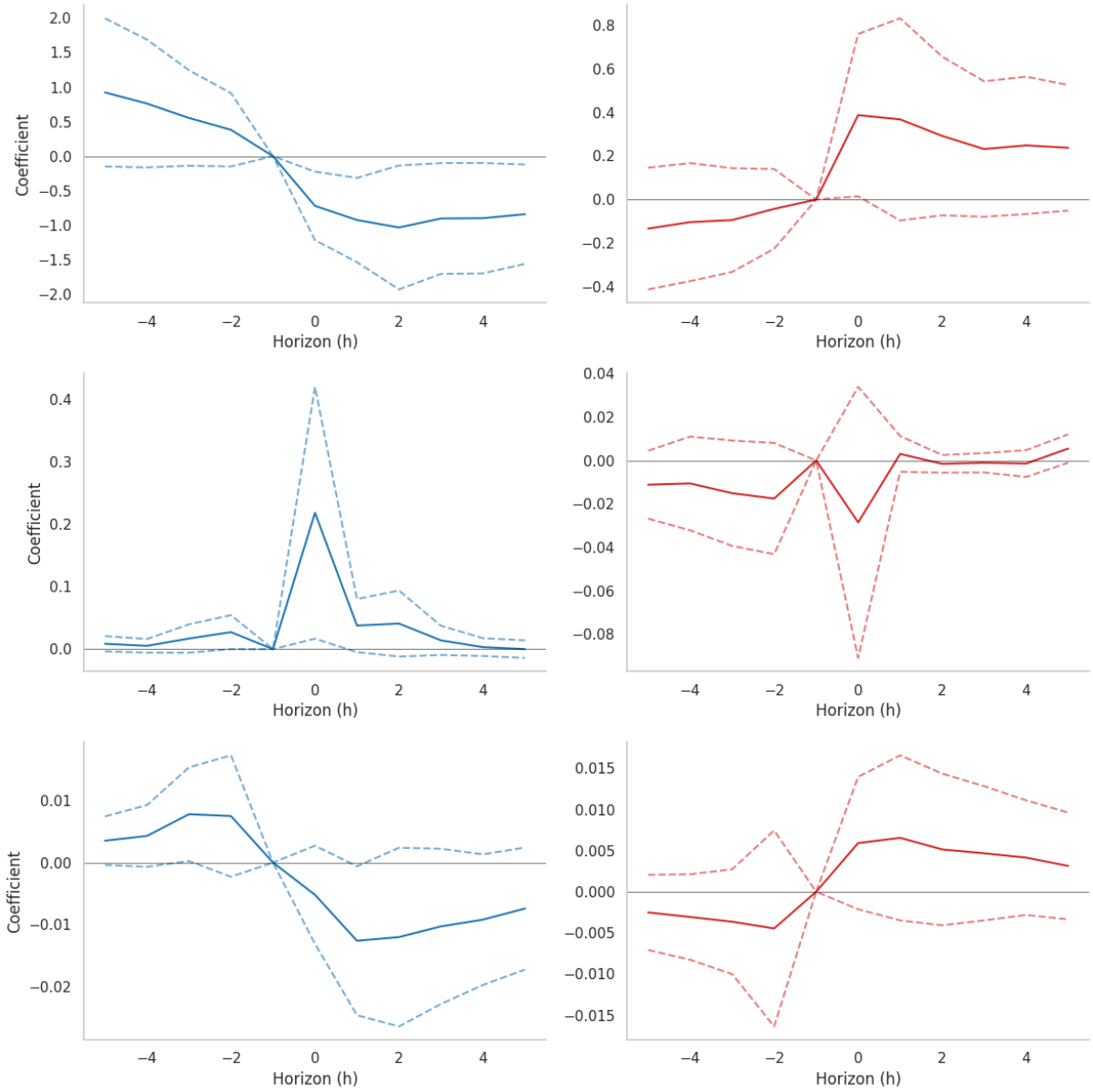
Notes: This figure plots the cumulative retail inflows into the treatment (liquidated) versus control (retained) long stocks held by Melvin Capital.

Figure 12. Retail Flow and Next-Month Returns by Flow Quintile



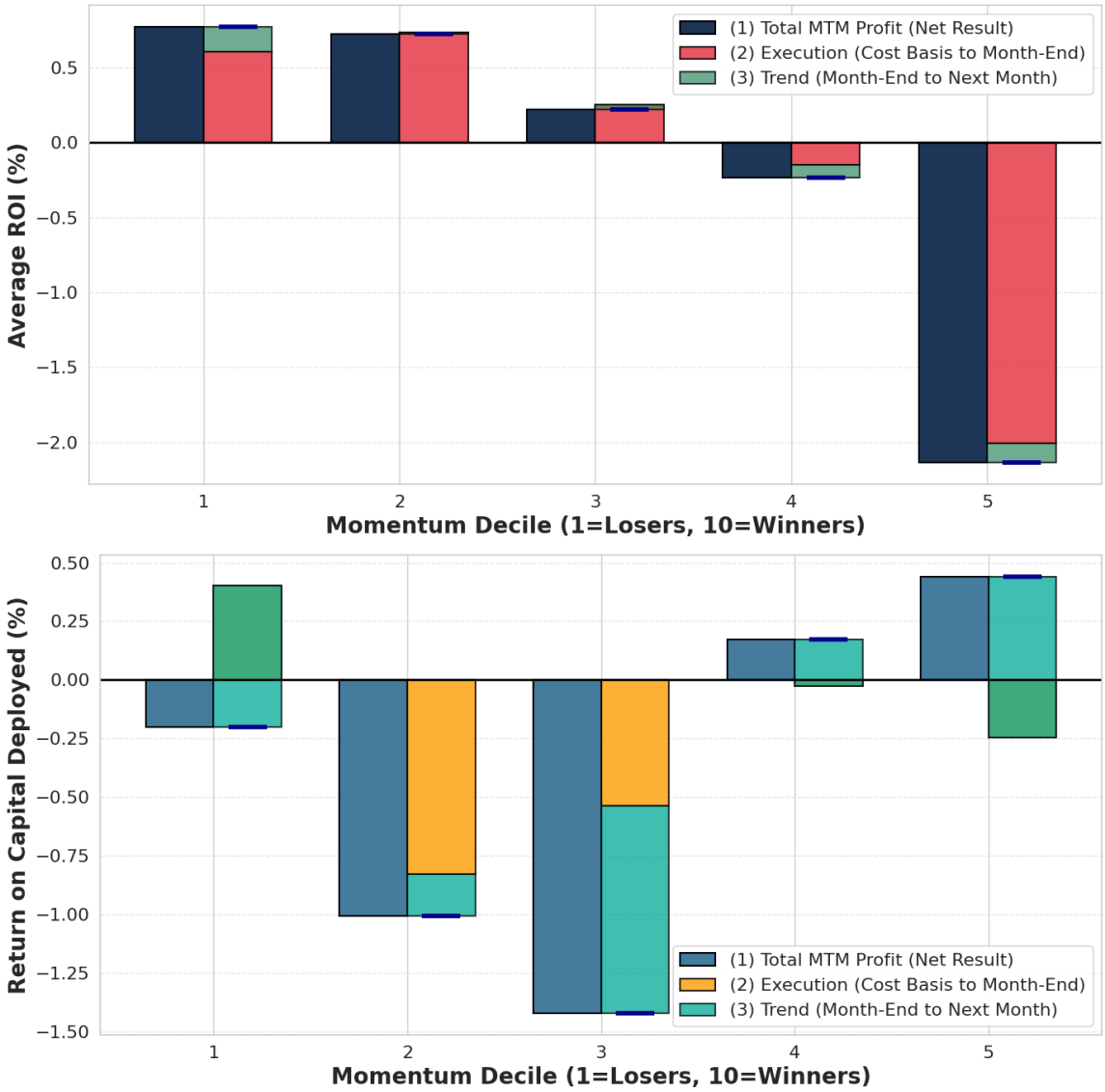
Notes: This figure presents next-month returns for 5×5 portfolios sorted independently by past one-month returns and concurrent retail flow intensity.

Figure 13. Dynamic Responses: Returns, Volatility, and Bid-Ask Spreads



Notes: This figure plots dynamic panel-regression coefficients for horizons $h \in \{-5, \ldots, 5\}$, estimating the response of monthly log returns (top row), return volatility (middle row), and bid-ask spreads (bottom row) to retail flow (left column) and total flow (right column). Shaded bands are 95% confidence intervals. Negative horizons serve as placebo pre-trend tests. Standard errors are clustered by firm.

Figure 14. Profitability of Retail Investors



Notes: This figure plots retail profitability by momentum decile (1 = losers, 10 = winners). The top panel shows the simple average return on investment (ROI) per dollar deployed within each decile. The bottom panel shows the dollar-weighted aggregate return on capital deployed, computed as the ratio of total mark-to-market profit to total capital deployed.

Table 1. Aggregate Time-Series Correlations of Order Flows

	Retail	BJZZ	BHJOS	Total	Inst50k	Instsec
Retail	1.00					
BJZZ	-0.01	1.00				
BHJOS	0.33	0.29	1.00			
Total	-0.17	-0.15	0.19	1.00		
Inst50k	-0.23	-0.20	0.03	0.85	1.00	
Instsec	-0.07	0.31	0.20	0.06	-0.22	1.00

Notes: This table reports Pearson correlations between aggregate monthly time-series of six flow measures over 2017–2024. *Retail* is the S&P Retail net order flow. *BJZZ* is the sub-penny algorithm measure of Boehmer et al. (2021). *BHJOS* is the modified measure of Barber et al. (2024). *Total* is total market net order flow. *Inst50k* proxies institutional flow using trades above $50,000 (Hvidkjaer, 2008a). *Instsec* proxies institutional flow from securities lending supply (Barardehi et al., 2024). All flows are expressed as a percentage of shares outstanding and aggregated to the market level.

Table 2. Cross-Sectional Correlations of Order Flows

	Retail	BJZZ	BHJOS	Total	Inst50k	Instsec
Retail	1.00					
BJZZ	0.19	1.00				
BHJOS	0.34	0.67	1.00			
Total	-0.32	0.11	0.11	1.00		
Inst50k	-0.05	-0.05	-0.02	0.61	1.00	
Instsec	0.11	0.14	-0.02	-0.26	-0.25	1.00

Notes: This table reports average cross-sectional Pearson correlations between monthly stock-level flow measures over 2017–2024. Variable definitions follow Table 1. Cross-sectional correlations are computed by averaging the flows across all stocks in the sample.

Table 3. Retail Flow Persistence

	Daily		**Monthly**	
Model	ϕ	$t(\phi)$	ϕ	$t(\phi)$
Retail Flow	0.005	0.05	0.133	6.88
Total Flow	0.010	0.57	0.117	3.93

Notes: This table reports AR(1) coefficient estimates ϕ from the stock-level regression $Flow_{i,t} = \alpha_i + \phi\, Flow_{i,t-1} + \varepsilon_{i,t}$, estimated separately at daily and monthly frequencies. *Retail Flow* is the retail net order flow; *Total Flow* is total market net order flow. Both are scaled by shares outstanding. Stock fixed effects α_i are included. t-statistics in parentheses; standard errors clustered by stock.

Table 4. Institutional Selling Pressure: Fund Outflows

	Base		**Interact**	
	$R_{i,t+1}$	$RetailFlow_{i,t}$	$R_{i,t+1}$	$RetailFlow_{i,t}$
$Pressure_{i,t}$	0.0085	0.0174**	-0.0335**	0.0035
	(0.63)	(2.57)	(-2.33)	(0.57)
$\mathbb{I}[R_{i,t} < 0]$			0.0084***	0.0033***
			(11.84)	(10.17)
$\mathbb{I}[R_{i,t} < 0] \times Pressure_{i,t}$			0.1270***	**0.0477*** **
			(5.38)	**(3.35)**
Stock FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	330,779	378,456	330,779	378,456

Notes: This table reports panel regressions of next-month stock return $R_{i,t+1}$ and concurrent retail net order flow $RetailFlow_{i,t}$ on mutual fund flow-induced selling pressure. $Pressure_{i,t}$ is the stock-level measure of Coval and Stafford (2007): net selling by funds experiencing investor outflows exceeding 5% of assets under management, minus net buying by funds with inflows above 5%, scaled by lagged total institutional ownership. $\mathbb{I}[R_{i,t} < 0]$ is an indicator equal to one when the stock's past one-month return is negative. The Base columns include $Pressure_{i,t}$ alone; the Interact columns add $\mathbb{I}[R_{i,t} < 0]$ and its interaction with $Pressure_{i,t}$. All specifications include stock and time fixed effects. t-statistics in parentheses; standard errors clustered by stock. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 5. Institutional Selling Pressure: Fund Under-performance

	Fund Return		**Rolling Alpha**		**Expense Ratio**	
	R_{t+1}	$Flow_t$	R_{t+1}	$Flow_t$	R_{t+1}	$Flow_t$
$Alpha_{i,t}$	0.0116	-0.12	0.0039	-0.21***	-1.4835	429.00***
	(0.95)	(-1.58)	(0.72)	(-2.74)	(-0.62)	(3.07)
$\mathbb{I}[R_{i,t} < 0]$	0.0081***	0.22***	0.0073***	0.19***	0.0063***	0.59***
	(11.11)	(5.21)	(10.40)	(5.64)	(2.75)	(5.13)
$\mathbb{I}[R_{i,t} < 0] \times Alpha_{i,t}$	-0.0353***	**-2.00*****	0.1274***	**-4.00*****	0.8409	**-584.00*****
	(-2.83)	**(-3.79)**	(5.30)	**(-2.97)**	(0.27)	**(-3.55)**
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	345,410	345,410	327,903	327,903	328,325	328,325

Notes: This table reports panel regressions of next-month stock return R_{t+1} and concurrent retail net order flow $Flow_t$ on fund under-performance pressure interacted with a negative-return indicator. The three column pairs use fund raw return, rolling 24-month net-of-fee alpha, and expense ratio as alternative under-performance proxies. $Alpha_{i,t}$ denotes the relevant proxy aggregated to the stock level; to avoid mechanical correlation, each fund's return is computed excluding stock i's own contribution before aggregation. $\mathbb{I}[R_{i,t} < 0]$ is an indicator equal to one when past return is negative. All specifications include stock and time fixed effects. t-statistics in parentheses; standard errors clustered by stock. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

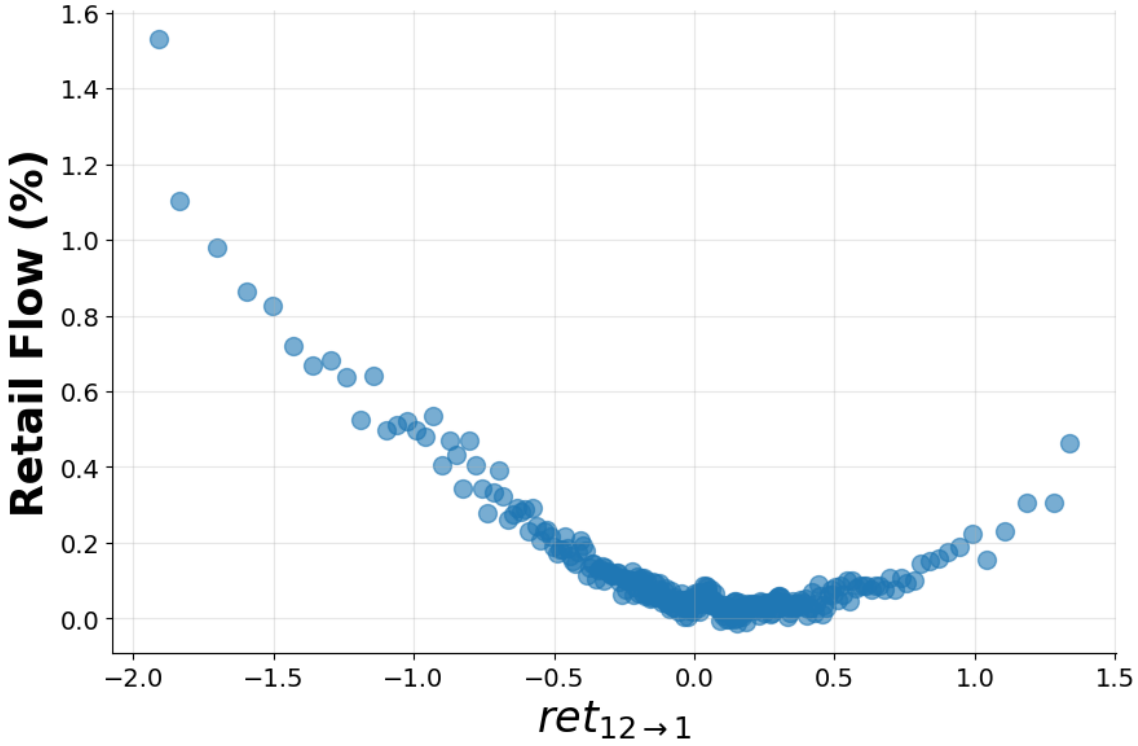
Table 6. Institutional Frictions: VaR and Ownership Concentration

	VOL		HHI^{IO}	
	$R_{i,t+1}$	$RetailFlow_{i,t}$	$R_{i,t+1}$	$RetailFlow_{i,t}$
$\mathbb{I}[R_{i,t} < 0] \times X_{i,t}$	0.0749***	0.0004***	0.0091***	0.0002***
	(3.27)	(2.87)	(2.76)	(5.25)
$X_{i,t}$	-0.0476***	0.0002***	-0.0079***	0.0002***
	(-3.37)	(2.93)	(-2.69)	(7.60)
$\mathbb{I}[R_{i,t} < 0]$	-0.0023***	0.0001***	-0.0033	0.0000*
	(-3.55)	(9.55)	(-0.92)	(1.99)
Time FE	Yes	Yes	Yes	Yes
Observations	301,290	343,694	408,322	475,185

Notes: This table reports panel regressions of next-month stock return $R_{i,t+1}$ and concurrent retail net order flow $RetailFlow_{i,t}$ on institutional friction proxies and their interaction with a negative-return indicator. $X_{i,t}$ denotes either realized return volatility ($VOL_{i,t}$, columns 1–2), computed as the standard deviation of daily returns over the prior month, or the Herfindahl-Hirschman index of institutional ownership concentration ($HHI^{IO}_{i,t}$, columns 3–4), constructed from quarterly 13F filings. $\mathbb{I}[R_{i,t} < 0]$ is an indicator equal to one when past return is negative. All specifications include time fixed effects. t-statistics in parentheses; standard errors clustered by stock. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

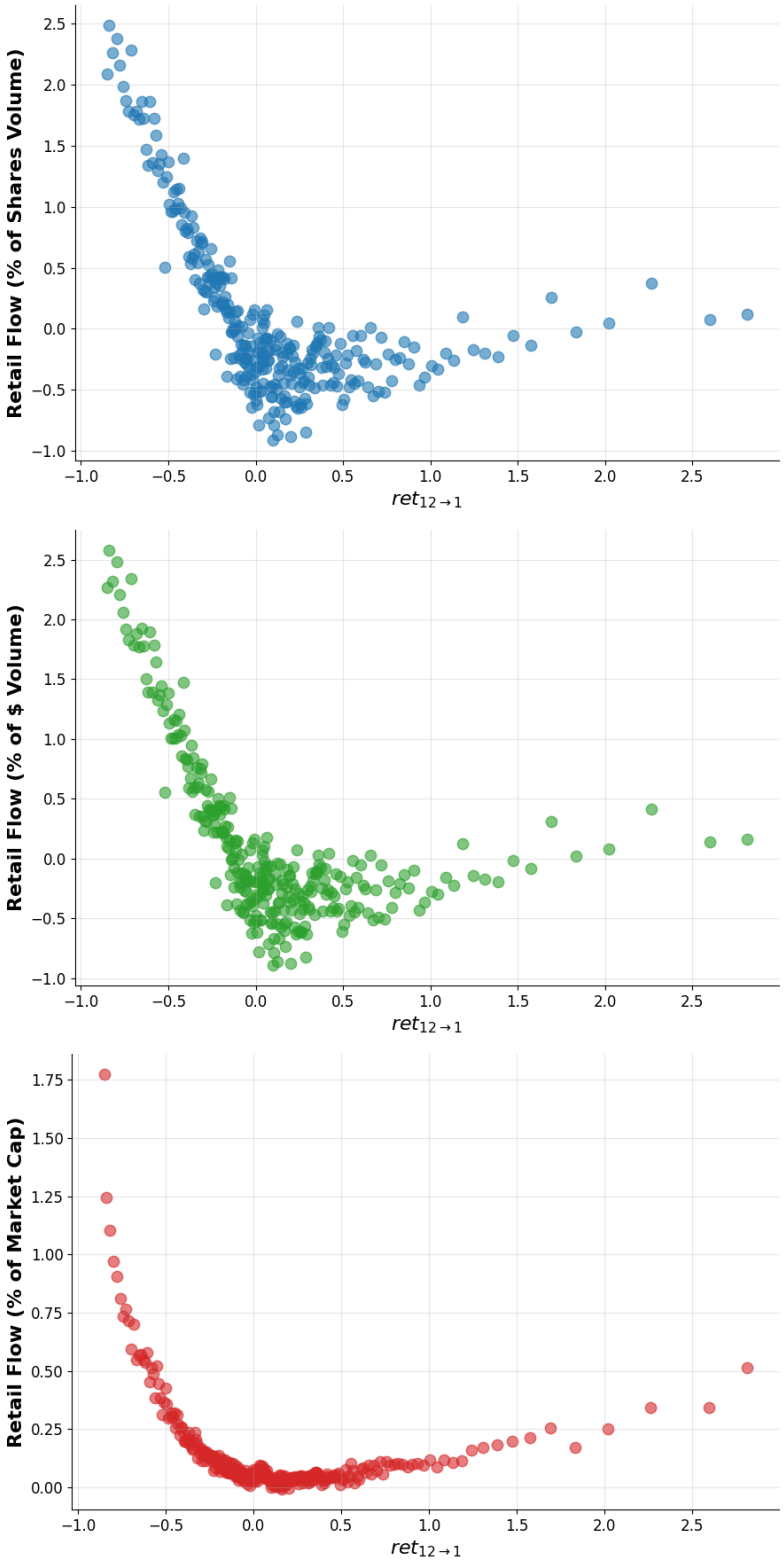
Appendix: Additional Figures and Tables

Figure A.1. Buying the Dip: Log Returns



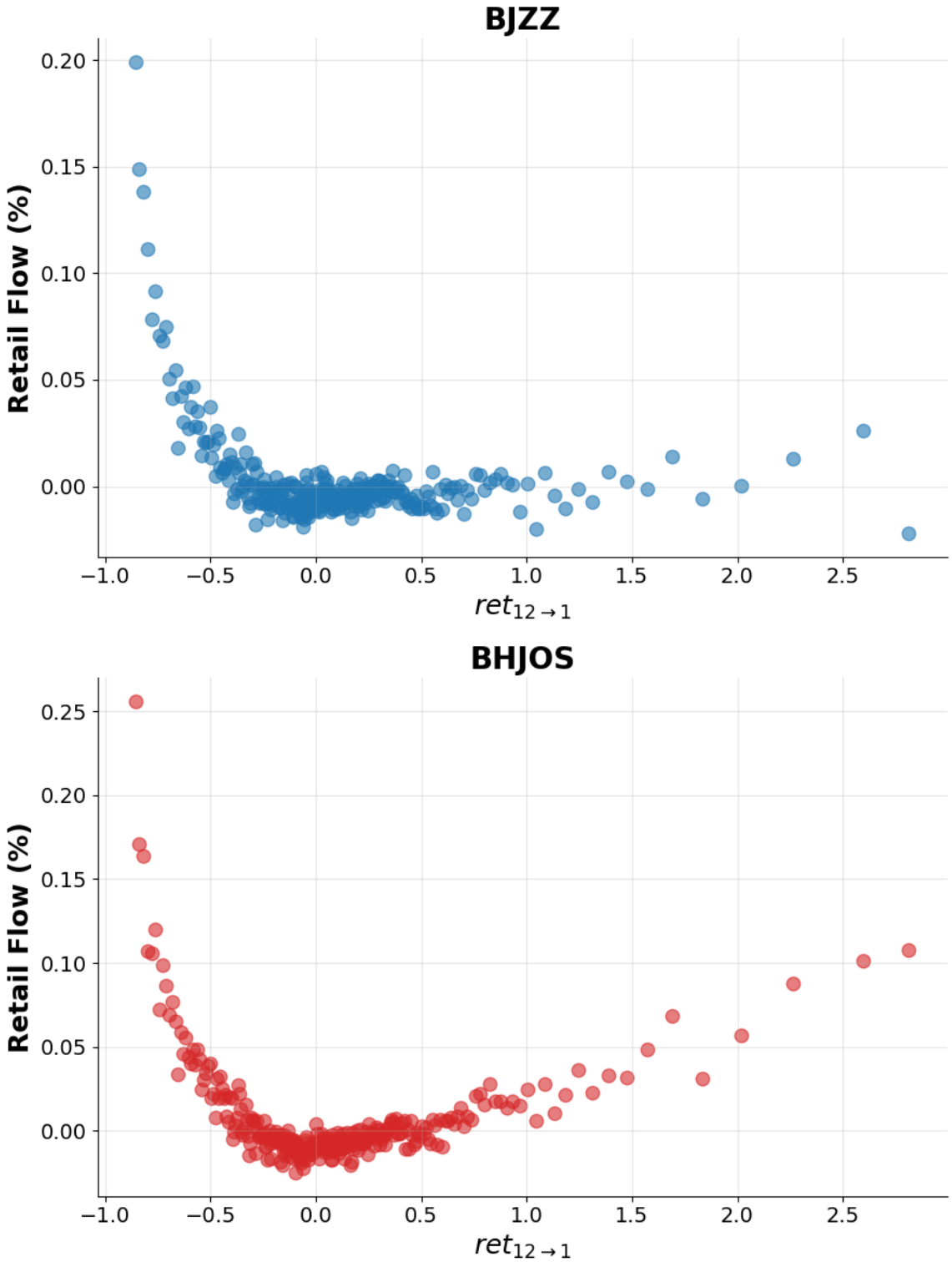
Notes: This figure replicates the main binscatter of retail order flow against lagged 12-to-1 month returns using a log-transformed return specification.

Figure A.2. Buying the Dip: Alternative Flow Denominators



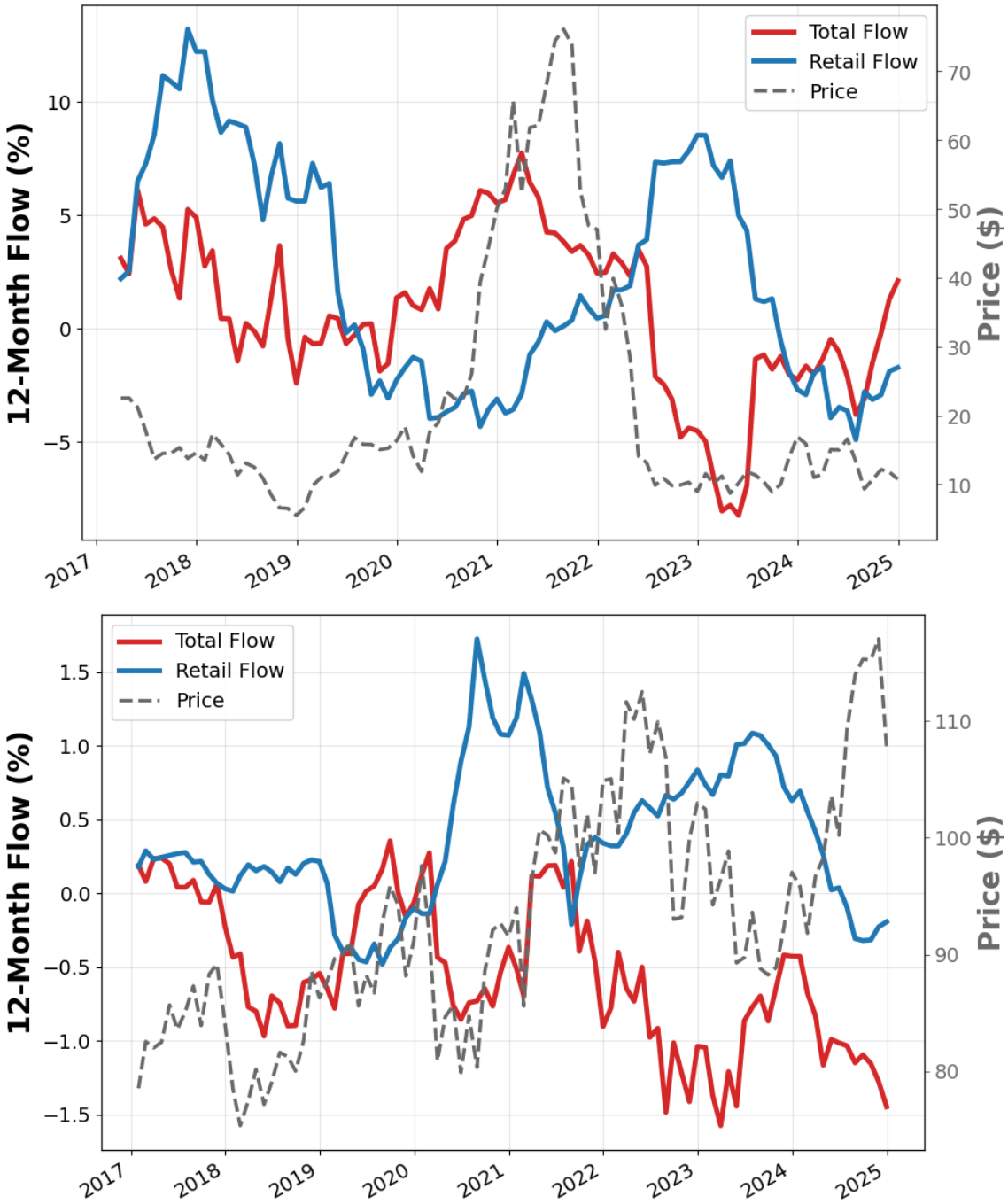
Notes: This figure replicates the main "buy the dip" binscatter using alternative retail flow denominators: flow as a percentage of shares volume (top), dollar volume (middle), and market capitalization (bottom).

Figure A.3. Buying the Dip: TAQ-Based Retail Flow Proxies



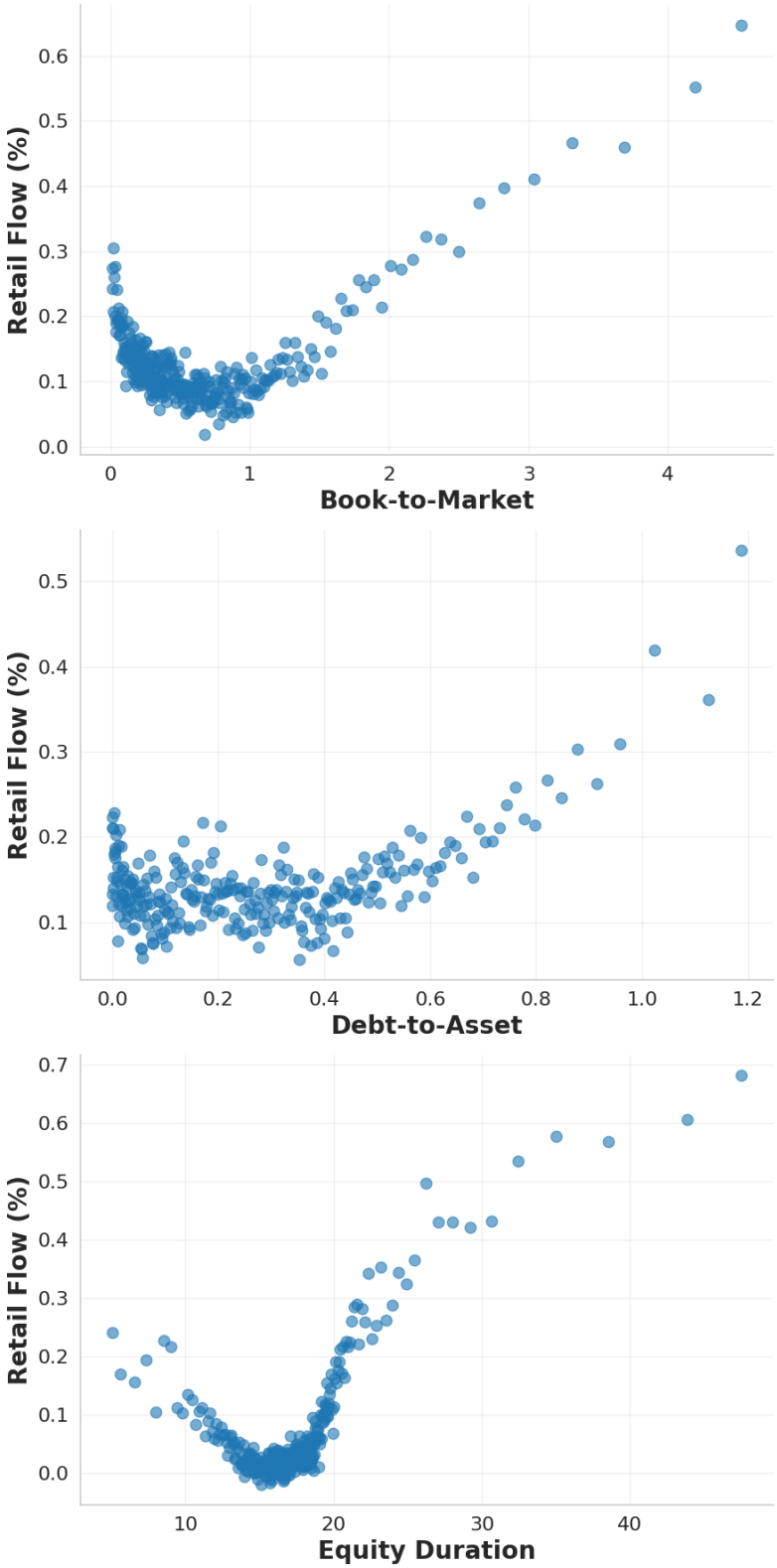
Notes: This figure plots binscatter estimates of the BJZZ net order flow measure (left) and the BHJOS net order flow measure (right) against lagged 12-to-1 month returns ($ret_{12\to1}$), expressed as a percentage of shares outstanding. Both panels can be compared directly to the main retail flow binscatter (Figure 4).

Figure A.4. Buying the Dip: Examples



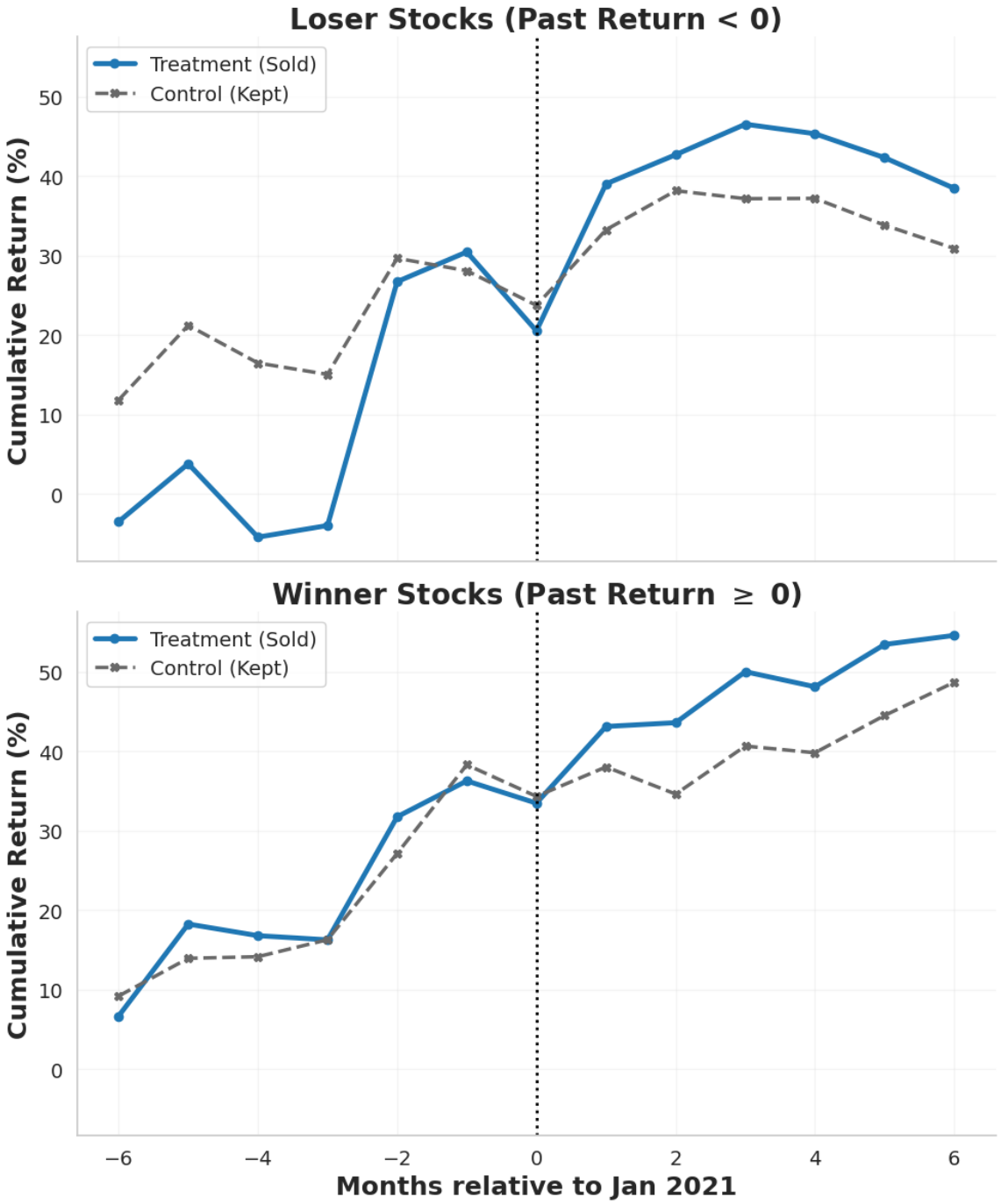
Notes: This figure illustrates the "buy the dip" behavior at the stock level for two examples: Snapchat Inc. (SNAP, top) and Duke Energy Corp. (DUK, bottom). Each panel plots the 12-month rolling retail flow (blue) and total flow (red) against the stock price (grey, right axis).

Figure A.5. Buying the Dip in Other Variables



Notes: This figure plots binscatter estimates of retail buy flow and retail sell flow separately against book-to-market ratio (top), debt-to-asset ratio (middle), and equity duration (bottom).

Figure A.6. Event Study of Melvin Capital: Cumulative Returns



Notes: This figure plots the cumulative returns of the treatment and control stocks during the Melvin Capital liquidation event.