Inelastic Demand at the Extensive and Intensive Margins

Xindi He — *Georgia Tech*
Allen Hu — *UBC*
Zigang Li * — *U of Toronto*

March 2026

Abstract

We decompose investor demand into its extensive margin—initiating or liquidating a position—and intensive margin—scaling an existing position. The extensive demand is economically large, accounting for 40% of institutional and 80% of retail flows, and its estimated price elasticity is closer to zero, and often positive. Extensive demand is primarily driven by past returns and attention, whereas intensive demand reflects fundamentals. To account for these findings, we develop a Grossman–Stiglitz–style model of inelastic markets where extensive flows are unobservable endogenous demand shocks, and investors learn from prices based on their subjective expectations about these flows. Learning from prices leads to a more inelastic, and even upward-sloping, extensive demand curve. Embedding the model in an asset demand system, we structurally estimate both elasticities under alternative specifications of flow expectations. In a counterfactual of reallocating capital from intensive to extensive demand, prices, volatility, and firm-specific price informativeness increase, while market-wide informativeness declines.

Keywords: Trading Flow, Demand System, Learning from Prices, Subjective Expectation
JEL Classification: G12, G14, G23, G40

*First version: November 2025. This version: March 2026. Latest version. We are grateful for comments from Nick Barberis, Carter Davis, Stefano Giglio, Zhiguo He, Jiacui Li, Song Ma, Kelly Shue, Stijn Van Nieuwerburgh, and seminar participants at HEC-McGill Winter Finance Workshop, Atlanta Fed, Chicago Booth, and Yale SOM. All remaining errors are ours. He: Georgia Tech Scheller, email: xindi.he@scheller.gatech.edu. Hu: University of British Columbia, Sauder School of Business, email: allen.hu@sauder.ubc.ca. Li: Ph.D. Candidate in Finance, Toronto Rotman, email: zigang.li@rotman.utoronto.ca.

1 Introduction

Recent work on demand system asset pricing underscores the value of directly studying, modeling, and estimating investor demand (Koijen and Yogo, 2019; Gabaix and Koijen, 2021). [1] Asset market quantities—especially demand and trading flows—are central to understanding price dynamics, market efficiency, and the impact of financial markets on the real economy (Lou, 2012; An, Su, and Wang, 2025; Choi et al., 2025).[2] A recurring theme in this literature is the substantial heterogeneity in investor demand and estimated price elasticities (Haddad, Huebner, and Loualiche, 2025).

Building on previous work, this paper studies investor demand and its heterogeneity through two novel perspectives. First, we decompose investor demand into its extensive and intensive margins: extensive flows are trades associated with investors' entry into or exit from stocks, while intensive flows are quantity adjustments to existing holdings. Our empirical evidence indicates that extensive and intensive flows have differential origins and impacts. And such distinctions likely originate from the heterogeneity in the informational content of these flows. Therefore, our second perspective highlights information dispersion as one important source of demand heterogeneity. We focus on a specific mechanism in which investors infer information about fundamentals from prices (Grossman and Stiglitz, 1980; Mendel and Shleifer, 2012; Banerjee and Green, 2015; Eyster, Rabin, and Vayanos, 2019; Bastianello and Fontanier, 2025a). In our framework, changes in prices not only affect investors' demand along a given demand curve but also shift the curve itself by altering investors' perceived information content of asset prices.

We begin by empirically establishing that extensive and intensive flows have different origins and impacts. Moreover, we find that the composition of flows matters for both return dynamics and price informativeness. We show that the extensive margin of investor demand is economically large, accounting for 40% of institutional and 80% of retail flows. Therefore, in the data, investors constantly switch among stocks, initiating new positions and liquidating old ones. Such behaviors deviate from the predictions of classic theories, such as the CAPM, that investors hold diversified portfolios and only rebalance at the intensive margin. This paper takes an information perspective

[1] For more work on demand system asset pricing, see Darmouni, Siani, and Xiao (2022); Bretscher et al. (2025); Chaudhary, Fu, and Zhou (2024); Chaudhry and Li (2025). There is a burgeoning literature on the estimation of these demand systems, including but not limited to Fuchs, Fukuda, and Neuhann (2023); Haddad et al. (2025); He, Kondor, and Li (2025); Koijen and Yogo (2025); van Binsbergen, David, and Opp (2025).

[2] For more work on trading flows, see Vayanos and Woolley (2011); Dou, Kogan, and Wu (2022); Hartzmark and Solomon (2025).

to explain this deviation: we develop a model in the spirit of Grossman and Stiglitz (1980) to highlight how learning from prices, information heterogeneity, attention, and specialization (Van Nieuwerburgh and Veldkamp, 2010; Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016) naturally lead to investor capital reallocation at the extensive margin. In our framework, "extensive" investors switch among stocks, generating "extensive" flows into or out of assets. Such flows cannot be observed in real time. As a result, investors, when inferring information from prices, need to disentangle price movements driven by changes in fundamentals from those originating from flow fluctuations, based on their subjective expectations about these flows.

Our model highlights two opposing effects of price movements: a standard substitution effect, which results in a downward sloping demand curve (Shleifer, 1986), and an information effect, which arises when investors infer fundamentals from prices (Grossman and Stiglitz, 1980) and pushes towards an upward sloping demand curve. Intuitively, after a price increase, the stock becomes "more expensive" to investors, so they want to substitute away and hold less of it. At the same time, investors also interpret the observed price increase as a positive signal on the stock's fundamentals, so they believe the stock is "better" than previously expected and want to hold more of it. Taken together, when investors determine how to shift their demand for an asset after a price increase, they face a tradeoff: the asset is now "better" but also "more expensive." As a result, when investors learn from prices, the demand curve becomes steeper and potentially upward sloping.

Empirically, to measure extensive and intensive flows, we use holdings data of both institutional investors from Thomson Reuters and retail investors from a large discount broker (Barber and Odean, 2000). We categorize each holding (i.e., an investor-stock-time observation) as either extensive or intensive demand, depending on whether the investor enters the stock from zero holding or exits the stock completely in this period. Intuitively, extensive demand is driven by investors' selection of stocks: whether an investor decides to hold a stock in her portfolio. And intensive demand represents adjustments to stocks already in the portfolio. We then measure flows as changes in holdings and net flows as the difference between buying flows and selling flows. To calculate percentage net flows, we divide net flows by the stock's lagged total shares outstanding (for net share flows) or by lagged market capitalization (for net dollar flows). These holding-level flow measures are then aggregated to the stock level to test their relations to returns and volatility.

We start the empirical analysis by documenting four stylized facts on extensive and intensive flows, from the levels of holdings to the aggregate. First, we define a stock holding cycle as an array of an investor's consecutive holdings in the same stock, from the opening of a position until its end. We show that, within a stock holding cycle, flows are concentrated at entry and exit: institutions buy 70% of their peak position on entry and sell 60% on exit, while retail investors buy 77% and sell 71% at those endpoints. Second, extensive demand makes up a large share of investor holdings. Retail trading is overwhelmingly extensive (84% of their flows), and even among institutions, the extensive flow is sizable—about 38% for investment advisors and roughly 25% for pension funds and insurance companies. Third, in the aggregate, the extensive margin remains economically important: for institutions, it averages 40% of quarterly flow, and for retail, it averages 80%. Finally, yet potentially most interestingly, extensive and intensive flows relate to returns in different ways. Net extensive flows are positively associated with contemporaneous returns but negatively with next-quarter returns; by contrast, net intensive flows are negatively associated with contemporaneous returns and display a weak positive relation with future returns. These facts highlight the conceptual and empirical differences between the extensive and intensive margins of investor demand.

To make the case for the distinction between extensive and intensive flows, we first shed light on their origins. We begin with a simple variation decomposition exercise for the flow type (extensive vs. intensive) and flow size using a Shapley-Owen R^2 decomposition with a rich specification of fixed effects. We demonstrate that whether investors trade at the extensive or intensive margins is primarily driven by investor-related dimensions (e.g., investor, investor-time, and investor-stock fixed effects). Therefore, whether an investor constantly switches among stocks or keeps holding the same set of stocks and rebalances is a persistent dimension of investor heterogeneity. In contrast, the flow size—especially at the extensive margin—is primarily driven by time-varying stock characteristics: stock-time fixed effects account for a sizable share of the variation in extensive flow intensity. These patterns suggest that the choice of margin is largely an investor trait, whereas the amount to trade is shaped by stock-level shocks, motivating our subsequent tests linking flows to attention, returns, and fundamentals.

We implement a within-portfolio empirical design on quarterly investor holdings to identify the driving forces behind investors' choice to trade at the extensive or intensive margins. Specifically, we regress the probability that a given demand is extensive (or intensive) on time-varying stock-level measures of attention, sentiment, past price movements, and fundamentals, controlling for investor-stock and investor-time fixed effects. This within-portfolio design captures how the same investor trades, at either extensive or intensive margins, across stocks with different characteristics. Such a design also rules out the influence of time-invariant investor preferences for specific stocks and investor-level shocks such as liquidity needs or fund flows. Three findings emerge. First, extensive flows are affected by social media attention and sentiment, whereas this social media channel is muted for intensive flows. Second, both margins respond to news coverage and news sentiment, but the intensive flow responds more strongly, consistent with its greater sensitivity to fundamentals. Third, extensive demand increases in past price movements (e.g., absolute lagged returns), while intensive demand tracks changes in fundamentals. Together, these results show that extensive flows originate in attention and recent price movements, whereas intensive flows reflect shocks to stock fundamentals, implying distinct informational content across the two margins.

We then document the impact of extensive and intensive flows on stock returns, volatility, and price informativeness. The intensity of investor extensive flow—measured by the proportion of shares traded at the extensive margin—has a significant impact on returns and volatility. In regressions with stock and time fixed effects and comprehensive controls, a higher extensive share is associated with higher contemporaneous returns and higher total and idiosyncratic volatility. In terms of the magnitude, a one-percentage-point increase in the extensive share corresponds to roughly a four-basis-point rise in the quarterly return and a five-basis-point rise in return volatility. These results underscore that the extensive and intensive margins of investor demand are not created equal for asset pricing: which margin dominates a stock's flow helps explain contemporaneous price movements.

Price informativeness responds to flow composition in a state-dependent way. On average, the relation between extensive flow intensity and informativeness is weakly positive. The relation becomes strongly positive when informativeness is already high: a one-percentage-point increase in the extensive share is associated with a 0.3-percentage-point rise in informativeness (a 3.4% increase relative to the mean). However, the correlation turns negative in low-informativeness states: the same increase in extensive margins leads to a 0.03-percentage-point decline in informativeness (a 4.5% decrease relative to the mean). We argue that this finding is consistent with the literature (Dávila and Parlatore, 2023) and with the economic mechanism that investors infer information from prices.

So far, we have been using flow composition as an explanatory variable for stock prices. We now treat flow composition as a dimension of stock heterogeneity. We document that extensive flows amplify the impact of news on fundamentals, measured by standardized unexpected earnings (SUE), on returns and volatility. At the average extensive share (30%), a one-percentage-point increase in SUE raises returns by an additional 0.09 percentage points beyond the baseline 0.23-percentage-point effect (i.e., a 39% amplification effect). The same increase in SUE reduces volatility by an additional 6.0 basis points beyond the baseline effect of −8.7 basis points (i.e., a 69% amplification effect). Thus, the composition of flows not only moves prices on its own but also shapes how the market incorporates news about fundamentals into prices.

As the last empirical exercise, we use local projections to trace the two-way dynamic relations between returns and net flows at both extensive and intensive margins. The estimated impulse response functions (IRFs) indicate that the contemporaneous link is strong in both directions: a one-percentage-point rise in net extensive flows increases returns by about 1.2 percentage points, while a one-percentage-point increase in returns boosts net extensive flows by roughly 2 basis points with a persistent effect in the next quarter. Echoing our previous cross-sectional evidence, we find extensive and intensive flows have opposite signs for their impact on returns. These empirical results anchor modeling choices and guide the design of the structural estimation and counterfactual analysis.

Motivated by our empirical findings, we develop a conceptual framework for extensive and intensive demand. We begin with a stylized model, a two-asset variant of Grossman and Stiglitz (1980), to introduce the key economic mechanisms, such as learning from prices and information heterogeneity, and illustrate the role of subjective expectations about unobserved extensive flows. In the model, a group of "extensive" investors can hold only one of two risky assets and switch between them, while a group of "intensive" investors can hold both assets. In a given period, only continuing investors observe a noisy signal about the asset's fundamentals. Meanwhile, non-holders of that asset try to infer the asset's fundamentals from its prices. The market cannot observe flows in real time. So non-holders face a key inference problem: they cannot fully tell to what extent an observed price change is due to fundamental news or due to unobserved flows. To solve such an inference problem, investors form subjective expectations about the intensity of extensive flow.

These subjective and potentially inaccurate beliefs shape how they infer information from prices. We show that, when investors underestimate the intensity of extensive flows, they overattribute observed price movements to changes in fundamentals and inaccurately update their beliefs about fundamentals. Such belief updating partly offsets the substitution effect from price movements and results in a more inelastic and potentially upward sloping demand curve. Under our framework with unobserved flows, extensive investors play a role similar to that of noise traders in Grossman and Stiglitz (1980), impeding prices from being fully revealing. The difference is that our model endogenizes the behavior of extensive investors, and their trading flows serve as unobservable endogenous demand shocks to the inelastic market.

Next, we develop a multi-asset extension of this stylized model. In this model, investors hold stocks at both the extensive and intensive margins. The information structure is the same as in the stylized model, and investors form subjective expectations about extensive flows. Flow expectations affect how investors learn from prices and, in turn, their portfolio choices. To structurally estimate the model, we derive a characteristics-based asset demand system in which expected payoffs and risks depend on stock characteristics. We model flow expectations under several specifications of belief formation, including statistically optimal beliefs based on public information (Bianchi, Ludvigson, and Ma, 2022), extrapolative beliefs (Barberis et al., 2015, 2018), and beliefs proxied by social media attention (Cookson et al., 2024). We then solve for the optimal asset demand: it is a linear function of stock prices, investor subjective expectations about extensive flows, public signals about idiosyncratic shocks to payoffs, and stock characteristics. The derived investor demand exhibits different price elasticities at the extensive and intensive margins.

We estimate the model on quarterly institutional holdings data. For identification, we adopt the optimal granular instrumental variables (GIV) proposed by Chaudhary, Fu, and Zhou (2024), which extends the approach of Gabaix and Koijen (2024). The estimation results are consistent with our conceptual framework. First, demand at the extensive margin exhibits a less negative, and often positive, price elasticity than demand at the intensive margin. The mean of the estimated price elasticity for extensive demand is -0.34, compared with -0.43 for intensive demand. There is a non-trivial mass of estimates with positive price elasticities, and the fraction of positive elasticities is larger for extensive demand than for intensive demand. Second, a higher expectation about extensive flow reduces investor demand. The estimated sensitivity of demand to the extensive flow expectation is -0.8. That is, a one-percentage-point rise in extensive flow expectations is associated with a 0.8-percentage-point decline in investor demand. Such a negative relation between flow expectation and demand is consistent with our mechanism of investor learning from prices. When investors believe it is the extensive inflow that pushes the price up, they don't update or even negatively update their beliefs about the fundamentals, decreasing their demand for that stock. We also find that estimation results based on extrapolative expectations are most consistent with our framework. Such findings suggest that investors, when forming their expectations about flows, deviate from an objective rational benchmark and overreact to recent realized flows.

Finally, we implement a counterfactual analysis to quantify the impact of a rise in investor demand at the extensive margin. Based on their average extensive demand share, we first assign investors to five quintiles. The investors in the top quintile with the highest extensive demand share are labeled as "extensive investors." We implement a budget-neutral reallocation that proportionally increases the AUM of extensive investors by 10% based on their initial AUM. The reallocation is financed by an equal-percentage AUM haircut on investors in the other four quintiles. Based on market clearing conditions, we calculate the counterfactual equilibrium prices after the reallocation. On average, the annualized stock return increases by 3.3 percentage points, and the annualized volatility rises by 3.7 percentage points. The results from the counterfactual analysis are consistent with our reduced-form empirical findings.

We also examine how this counterfactual increase in extensive demand would affect market efficiency. On average, the stock-specific price informativeness (Dávila and Parlatore, 2025) increases by 0.4 percentage points, a 9.3% increase relative to the sample mean. However, this effect is state-dependent. For stocks with above-median informativeness, there is a one percentage point increase in stock-specific price informativeness (a 12.1% rise), while for the below-median group, the effect turns negative—stock-specific price informativeness goes down by 0.2 percentage points, a 26.9% decline. More interestingly, the market-wide price informativeness, as measured in Bai, Philippon, and Savov (2016), declines after the capital reallocation from intensive to extensive demand, across all horizons of one, three, and five years. The effects on stock-specific and market-wide price informativeness have opposite signs because these two measures capture distinct aspects of market informational efficiency. Stock-specific price informativeness gauges, given a stock, how well its prices reflect its future fundamentals. As a result, stock-specific price informativeness is driven by within-stock time series variations. By contrast, market-wide price informativeness captures, given a cross section, whether the market assigns higher prices to stocks with better future fundamentals. Therefore, market-wide price informativeness is driven by cross-sectional variations across stocks. Relating back to our counterfactual analysis, on the one hand, higher extensive flows (i.e., more investors entering or exiting stocks) introduce additional unobserved demand shocks to the market, impeding across-stock price discovery and lowering market-wide price informativeness. On the other hand, investors infer fundamentals from prices when trading at the extensive margin. It is easier for them to do so when the price is already highly informative, and harder when the price is less informative. As a result, the relation between extensive flows and stock-specific price informativeness depends on the original level of informativeness, as we find in the counterfactual and reduced-form analyses.

The remainder of the paper proceeds as follows. Section 2 introduces data and establishes motivating stylized facts. Section 3 presents empirical results on the origin and impact of extensive and intensive flows. Section 4 develops the stylized model. Section 5 develops and estimates the multi-asset structural demand system, and Section 6 presents the structural results and counterfactuals. Section 7 concludes.

Related literature. Our paper is related to four strands of literature.

Learning from prices and the informational content of order flow. Foundational work shows that prices do not always fully reveal private information in equilibrium (Grossman and Stiglitz, 1980; Hellwig, 1980). Empirically, trades and order imbalances convey information and move prices in ways that blend fundamentals with contemporaneous flow (Hasbrouck, 1991; Llorente et al., 2002; Chordia, Roll, and Subrahmanyam, 2002). A more recent strand of literature emphasizes investor errors and biases when inferring information from prices. For example, Eyster, Rabin, and Vayanos (2019) study asset markets where investors neglect the information content of prices. Mondria, Vives, and Yang (2022) study the impacts on return volatility and volume if it is costly for investors to process information from asset prices. Bastianello and Fontanier (2025a,b) examine the price and volume implications if investors neglect that other investors may learn about fundamentals from prices. The mechanism in our paper is different. Investors in our model explicitly account for how other investors learn from prices, depending on whether they are extensive or intensive investors.

They correctly understand the informational differences between these two types of investors and how these differences affect learning from prices. Schmidt-Engelbertz and Vasudevan (2025) highlight that investors' higher-order beliefs can amplify stock market overreaction and excess volatility. Hartzmark and Sussman (2025), in an experimental setting, document that investors lack conviction as to what the stock market price should be, which they label as "price agnostic demand." We build on these papers but emphasize a distinct identification problem: because extensive reallocation, such as entry and exit, is unobserved in real time, investors who learn from prices rationally but potentially incorrectly attribute part of a price move to news on fundamentals. This misinference from prices is central to our stylized and estimation models and leads to heterogeneity in information among investors.

Demand curves for stocks and flow-based price pressure. Event studies around index changes establish downward-sloping demand and sizable short-horizon price impacts from non-fundamental flows (Shleifer, 1986; Harris and Gurel, 1986; Kaul, Mehrotra, and Morck, 2000; Chen, Noronha, and Singal, 2004). Beyond index events, a broader literature shows that institutional and fund flows are associated with price pressure, fragility, and predictability across assets (Greenwood, 2005; Coval and Stafford, 2007; Frazzini and Lamont, 2008; Campbell, Ramadorai, and Schwartz, 2009; Greenwood and Thesmar, 2011; Lou, 2012; Pavlova and Sikorskaya, 2023; Hendershott and Menkveld, 2014; Hartzmark and Solomon, 2025). We contribute by decomposing total trading into extensive and intensive flows and showing theoretically and empirically that the extensive and intensive flows have differential origins and impacts.

Investor heterogeneity. Micro evidence highlights systematic differences in who trades and how those trades affect prices (Barber and Odean, 2000, 2008; Kaniel, Saar, and Titman, 2008; Kelley and Tetlock, 2013). Our decomposition maps directly into these participation margins and helps explain why reallocations by certain investor groups generate stronger short-run return and volatility dynamics.

Demand system asset pricing. We extend the demand system asset pricing literature (Koijen and Yogo, 2019; van der Beck, 2022; Huebner, 2023; Koijen, Richmond, and Yogo, 2024; Gabaix et al., 2025; Davis, Kargar, and Li, 2025) by embedding information heterogeneity and subjective beliefs about extensive flow into a multi-asset demand system. The extension predicts attenuated price elasticities when investors perceive extensive flow as negatively correlated with prices. Such an economic mechanism provides a belief-based microfoundation for why asset markets are inelastic (Gabaix and Koijen, 2021). A related insight is provided by Chaudhry and Li (2025), who show that stock-level price multipliers—the per-unit price impact of uninformed demand shocks—are smaller for larger shocks, both contemporaneous and cumulative. Their empirical evidence indicates that price elasticities rise for securities experiencing larger price dislocations, a pattern consistent with models in which existing holders adjust their positions more aggressively in response to large shocks. Our framework is consistent with this evidence but emphasizes a different dimension of heterogeneity. Holding shock size fixed, the presence of extensive investors—who do not observe fundamentals and instead infer signals from prices—can mechanically lower effective price elasticity relative to a counterfactual market composed purely of intensive holders. Recent research also emphasizes that different modeling assumptions in the demand-system literature can yield different estimates of price elasticities. For example, He, Kondor, and Li (2025) show that in dynamic general-equilibrium environments, local demand curves may shift with shocks due to intertemporal hedging motives, implying that static demand systems identify short-run rather than long-run dynamic elasticities. Our model is explicitly static and focuses on within-period portfolio decisions. We therefore interpret the elasticities arising from our framework as cross-sectional, short-run elasticities, holding investors' information sets and subjective beliefs fixed. Graves (2025) also studies the role of beliefs and information in an asset demand system. Similar to our structural estimation approach, Graves (2025) recovers investor beliefs from holdings and highlights the difference in informational content between investors owning and not owning a stock, an angle analogous to the extensive and intensive margins of investor demand studied in our paper.

2 Data and Motivating Stylized Facts

This section first describes the data sources and sample construction, and then establishes four motivating stylized empirical facts about the extensive and intensive margins of investor demand.

2.1 Data Sources

Our data on investor holdings come from two sources: Thomson Reuters for institutional holdings, as used in Koijen and Yogo (2019), and a large discount broker for retail trading, as used in Barber and Odean (2000). We use CRSP, Compustat, and IBES to construct characteristics, data shared by Cookson et al. (2024) for social media attention and sentiment, and RavenPack for news coverage and news sentiment.

We obtain institutional stock holdings data from Thomson Reuters Institutional Holdings Database (S34) from 1980 to 2022. This database compiles the quarterly 13F filings from the Securities and Exchange Commission (SEC). All institutional investors that have assets under management (AUM) of more than $100 million are required to file the form to report stock positions. Institutional holdings data are merged with CRSP and Compustat data by CUSIP. Dollar holdings are computed as price times shares held. We follow Koijen and Yogo (2019) to classify institutional investors into six types based on their reported institution type in the Thomson Reuters database: investment advisors, mutual funds, pension funds, insurance companies, banks, and other 13F institutions. Investment advisors are firms that register with SEC Form ADV. Mutual funds are separated from investment advisors and form their own group. Other 13F institutions include endowments, foundations, and nonfinancial corporations. We define the household shares for any given stock as the difference between the total shares outstanding and the total shares held by all other 13F institutions. The household sector comprises direct household holdings and smaller institutions that are exempt from filing Form 13F.

We also draw on the retail investor dataset used in Barber and Odean (2000), which provides detailed trading records for a large set of self-directed brokerage accounts between 1991 and 1996. The sample allows us to observe how retail investors adjust their holdings at both the extensive and intensive margins. The retail holdings data complements our analyses on the institutional investors and strengthens the external validity of the results presented in this paper.

2.2 Sample and Variable Construction

We build three data sets: an investor-by-stock quarterly panel for holdings of institutional investors, an investor-by-stock monthly panel for holdings of retail investors, and a quarterly panel of stocks that aggregates investor holdings and augments stock and firm characteristics.

Investor demand. Using the holdings data described above, we construct investor demand variables. First, we categorize each position (i.e., an investor-stock-time observation) into one of the five types: extensive in, extensive out, intensive in, intensive out, and no trading. Extensive in means an investor opens a new position in a stock in this period. Similarly, extensive out means an investor exits a stock. Intensive in and out represent an investor's continuing holdings of a stock, with buying or selling shares in this period, respectively. And no trading indicates an investor has a position in a stock whose shares do not change. We then calculate the dollar value and the number of shares for each holding.

Based on holdings data, the flow measures are constructed as changes in holdings relative to the previous period's holdings. Net flow is the difference between buying flows and selling flows, in both dollar value and number of shares. To obtain the percentage net flows, we normalize dollar net flows by the stock's lagged market capitalization and normalize share net flows by the stock's lagged total shares outstanding.

For each institutional investor, we follow Koijen, Richmond, and Yogo (2024) and construct their active shares. The active share is defined as the total share of the investor's portfolio that deviates from the market weights (Cremers and Petajisto, 2009). Specifically, the active share is calculated as one-half the sum of the absolute differences between the portfolio weights and the market weights for the set of stocks in the investor's investment universe. We then average an investor's active share measures over time to generate an investor-level, time-invariant average activeness measure, which we used to study investor heterogeneity.

Finally, we aggregate variables constructed above from the level of investor holdings (i.e., investor-stock-quarter) to the level of stocks (i.e., stock-quarter). We obtain stock-level measures of aggregate investor demand and flows, such as percentage net flows, by demand type, including extensive in/out and intensive in/out. The composition of flows is defined as the percentage of extensive and intensive flows in shares, divided by the number of shares of total flows. As a result, the compositions of extensive flows and intensive flows always sum up to one hundred percent by definition.

Stock characteristics. We use CRSP, Compustat, and IBES to construct stock characteristics. Stock size is measured by the log of a stock's market capitalization. B/M is the ratio of book equity to market equity. Profitability is measured by operating cash flow to assets, following Bouchaud et al. (2019). Standardized unexpected earnings (SUE) is calculated as unexpected earnings (IBES actual earnings minus consensus analyst earnings forecast)[3] divided by fiscal-quarter-end market capitalization, following Rendleman Jr, Jones, and Latane (1982).

We use data from Cookson et al. (2024) on social media attention and sentiment. Both indices are standardized to have a mean of zero and a standard deviation of one. RavenPack provides measures of news coverage and news sentiment. News coverage is the log of the number of news articles covering a stock in a quarter. We focus on news articles with a relevance score above 75 in the RavenPack sample. Following Cookson et al. (2024), we use the Event Sentiment Score (ESS) calculated by Ravenpack as our measure for news sentiment.

For each stock in a given quarter, we calculate realized total volatility as the annualized monthly average of the standard deviation of cum-dividend daily stock returns from CRSP. Idiosyncratic volatility is constructed as the standard deviation of residuals in a regression of daily returns on daily market returns for the three years preceding the quarter end (Ali, Hwang, and Trombley, 2003).

We construct stock-specific price informativeness, which was recently developed by Dávila and Parlatore (2023, 2025). The data and programs are downloaded from the author's webpage. We extend the sample to 2022 using the programs shared by the authors. Under the framework of Dávila and Parlatore (2025), price informativeness τ_s measures the relative precision of information that asset prices contain about future payoffs. The measure is identified by comparing the R^2 from two OLS regressions of log-price changes (Δp_t) on contemporary and future log-payoff changes (Δx_t and Δx_{t+1}, with $\Delta \chi_t$ as the change in public signals):

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + \beta_2 \cdot \Delta \chi_t + e_t, \tag{1}$$

$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + \zeta_2 \cdot \Delta \chi_t + e_t^{\zeta}. \tag{2}$$

After collecting $R^2_{\Delta x, \Delta x'}$ and $R^2_{\Delta x}$ from the two regressions respectively. The relative price informativeness is then calculated as the normalized difference:

$$\tau_s = \frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}} \tag{3}$$

[3] If the consensus analyst earnings forecast is not available, then use the seasonally differenced quarterly earnings before extraordinary items from Compustat quarterly file

The numerator captures the percentage reduction in uncertainty about future payoffs after observing the price, while the denominator represents the residual uncertainty about future payoffs not explained by current public information and the realized payoff.

Table 1 presents the descriptive statistics for investor holdings (i.e., the investor-stock-quarter panel) in Panel (a) and stocks (i.e., the stock-quarter panel) in Panel (b). All variables are winsorized at the 1% and 99% cutoffs.

[Insert Table 1 Here.]

2.3 Motivating Stylized Facts

Using the samples constructed above, we document the following four stylized facts on investor extensive and intensive flows.

Fact 1: Extensive flow is large within a stock holding cycle.

We begin by examining the size of extensive and intensive flows at the finest level: stock holding cycles. We define a stock holding cycle as an array of an investor's consecutive holdings in the same stock, which always begins with an extensive inflow and ends with an extensive outflow. We normalize the holdings in shares by the maximum shares held within a stock holding cycle. Figure 1 presents how flows distribute within an average holding cycle.

[Insert Figure 1 Here.]

A strong empirical pattern emerges for both institutional investors in Panel (a) and retail investors in Panel (b). Extensive flows dominate in size within a stock holding cycle: flows are concentrated in the first and the last periods (F1 and L1 in the graphs) when an investor enters or exits a stock. When an institutional investor decides to enter a stock, she, on average, buys around 70% of her shares in the first period. Similarly, when she decides to exit a stock, she sells around 60% of her shares in the last period. The magnitude is even larger for retail investors, who buy 77% via extensive inflows and sell 71% via extensive outflows. As a robustness check, we exclude holding cycles that consist of only one buying observation and one selling observation. The results are presented in Figure A1. Although the magnitudes are slightly smaller, the conclusion remains unchanged: extensive flows still dominate within an average holding cycle.

Fact 2: Extensive flow is large in percentage composition.

We now turn to the level of investor holdings (i.e., investor-stock-time observations). We calculate the average dollar values and the percentage composition of extensive and intensive flows. Table 2 presents these statistics by investor type in Panel (a) and by active share quintile in Panel (b). To make the results more intuitive, we also visualize the percentage composition of flows in Figure 2.

[Insert Table 2 and Figure 2 Here.]

Retail investors own the largest share of the extensive demand, around 84%, indicating that they frequently switch stocks in their potentially underdiversified portfolios. For institutional investors, investment advisors implement more extensive trading, comprising 38%. Namely, around one-third of investment advisors' trading involves entering a new stock or completely exiting a stock in their portfolios. The extensive demand remains considerable, even for relatively passive investors such as pension funds and insurance companies, whose extensive investment compositions are 25% and 22%, respectively. Figure A2 presents graphs of the average dollar values of extensive and intensive flows, supporting the observation that extensive demand plays an important role in both institutional and retail investor holdings.

Fact 3: Extensive flow is large in aggregate.

We then aggregate investor holdings to the market level. We construct the quarterly time series of aggregate dollar values and the percentage composition of extensive and intensive flows. Figure 3 presents the aggregate flow percentage composition.

[Insert Figure 3 Here.]

For institutional investors in Panel (a), extensive demand accounts for up to 50%, with an average of around 40% across quarters. Extensive demand is more salient for retail investors in Panel (b), taking up to 90% of the total flow, with an average of around 80%. The large magnitude of extensive flows highlights the importance of decomposing total flows into extensive and intensive components and studying them separately. Figure A3 presents the time series of the aggregate dollar values, yielding a similar conclusion.

Fact 4: Extensive and intensive flows relate to stock returns differently.

The last but potentially the most interesting fact associates net percentage flows of extensive and intensive demand with contemporaneous and future returns. For each stock in a quarter, we construct net percentage extensive (intensive) flows as the number of shares bought minus those sold by extensive (intensive) investors, scaled by the lagged total number of shares outstanding for that stock. We then estimate the specification below on a stock (n)-by-quarter (t) panel to link demand flows to returns.

$$\text{Return}_{n,t+h} = \beta^h \cdot \text{Net Flow}_{nt} + \gamma \cdot \text{Controls}_{nt} + \alpha_n + \alpha_t + \epsilon_{nt}, \tag{4}$$

for $h = 0$ (contemporaneous returns) and $h = 1$ (returns next quarter). The controls include standardized unexpected earnings (SUE), size, book-to-market ratio, lagged return, and profitability. We also include stock fixed effects (α_n) and quarter fixed effects (α_t). β^h is the coefficient of interest, capturing how net extensive and intensive flows relate to returns in the h-th period. We present the binned scatterplots in Figure 4 to illustrate the relationship between returns and net extensive and intensive flows. In Section 3.3, we implement this intuitive regression specification (4) using a local projection framework and estimate the orthogonalized impulse response functions (IRFs). The results are similar, qualitatively and quantitatively.

[Insert Figure 4 Here.]

Figure 4 Panel (a) presents the results on contemporaneous returns: net extensive flows are significantly positively correlated with contemporaneous returns, while the intensive flows exhibit a significant negative association with returns. However, when we implement the same specification on future returns in Panel (b), the signs above flip. Net extensive flows are significantly negatively correlated with future returns, while the intensive flows are now positively related to future returns. Such a contrast highlights the distinction between extensive and intensive flows, suggesting different information content and asset pricing implications. Together with Facts 1–3, these stylized empirical patterns motivate our empirical analyses in Section 3.

3 Empirical Analysis

In this section, we examine the differential origins and impacts of extensive and intensive flows. We show that extensive flows are driven by past returns and attention, while intensive flows are more affected by fundamentals. Stocks with a higher composition of extensive demand have higher returns and volatility. When price informativeness is high (low), the extensive demand composition is positively (negatively) related to price informativeness. At the end of this section, we use a local projection framework to study the dynamic two-way relationship between net flows and returns.

3.1 The Origin of Extensive and Intensive Flows

In this section, we implement our analyses at the level of holdings. To fix ideas, the data consists of investor (i)-stock (n)-time (t) observations, with t being a quarter for institutional holdings and a month for retail holdings.

3.1.1 Variation Decomposition

To understand what drives extensive and intensive flows, we begin by decomposing their variations. Intuitively, we estimate OLS regressions of the flow type (i.e., indicator variables of whether a flow is extensive or intensive) and flow intensity (i.e., a flow's dollar value) on a comprehensive list of fixed effects, including time fixed effects (η_t), investor fixed effects (α_i), stock fixed effects (α_n), investor-time fixed effects (η_{it}), stock-time fixed effects (η_{nt}), and investor-stock fixed effects (α_{in}). To evaluate the independent contribution of each set of fixed effects to the variations in flows, we implement a Shapley-Owen variance decomposition (Audoly et al., 2025; Shorrocks et al., 2013). Intuitively, the procedure considers all possible orderings of introducing the fixed effects into the OLS regressions and averages their incremental contributions to the model fit (R^2), ensuring that no set of fixed effects is given undue priority simply because of the order in which it is added.

[Insert Figure 5 Here.]

Figure 5 visualizes the variations explained for flow type in Panel (a) and dollar values of extensive flows in Panel (b) and intensive flows in Panel (c). Figure 5 depicts a clear picture of which dimensions drive investors' decisions of investing at extensive or intensive margins and by how much. A key takeaway from the exercise is that investor-related dimensions account for the majority of the variations in flow type, suggesting that whether investors constantly switch across stocks ("extensive trading") or maintain a diversified portfolio and rebalance ("intensive trading") is an important dimension of investor heterogeneity and market dynamics. However, when examining the dollar values of investor flows, only the stock-time fixed effects can explain a considerable portion of the variations in extensive flows. This pattern motivates our analysis below to discover which dimensions of time-varying stock heterogeneity drive investor flows.

3.1.2 Relations to Sentiment, Attention, Returns, and Fundamentals

We augment investor holdings data with time-varying stock characteristics to study which factors drive investors' decisions in trading at the extensive or intensive margins. We focus on stock-level attention and sentiment from social media and news media, past price movement (e.g., the absolute value of lagged returns), and changes in fundamentals (e.g., standardized unexpected earnings). We estimate the specification below, linking these stock characteristics to the flow type at the level of investor holdings, using the investor (i)-stock (n)-time (t) observations.

$$\mathbf{1}(\text{Extensive/Intensive})_{int} = \beta \cdot X_{nt} + \eta_{it} + \alpha_{in} + \epsilon_{int}, \tag{5}$$

where $\mathbf{1}(\text{Extensive/Intensive})_{int}$ represents indicator variables in percentage terms for whether investor i's demand for stock n at time t is extensive or intensive. X_{nt} is the explanatory variable, such as social media attention and sentiment from Cookson et al. (2024), news coverage and sentiment from RavenPack, and lagged absolute values for stock returns and standardized unexpected earnings (SUE). Social media attention and sentiment, as well as news media sentiment, are standardized to have a mean of zero and a standard deviation of one. News coverage is measured by the log of the number of news articles covering a stock in a quarter. We include α_{in}, investor-stock fixed effects, accounting for unobserved time-invariant factors affecting investor i's demand on stock n, such as preferences, the time-invariant component of beliefs, proximity, and investor habitats. We also include η_{it} investor-time fixed effects, which capture investor-level shocks, such as liquidity, fund flows, fund structure, and investor-specific exposure to market fluctuations. By adding investor-time fixed effects, we effectively estimate a within-portfolio effect, which compares the same investor's choices between extensive and intensive trading, within the same period and within her portfolio, but across different stocks that have varying levels of attention, sentiment, price movement, and changes in fundamentals.

[Insert Table 3 Here.]

Table 3 reveals three main findings on the different origins of extensive and intensive flows. First, the extensive flow is significantly driven by social media attention and sentiment, as shown in Column (1). However, this social media pathway is muted for the intensive flow in Column (5). In terms of the magnitude, social media attention has a sizable impact: a one-standard-deviation increase in social media attention increases the probability of a flow being extensive by 0.68 percentage points, a 3% rise relative to the sample mean of 20.4 percentage points.

Second, both flow types react to news coverage and news sentiment, as presented in Columns (2) and (6). The intensive flow reacts more strongly to news media, suggesting a high sensitivity to asset fundamentals. Finally, the extensive flow is driven by past returns while the intensive flow reflects changes in fundamentals. For example, Column (3) shows a significant positive relationship between the probability of extensive flows and the size of past price movement, measured by the absolute values of lagged returns—the larger a stock's price moved last quarter, the higher the probability for an investor to enter or exit that stock, relative to other stocks in her portfolio with smaller price movement. On the contrary, the intensive flow has a significantly positive relationship with changes in fundamentals, as captured by the absolute values of standardized unexpected earnings (SUE). SUE measures the difference between realized earnings and the consensus analyst forecasts, which can be viewed as an unexpected shock to stock fundamentals. Column (8) indicates that a one-percentage-point increase in SUE results in a positive effect of 1.3 percentage points on the intensive flow probability.

Taken together, empirical results suggest that extensive and intensive flows originate from different sources and may have distinct information contents, with extensive flows driven by attention and past returns, and intensive flows reflecting asset fundamentals. Such findings motivate our stylized and estimation models and support the modeling assumptions on the distinction between extensive and intensive flows.

3.2 The Impact of Extensive and Intensive Flows

In this section, we examine the impacts of extensive and intensive flows on stock returns, volatility, and price informativeness. We aggregate investor holdings to the stock level, yielding a stock (n)-by-time (t) panel data for empirical analysis. The main finding suggests that the flow composition of a stock explains its return, volatility, and price informativeness—the larger the portion of extensive flows, the higher the contemporaneous return and volatility. The impact of extensive flow composition on price informativeness depends on the level of price informativeness itself. If a stock's price informativeness is high, extensive flow is positively correlated with informativeness. On the contrary, if a stock's price informativeness is low, extensive flow is then negatively correlated with informativeness.

3.2.1 Relations to Returns and Volatility

We first examine how the composition of investor flow relates to stock returns and volatility by estimating the specification below.

$$Y_{nt} = \beta \cdot \text{Composition}_{nt} + \gamma \cdot \text{Controls}_{nt} + \alpha_n + \alpha_t + \epsilon_{nt}, \tag{6}$$

where the outcome variable Y_{nt} is stock return, total volatility, and idiosyncratic volatility. The flow composition Composition_{nt} is the percentage composition of extensive flows. The percentage composition of intensive flow is omitted due to its perfect colinearity with extensive flow composition. We include stock fixed effects (α_n) and time fixed effects (α_t). Control variables include standardized unexpected earnings (SUE), size, book-to-market ratio, lagged return, and profitability.

[Insert Table 4 Here.]

Panel (a) in Table 4 shows the results. From Column (1) to Column (3), a consistent pattern emerges: a rise in the composition of extensive flows significantly increases contemporaneous return, volatility, and idiosyncratic volatility. Quantitatively, a one-percentage-point rise in extensive flow composition is associated with a four-basis-point increase in quarterly return and a five-basis-point increase in return volatility. If we shift the level of extensive flow composition from its 25th percentile, 19%, to its 75th percentile, 42%, the effect on return is 0.87 percentage points (i.e., a 32.5% increase from the sample mean of 2.68%) and the effect on volatility is 1.15 percentage points (i.e., a 2.45% increase from the sample mean of 46.96%). These results highlight the conceptual and empirical differences between extensive and intensive flows, demonstrating that the composition of flows has significant implications for asset pricing.

3.2.2 Relation to Price Informativeness

We now turn to the impact of flow composition on price informativeness, using the same specification (6) in the previous section. We focus on the stock-specific price informativeness (Dávila and Parlatore, 2023, 2025) whose construction is discussed in detail in Section 2.2.

Panel (b) in Table 4 displays the relationship between flow composition and price informativeness. We estimate the same regression specification first on the full sample and then on the samples with above-median price informativeness (the "high" sample) and below-median price informativeness (the "low" sample). Column (1) presents results estimated on the full sample, indicating a weakly positive relationship between extensive flow composition and price informativeness. If we focus on the sample of high informativeness, this positive relationship is strengthened significantly. A one-percentage-point increase in extensive flow is associated with a 0.3-percentage-point increase in price informativeness, representing a 3.4% increase from the sample mean. However, this positive relation flips sign in the sample of low informativeness—extensive flow is negatively correlated with price informativeness. A one-percentage-point increase in extensive flow is now associated with a 0.03-percentage-point decline in price informativeness, indicating a 4.5% decrease from the sample mean. Such a contrast between high- and low-informativeness samples is consistent with Dávila and Parlatore (2023): the papers argue that volatility and informativeness have a positive (negative) correlation when informativeness is sufficiently high (low). This argument naturally extends to our setting of extensive flows and corroborates our empirical findings, as we have shown in Panel (a) of Table 4 that extensive composition is positively correlated with volatility.

3.2.3 Flow Composition as Stock Heterogeneity

So far, we have been using flow composition as an explanatory variable for stock prices. Besides its direct impact on return and volatility, can flow composition function as a dimension of stock heterogeneity and alter how prices react to shifts in fundamentals? To answer this question, we estimate the regression specification below, which interacts flow composition with fundamentals and ties both to returns and volatility.

$$\begin{aligned} Y_{nt} = \beta_0 \cdot \text{SUE}_{nt} + \beta_1 \cdot \text{Composition}_{nt} + \beta_2 \cdot \text{SUE}_{nt} \times \text{Composition}_{nt} \\ + \gamma \cdot \text{Controls}_{nt} + \alpha_n + \alpha_t + \epsilon_{nt}, \end{aligned} \tag{7}$$

where the outcome variable Y_{nt} is stock return, total volatility, and idiosyncratic volatility. We measure a stock's changes in fundamentals by standardized unexpected earnings (SUE). We then include the percentage composition of extensive flow and interact it with SUE. The flow composition Composition_{nt} is the percentage composition of extensive flows. The percentage composition of intensive flow is omitted due to its perfect colinearity with extensive flow composition. We include stock fixed effects (α_n) and time fixed effects (α_t). Control variables include size, book-to-market ratio, lagged return, and profitability.

The coefficient of interest is β_2, the one for the interaction term of SUE and flow composition. β_2 captures the heterogeneous effects of SUE on outcomes with varying levels of extensive flows, in addition to SUE's baseline effect on outcomes, β_0.

[Insert Table 5 Here.]

Table 5 documents that the extensive flow composition, as a measure for stock heterogeneity, amplifies the effect of SUE on returns and volatility—all coefficients for the interaction term are highly statistically significant. In terms of the magnitudes, for a stock with the average extensive flow composition around 30 percentage points, a one-percentage-point increase in SUE will raise returns by 0.23 percentage points through its baseline effect, as well as by an additional 0.09 (=0.003×30×1) percentage points through the amplifying heterogeneous effect via extensive flow composition. As such, the average extensive flow composition amplifies the impact of fundamentals on returns by 39% (=0.09/0.23). Similarly, for a stock with the average extensive flow composition, a one-percentage-point increase in SUE will lower volatility by 8.7 basis points through its baseline effect, as well as by an additional 6.0 (=0.002×30×1) basis points through the amplifying heterogeneous effect via extensive flow composition. As such, the average extensive flow composition amplifies the impact of fundamentals on volatility by 69% (=6.0/8.7).

3.3 Dynamic Flow-Return Relation: Local Projections

In the final part of this section, we utilize local projections to study the dynamic two-way relationship between stock returns and net extensive and intensive flows. We use the following specification to estimate the orthogonalized impulse response functions (IRFs).

$$Y_{n,t+h} = \beta^h \cdot X_{nt} + \sum_{\tau=1}^{L} \psi^\tau \cdot Y_{n,t-\tau} + \sum_{\tau=1}^{L} \phi^\tau \cdot X_{n,t-\tau} + \gamma \cdot \text{Controls}_{nt} + \alpha_n + \alpha_t + \epsilon_{nt}, \tag{8}$$

for $h \in \{0, 1, ..., 6\}$ and $L = 3$. Y and X are returns, net extensive flows, and net intensive flows. The coefficients of interest are $\{\beta^h\}_{h=1}^{6}$, which capture the dynamic effects of net extensive and intensive flows on stock returns in the next h quarter or the effects in the other direction. The specification includes stock fixed effects (α_n) and time fixed effects (α_t), and controls for standardized unexpected earnings (SUE), size, book-to-market ratio, and profitability. We present the estimated effects visually in Figure 6.

[Insert Figure 6 Here.]

Figure 6 plots the orthogonalized impulse response functions (IRFs) of returns to a one-percentage-point increase in net extensive and intensive flows in Panel (a), and net extensive and intensive flows to a one-percentage-point increase in returns in Panel (b).

The figure reveals three main findings. First, the contemporaneous relationship between net flows and returns is strong, with high statistical significance and large economic magnitude. Notably, Panel (a) indicates that a one-percentage-point increase in net extensive flows leads to an immediate positive effect of 1.2 percentage points on stock return. In the other direction, Panel (b) shows that a one-percentage-point increase in stock returns results in an immediate positive effect of two basis points on net extensive flows, and such an effect persists for at least one quarter.

Second, the effects for extensive and intensive flows generally have opposite signs. Take the contemporaneous effects of flows on returns in Panel (a) as an example. The extensive flows are positively related to stock returns, while a rise in net intensive flows has a negative, though relatively small, impact on stock returns. This finding is consistent with the motivating empirical Fact 4 presented in the previous section.

Finally, effects for extensive flows are strong in the short run but exhibit a reversal in the long run. As Panel (a) indicates, the effect of extensive flows on returns is positive in the same period, but the effects on future returns all turn negative. A similar pattern holds for the effect of returns on extensive flows: it is significantly positive within the first three periods but becomes negative starting from the fourth period, suggesting potential overshooting over the short run and then reversal over the long run.

4 Stylized Model

This section develops a stylized Grossman-Stiglitz-style model to clarify why investor demand at the extensive margin can be less downward-sloping (and sometimes upward-sloping) than demand on the intensive margin. Motivated by our empirical results, we build in two ingredients: (i) an information gap between continuing holders and non-holders; and (ii) unobserved extensive reallocation (entry/exit/switching) that investors must disentangle from fundamentals when learning from prices.

4.1 Set-up

Consider an economy with two risky assets, stocks A and B, each with unit supply, and a risk-free asset yielding zero net return. To transparently illustrate the mechanism, we use a simple three-period static model ($t = 1, 2, 3$). Stocks A and B pay dividends d_A and d_B only at $t = 3$, and their prices $p_{A,t}$ and $p_{B,t}$ for $t = 1, 2$ are endogenously determined in equilibrium. The two-asset structure captures extensive reallocation while keeping the model tractable.[4]

In each period, any given investor exits the market with probability δ. An equal mass of new investors enters, so that the total mass of investors is normalized to one at all dates. This structure generates investor turnover while preserving a constant market size, and allows for unobserved

[4]For instance, stock A may represent a group of stocks directly hit by a shock, whereas stock B represents the remaining market segment unaffected by it.

extensive flows even in the two-asset environment.

There are two investor types: "intensive" investors (I) with mass $1-\phi$ and "extensive" investors (E) with mass ϕ, where $\phi \in (0,1)$. Intensive investors can invest in both assets. Extensive investors may hold only one of the two risky assets at a time, and can switch between them across periods. Let $E_{A,t}$ and $E_{B,t}$ denote the subsets of extensive investors holding A and B in period t, respectively, and let θ_t denote the fraction of extensive investors holding A in period t.[5]

All investors are atomic price takers with identical initial wealth W_0 and share homogeneous mean-variance preferences over terminal wealth W_3, under standard CARA-normal assumptions.[6] For $t = 1, 2$, each investor maximizes

$$V_t = \mathbb{E}_t [W_3] - \frac{1}{2}\alpha \mathrm{Var}_t (W_3) .$$

At $t = 1$, all investors share common priors about risky payoffs at $t = 3$,

$$\begin{cases} d_A = \mu_A + \varepsilon_A \\ d_B = \mu_B + \varepsilon_B \end{cases} \quad \text{with } \varepsilon_A \sim \mathcal{N}\left(0, \sigma_A^2\right),\ \varepsilon_B \sim \mathcal{N}\left(0, \sigma_B^2\right),\ \varepsilon_A \perp \varepsilon_B,$$

where $\mu_A, \mu_B, \sigma_A, \sigma_B$ are known constants. Intensive investors choose optimal holdings $q^I_{A,1}$ and $q^I_{B,1}$, while extensive investors first select which stock to hold and then choose quantities.

In period $t = 2$, an unexpected noisy signal about each asset's dividend is revealed:

$$\begin{cases} \eta_A = d_A + e_A, \\ \eta_B = d_B + e_B, \end{cases} \quad \text{with } e_A \sim \mathcal{N}(0, v_A^2),\ e_B \sim \mathcal{N}(0, v_B^2),\ e_A \perp e_B \perp \varepsilon_A \perp \varepsilon_B. \tag{9}$$

Only intensive investors observe a stock's signal: at $t = 2$, investors holding A at $t = 1$ observe η_A but not η_B, and vice versa. Investors who do not observe a signal must infer it from the corresponding price at $t = 2$.

In period $t = 3$, dividends d_A and d_B are realized and investors consume terminal wealth W_3.

[5] For example, at $t = 1$, the total mass of extensive investors holding A is $\phi\theta_1$.

[6] Under CARA-normal preferences, the wealth distribution plays no essential role.

4.2 Pre-Signal Equilibrium ($t = 1$)

Given the common information set at $t = 1$, intensive investors choose $\boldsymbol{q}_{I,1} = (q^I_{A,1}, q^I_{B,1})'$, while extensive investors choose $q^{E_A}_{A,1}$ or $q^{E_B}_{B,1}$ depending on which stock they hold. Prices $\{p_{A,1}, p_{B,1}\}$ are determined such that both risky asset markets clear and extensive investors are indifferent between holding A or B.

Definition 1. *$\{p^\star_{A,1}, p^\star_{B,1}\}$ is a pre-signal equilibrium if and only if there exist $\boldsymbol{q}^\star_{I,1}$, $(q^{E_A}_{A,1})^\star$, $(q^{E_B}_{B,1})^\star$, and $\theta^\star_1$ such that:*

(i) Utility maximization:

$$\boldsymbol{q}^\star_{I,1} \in \arg\max_{\boldsymbol{q}_{I,1}} V_1, \tag{10}$$

$$(q^{E_A}_{A,1})^\star \in \arg\max_{q^{E_A}_{A,1}} V_1, \tag{11}$$

$$(q^{E_B}_{B,1})^\star \in \arg\max_{q^{E_B}_{B,1}} V_1. \tag{12}$$

(ii) Market clearing:

$$(1-\phi)(q^I_{A,1})^\star + \theta^\star_1 \phi (q^{E_A}_{A,1})^\star = 1 \tag{13}$$

$$(1-\phi)(q^I_{B,1})^\star + (1-\theta^\star_1)\phi (q^{E_B}_{B,1})^\star = 1 \tag{14}$$

(iii) Indifference for extensive investors:

$$V_1(A) = V_1(B). \tag{15}$$

Proposition 1. *Under the regularity condition*[7]

$$(1-\phi)\,\sigma_B < \sigma_A < \frac{1}{1-\phi}\sigma_B, \tag{16}$$

[7]The regularity condition (16) ensures that market clearing admits an interior solution, i.e., $\theta^\star_1 \in (0, 1)$.

the equilibrium in Definition 1 is given by:

$$(q^I_{A,1})^\star = (q^{E_A}_{A,1})^\star = \frac{\mu_A - p^\star_{A,1}}{\alpha\sigma_A^2}, \qquad p^\star_{A,1} = \mu_A - \frac{\alpha\sigma_A^2}{(1-\phi)+\theta_1^\star\phi},$$
$$(q^I_{B,1})^\star = (q^{E_B}_{B,1})^\star = \frac{\mu_B - p^\star_{B,1}}{\alpha\sigma_B^2}, \qquad p^\star_{B,1} = \mu_B - \frac{\alpha\sigma_B^2}{(1-\phi)+(1-\theta_1^\star)\phi}, \tag{17}$$
$$\theta_1^\star = \frac{\sigma_A - (1-\phi)\sigma_B}{(\sigma_A+\sigma_B)\phi}.$$

Proof. See Appendix A.1. □

Two remarks are useful for later. First, since all investors share the same information at $t = 1$, intensive and extensive investors hold the same per-capita quantities. Second, the composition $\theta_1^\star$ affects prices mechanically through market clearing: shifting more extensive investors into A raises $p^\star_{A,1}$ and lowers $p^\star_{B,1}$.

4.3 Post-Signal Equilibrium ($t = 2$): learning from prices with subjective beliefs about extensive flows

At $t = 2$, holders of each stock observe its signal in (9). For stock $n \in \{A, B\}$, investors who observe η_n update their beliefs about d_n via Bayes' rule:

$$d_n \sim \mathcal{N}(\widehat{\mu}_n, \widehat{\sigma}_n^2),$$

where

$$\widehat{\mu}_n = \mu_n + g_n(\eta_n - \mu_n), \qquad g_n = \frac{\sigma_n^2}{\sigma_n^2 + v_n^2}, \qquad \widehat{\sigma}_n^2 = \frac{\sigma_n^2 v_n^2}{\sigma_n^2 + v_n^2}. \tag{18}$$

Unobserved extensive flow and the inference problem. The key friction is that non-holders do not observe η_n and must infer it from $p_{n,2}$. They understand, however, that $p_{n,2}$ reflects not only fundamentals but also unobserved reallocation at the extensive margins. Between $t = 1$ and $t = 2$, a fraction δ of investors exits the market. For stock A, both holders and non-holders exit proportionally. All entrants are new investors who arrive as non-holders and whose initial portfolio choices across A and B are not observed. This entry–exit process generates extensive trading flows that are hidden from non-holders but nevertheless impounded into prices.

To align this force with the traditional Grossman–Stiglitz noise-trade channel and to keep signal extraction tractable, we define extensive flow in quantity (shares). Specifically, let the (net) extensive inflow into stock A at $t = 2$ be m_A, measured in shares normalized by total supply (recall supply is normalized to one). A positive m_A reflects net buying pressure into A by investors who did not hold A at $t = 1$, while $m_A < 0$ reflects net selling pressure. When continuing investors on the intensive margin do not exit under the indifference condition, aggregate extensive flows satisfy $m_A + m_B = \delta$. These flows are not observed in real time by investors.

We allow a non-holder to form a subjective belief about this extensive flow. Specifically, for an investor i who does not observe η_A, let her perceived extensive inflow be

$$\hat{m}_{A,i} = \tilde{m}_{A,i} + \tilde{\epsilon}_{m,A,i}, \tag{19}$$

where $\tilde{m}_{A,i}$ is subjective mean and $\tilde{\epsilon}_{m,A,i}$ is a mean-zero noise term capturing idiosyncratic belief errors (or coarse inference) about reallocation.[8] We define $\hat{m}_{B,i}$ analogously.

Optimal demand at $t = 2$. Conditional on a perceived signal (true or inferred), the mean-variance optimal demand keeps the same form. For any investor who uses perceived posterior $(\widehat{\mu}_n, \widehat{\sigma}_n^2)$ for asset n,

$$q_{n,2} = \frac{\widehat{\mu}_n - p_{n,2}}{\alpha\widehat{\sigma}_n^2}.$$

Thus, investors differ only through what posterior they plug in: holders use the true η_n, while non-holders infer an implied $\tilde{\eta}_n$ from price and their belief about $\tilde{m}_n$.

How non-holders infer the signal from prices. A non-holder understands that, under CARA-normal demand, market clearing links the equilibrium price to (i) posterior mean fundamentals and (ii) the residual supply that must be absorbed by informed holders after extensive reallocation. Let

$$\bar{\theta}_A \equiv \left[(1-\phi) + \phi\theta_1^\star\right](1-\delta) \qquad \text{and} \qquad \bar{\theta}_B \equiv \left[(1-\phi) + \phi(1-\theta_1^\star)\right](1-\delta)$$

[8] With $\hat{m}_{A,i}$ interpreted as a quantity (shares), the belief noise $\tilde{\epsilon}_{m,A,i}$ enters the market-clearing condition linearly, directly analogous to the standard noise-trader demand in Grossman–Stiglitz. This preserves normality of the inferred signal under normal belief noise.

denote the total mass of investors who held A and B at $t = 1$, survived to $t = 2$, and hence observed the signal, respectively. Then, from the non-holder's perspective, extensive flow $\tilde{m}_{A,i}$ shifts the residual supply for informed holders from 1 to $1 - \tilde{m}_{A,i}$, so she inverts the price into an implied signal.

Formally, the results below summarize (i) optimal demands and (ii) the implied signal extraction. We report the key expressions in the text and place derivations in Appendix A.2–A.3.

Proposition 2 (Optimal demands at $t = 2$)**.** *For investors who observe η_A and η_B, their optimal demands are*

$$q_{A,2}(\eta_A) = \frac{\mu_A + g_A(\eta_A - \mu_A) - p_{A,2}}{\alpha\widehat{\sigma}_A^2}, \tag{20}$$

$$q_{B,2}(\eta_B) = \frac{\mu_B + g_B(\eta_B - \mu_B) - p_{B,2}}{\alpha\widehat{\sigma}_B^2}. \tag{21}$$

For investors who do not observe η_A (resp. η_B) and instead use an inferred signal $\tilde{\eta}_A$ (resp. $\tilde{\eta}_B$), their demands take the same form with η replaced by $\tilde{\eta}$:

$$q_{A,2}(\tilde{\eta}_A) = \frac{\mu_A + g_A(\tilde{\eta}_A - \mu_A) - p_{A,2}}{\alpha\widehat{\sigma}_A^2}, \tag{22}$$

$$q_{B,2}(\tilde{\eta}_B) = \frac{\mu_B + g_B(\tilde{\eta}_B - \mu_B) - p_{B,2}}{\alpha\widehat{\sigma}_B^2}. \tag{23}$$

Proof. See Appendix A.2. □

Corollary 1 (Signal extraction with subjective beliefs on extensive flows)**.** *Consider an investor i who does not observe η_A at $t = 2$ and forms a belief $\tilde{m}_{A,i}$ as in (19). Her inferred signal $\tilde{\eta}_{A,i}$ satisfies:*

$$\tilde{\eta}_{A,i} = \frac{1}{g_A}\left(p_{A,2} + \alpha\widehat{\sigma}_A^2\tilde{\xi}_{A,i} - \mu_A\right) + \mu_A. \tag{24}$$

where $\tilde{\xi}_{A,i} \equiv \left(1 - \tilde{m}_{A,i}\right)/\bar{\theta}_A$. An analogous expression holds for $\tilde{\eta}_{B,i}$ with $\bar{\theta}_B$ and $\tilde{m}_{B,i}$.

Proof. See Appendix A.3. □

Why subjective flow beliefs attenuate (and can flip) price elasticity. Equation (24) clarifies how flow beliefs reshape learning from prices. If a non-holder underestimates extensive inflows into A, i.e., $\tilde{m}_{A,i}$ is too small relative to the true extensive buying pressure, she infers that informed holders must absorb a larger residual supply, $1 - \tilde{m}_{A,i}$. Holding the observed price $p_{A,2}$ fixed, she therefore attributes an overly large share of the high price to fundamentals and overestimates $\tilde{\eta}_{A,i}$. As a result, her extensive demand becomes less sensitive to price increases, mechanically attenuating the magnitude of the negative extensive price elasticity and, in some cases, potentially producing upward-sloping extensive demand (i.e., positive price elasticity).

One potential scenario arises when expected extensive flows are positively correlated with prices. higher price leads the investor to infer smaller extensive inflows, so she treats the price as more informative and overlearns from it. Consistent with the logic above, this channel makes the extensive elasticity less negative and can rationalize positive elasticities. This stylized comparative-static mirrors the mechanism we revisit in the structural estimation: if investors believe extensive flows are negatively related to price, the model can generate a non-trivial mass of positive estimated elasticities on the extensive margin.

Subjective Flows Equilibrium (SFE). We define an equilibrium at $t = 2$ allowing (i) unobserved extensive flows and (ii) subjective beliefs about them. In SFE, prices clear markets, intensive investors and informed extensive investors optimize using true signals, and non-holders optimize using inferred signals based on (19)–(24). We state the equilibrium in an abstract way as follows to keep the main text light; More discussion is provided in Appendix A.4.

Definition 2 (Subjective Flows Equilibrium)**.** *A Subjective Flows Equilibrium (SFE) at $t = 2$ is a tuple*

$$\left\{p_{A,2}^{SFE}, p_{B,2}^{SFE}, \theta_2^{SFE}, \{\tilde{m}_{A,n}, \tilde{m}_{B,n}\}_n\right\}$$

such that: (i) given perceived information (true signals for holders; inferred signals for non-holders), investors choose optimal demands as in Proposition 2; (ii) markets clear for both assets; and (iii) extensive investors are indifferent between holding A and B at $t = 2$, where non-holders form beliefs $\tilde{m}_i$ as in (19).

The key implication is that unobserved extensive reallocation limits price revelation: even if non-holders rationally learn from prices, they must take a stand on how much of a price move reflects fundamentals versus extensive flow. Subjective beliefs about extensive flow therefore generate heterogeneity in learning from prices and deliver distinct price elasticities across extensive and intensive margins.

5 Structural Estimation

In this section, we extend the two-asset stylized model in Section 4 to a more general setting with multiple assets and heterogeneous investors. Our goal is to obtain a tractable asset demand system that (i) preserves the core economic mechanisms distinguishing extensive from intensive demand and (ii) can be empirically estimated using stock characteristics and portfolio holdings.

Our structural model shares the same spirit but extends the characteristics-based demand system in Koijen and Yogo (2019) and Koijen, Richmond, and Yogo (2024). Specifically, we embed learning from prices (Grossman and Stiglitz, 1980) and subjective expectations about flows into a demand system. Expectations about extensive flows shape their demand by altering how they interpret the informational content of prices. We next present the full model setup and solve for investor demand and equilibrium prices. We then describe how we empirically estimate such a characteristics-based demand system, under alternative specifications of flow expectations.

5.1 Model Setup and Solution

We consider a static, two-period model ($t \in \{0, 1\}$). The economy consists of N risky assets and one riskless asset. The riskless asset offers a gross return R_f and is in perfectly elastic supply. Each risky asset n pays a dividend $d(n)$ at $t = 1$. The vector of dividends for all N assets, d, is determined by a factor model:

$$d = \mu + \rho F + \varepsilon \tag{25}$$

where d, μ, and ε are $N \times 1$ vectors. μ is the vector of expected dividends. F is a scalar systematic factor, assumed to follow a standard normal distribution, $F \sim N(0, 1)$. ρ is an $N \times 1$ vector of factor loadings. ε is an $N \times 1$ vector of idiosyncratic shocks. The shocks are normally distributed, $\varepsilon \sim N(0, \Sigma_\varepsilon)$, where $\Sigma_\varepsilon = \sigma_\varepsilon^2 I$ is a diagonal covariance matrix. The supply of risky assets is normalized to $\mathbf{1}$.

Investors. There is a continuum of investors, indexed by i. All investors have constant absolute risk aversion (CARA) utility over the period-1 wealth W_{1i}, with a risk aversion coefficient of γ_i. Their objective is to:

$$\max_{q_i} E[-\exp(-\gamma_i W_{1i})|\mathcal{I}_i] \tag{26}$$

where q_i is the vector of asset holdings for investor i and $\mathcal{I}_i$ is their information set.

At the beginning of $t = 0$, investors hold a subset of the risky assets. We denote holdings as q_{0i} and the subset of assets that investor i holds as $\mathcal{N}_{0i} \subseteq \{1, ..., N\}$. We define a selection vector s_i where $s_i(n) = 1$ if $n \in \mathcal{N}_{0i}$ and 0 otherwise. The corresponding selection matrices are defined as: $s^I = diag(s_i)$ and $s^E = diag(\mathbf{1} - s_i)$.[9] For an asset n, an investor i is classified as an "intensive investor" if $n \in \mathcal{N}_i$ (i.e., the investor holds asset n at the beginning of period 0), and as an "extensive investor" if $n \notin \mathcal{N}_i$ (i.e., the investor does not hold asset n at the beginning of period 0). Their budget constraints are:

$$W_{1i} = q'_{0i} p R_f + q'_i (d - p R_f) \tag{27}$$

Information gap between investors. A key feature of this model is the information gap between extensive and intensive investors. Investors who hold the asset n at period 0 receive a noisy signal about the asset's payoff. Such noisy signal, $\eta(n)$, is about the idiosyncratic shock to the asset payoff, $\varepsilon(n)$. We write the signals of different assets in a vector form:

$$\eta = \varepsilon + e, \quad \text{where } e \sim N(0, \Sigma_\eta). \tag{28}$$

The noise for each asset's signal is independent. The noise covariance matrix is a diagonal matrix and $\Sigma_\eta = \sigma_\eta^2 I$. The variance of the signal noise for each asset is σ_η^2.

Belief updating. Investors update their beliefs about ε using their signals and, in the case of extensive investors, the information from prices. Extensive investors infer information from market prices based on their perceived extensive flows. Following the setup in Section 4.3, the extensive flow of a stock is defined as the total quantity (shares) held by the extensive investors. An extensive investor i forms a subjective expectation about this extensive flow. Let $\hat{m}_i = \left[\hat{m}_i(1), \hat{m}_i(2), \ldots, \hat{m}_i(n)\right]^\top$

[9] For each investor, we sort the assets to concentrate the non-zero diagonal elements of s^I in the top-left corner.

denote the vector of investor i's expectations. Extending the definition in (19), we obtain

$$\hat{m}_i(n) = \tilde{m}_i(n) + \tilde{\epsilon}_{m,i}(n), \tag{29}$$

where $\tilde{m}_i(n)$ is the subjective mean and $\tilde{\epsilon}_{m,i}(n)$ is a zero-mean noise term capturing idiosyncratic belief errors. One potential microfoundation for such a functional form is noisy expectations (Patton and Timmermann, 2010; De Silva and Thesmar, 2024). Moreover, as we illustrate in later sections, investors acknowledge the noise in their forecasts and beliefs and are aware of their own cognitive uncertainty (Enke and Graeber, 2023). The perceived price signal by extensive investor i is denoted by η_{p_i}.

In vector-matrix form, the posterior beliefs for any investor i are:

$$E[\varepsilon|\mathcal{I}_i] = \hat{\varepsilon}_i = \hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}_i}^{-1}\bar{\eta}_i \tag{30}$$

$$\text{Var}(\varepsilon|\mathcal{I}_i) = \hat{\Sigma}_{\varepsilon i} = (\Sigma_{\varepsilon}^{-1} + \Sigma_{\bar{\eta}_i}^{-1})^{-1} \tag{31}$$

where $\bar{\eta}_i = s_i^I \eta + s_i^E \eta_{p_i}$, and $\Sigma_{\bar{\eta}_i} = \Sigma_\eta + \Sigma_{\eta p}^E$. The definition and derivation of η_{p_i} and $\Sigma_{\eta p}^E$ are in Appendix B. $\Sigma_{\bar{\eta}_i}$ can be written as the sum of a diagonal matrix, Σ_η and a block matrix, $\Sigma_{\eta p}^E$. The only non-zero block in $\Sigma_{\eta p}^E$ is its bottom-right block.

The conditional moments of the full dividend vector d are:

$$E[d|\mathcal{I}_i] \equiv \hat{\mu}_{di} = \mu + \hat{\varepsilon}_i \tag{32}$$

$$\text{Var}(d|\mathcal{I}_i) \equiv \hat{\Sigma}_{di} = \rho\rho' + \hat{\Sigma}_{\varepsilon i} \tag{33}$$

Equilibrium. An equilibrium is a set of price functions p and investors' portfolio choices $\{q_i\}$, such that

1. Given the information set $\mathcal{I}_i$ and their subjective expectations about extensive flows $\tilde{m}_i$, each investor i chooses asset demand q_i to maximize expected utility (26), subject to the budget constraint (27).

2. Markets clear. The vector of price function p equates all investors' demand to the supply:

$$\int q_i di = \mathbf{1}. \tag{34}$$

Solving investor demand. By substituting the posterior beliefs (32) and (33) into the first order conditions, we obtain the expression of investors' optimal demand.

Proposition 3 (Optimal Demand)**.** *Investors' optimal demand is*

$$q_i = \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\left((\mathbf{1}-\delta_i)\,\mu + g(\tilde{m}_i) + \hat{\Sigma}_{\varepsilon i}^{-1}\Sigma_{\bar{\eta}_i}^{-1} s^I \eta - (\mathbf{1}-\delta_i)\, R_f p\right). \tag{35}$$

The detailed proof is in Appendix B.1. $\delta_i = \hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}i}^{-1} s^E \Sigma_\eta \hat{\Sigma}_{\varepsilon I}^{-1}$ captures the perceived price signal precision by investor i.[10] If δ_i is higher, the investor perceives that the price is more informative and she puts more weight on the price signal. The term $g(\bar{m}_i) = \delta_i \bar{\gamma}_I \hat{\Sigma}_{dI}(\mathbf{1} - \bar{m}_i)$ captures the perceived level shift of prices driven by the extensive flow.

Solving equilibrium prices. The objective market clearing condition can be written as:

$$\int q_i di = \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}(\hat{\mu}_{di} - pR_f)di = \mathbf{1} \tag{36}$$

Define a matrix Ω as the risk-tolerance-weighted average posterior precision of all investors:

$$\Omega = \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1} di$$

The term Ω captures the aggregate sensitivity of investors' static demands to expected excess returns, holding their information sets and subjective beliefs fixed. The aggregate demand can then be written as:

$$\int q_i di = \Omega(\mu - pR_f) + \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\hat{\varepsilon}_i di \tag{37}$$

The second term captures the demand component driven by information about the idiosyncratic shocks. We have the following proposition.

[10] δ_i is a block matrix in which only the bottom-right block is nonzero.

Proposition 4 (Equilibrium Prices)**.** *Solving for the equilibrium price p gives the final linear price function:*

$$p = \frac{1}{R_f}\left[\mu - (\Omega - \Omega_p)^{-1}(\mathbf{1} - \Omega_m(\tilde{m})) + (\Omega - \Omega_p)^{-1}\Omega_\eta \eta\right] \tag{38}$$

The proof and definitions of Ω_p, Ω_η, and $\Omega_m(\tilde{m})$ are in Appendix B.2. The equilibrium price is a function of expected payoffs (μ), the public signal about idiosyncratic shocks (η), and the aggregated effect of subjective expectations about extensive flows, $\Omega_m(\tilde{m})$.

5.2 Characteristics-Based Demand System

Following Koijen and Yogo (2019), we assume that expected dividends $\mu(n)$ and factor loadings $\rho(n)$ can be represented as functions of observable characteristics. Let $x(n)$ denote the observable characteristics and $\xi_i(n)$ denote the unobservable characteristics to the econometrician, we have

$$\mu(n) = \Phi'_\mu x_i(n) + \xi_\mu, \tag{39}$$

$$\rho(n) = \Phi'_\rho x_i(n) + \xi_\rho, \tag{40}$$

where Φ_μ and Φ_μ are vectors of coefficients on firm characteristics.

Rewriting investors' optimal demand in Proposition 3 into a form that separates own and cross-asset substitutions, and combining it with equations (39) and (40), give us the following corollary to Proposition 3.

Corollary 2. *Investor i's optimal demand in* (3) *admits a linear functional form*

$$q_i(n) = \gamma_i^{-1}\left(\hat{\sigma}^2_{\varepsilon I} + (\hat{\sigma}^{E,d}_{\varepsilon i})^2\right)^{-1}\Bigg[\underbrace{-(1-\delta_i^{E,d})R_f p(n)}_{\text{price term}} + \underbrace{\delta_i^{E,d}\bar{\gamma}_{iI}\sigma^2_{dI}(1-\tilde{m}_i(n))}_{\text{flow expectation term}} + \underbrace{\left((1-\delta_i^{E,d})\Phi_\mu - \Phi_\rho c_{i\rho}\right)x(n)}_{\text{characteristics term}} + \underbrace{\hat{\sigma}^2_{\varepsilon I}\sigma_\eta^{-2}s_i(n)\eta(n)}_{\text{signal term}} + \underbrace{\zeta_i}_{\text{investor specific scalar}}\Bigg]. \tag{41}$$

The proof and the definitions of $(\hat{\sigma}^{E,d}_{\varepsilon i})^2$, $\delta_i^{E,d}$, and ζ_i are in Appendix B.1. We separate the posterior variance of idiosyncratic shocks, $\hat{\Sigma}_{\varepsilon i}$, into a diagonal matrix, $\hat{\Sigma}^{E,d}_{\varepsilon i}$, and a block matrix with only the bottom-right block being non-zero. The non-zero part is a rank-1 matrix and captures the posterior covariance among assets induced by learning from prices. We obtain a similar separation of δ_i and the diagonal matrix is $\delta_i^{E,d}$. ζ_i is an investor-specific vector that captures the cross-asset substitution induced by investors' posterior beliefs.

Following Haddad et al. (2025), we interpret the "price term" as the "relative" elasticity, under the assumption of homogeneous substitution conditional on observables. The "characteristics term" is standard in the literature, except that our variant embeds the heterogeneity introduced by extensive demand. The innovation of our model is reflected in the "signal" and "flow expectation" terms. The "signal" term captures the information gap between extensive and intensive investors, through the effect of the payoff signal received only by intensive investors. The "flow expectation" term captures the effect of subjective expectations about flows.

5.3 Flow Expectations

A key component in our demand system is investor i's expectation about the contemporaneous extensive flow into stock n at time t, $\tilde{m}_{it}(n)$. The actual extensive flow $m_{it}(n)$ is not observable within period t at the time of the portfolio choice. In this section, we first discuss how subjective expectations about extensive flows affect investor demand and its price elasticity. And then, we model flow expectation $\tilde{m}_{it}(n)$ using several specifications of belief formation widely studied and used in the literature.

Extensive flow expectations and demand function. Investors' extensive flow expectations enter their demand function (41) through the "flow expectation" term below:

$$\underbrace{\gamma_i^{-1}}_{\text{Risk tolerance}} \underbrace{\left(\hat{\sigma}_{\varepsilon I}^2 + (\hat{\sigma}_{\varepsilon i}^{E,d})^2\right)^{-1}}_{\text{Posterior precision}} \underbrace{\delta_i^{E,d}\bar{\gamma}_{iI}\sigma_{dI}^2}_{\text{Kalman gain}}(1 - \tilde{m}_i(n))$$

The demand sensitivity to flow expectations is determined by risk tolerance, posterior precision, and the Kalman gain, which is the precision weight investors use to update their posterior beliefs about fundamentals using price signals in Bayesian updating. Flow expectations $\tilde{m}_i(n)$ enter the term negatively through $1 - \tilde{m}_i(n)$, which represents the perceived total intensive demand for the stock.

Intuitively, when investors believe that extensive inflows are large, they view a price increase as being less driven by intensive demand. Therefore, when investors interpret the informational content of prices, they effectively net out this uninformed component, which lowers their expectation about fundamentals and leads them to hold less of the stock. Together, there is a negative relation between extensive flow expectation and investor demand, as shown in the equation above.

We next describe three different specifications of flow expectations that we use in structural estimation.

Specification 1: Statistically optimal expectations.

We begin with a rational benchmark and assume investors form extensive flow expectations based on statistically optimal forecasts (Bianchi, Ludvigson, and Ma, 2022). The forecasting model, estimated using rolling windows on historical data, uses a wide range of available information to predict contemporaneous extensive flows.

Specifically, we estimate a forecasting model on a training sample of past observations:

$$m_{it}(n) \;=\; h_{\theta,t}(y_{i,t}(n), z_t) + \varepsilon_{i,t},$$

where $h_\theta(\cdot)$ is the LASSO prediction of extensive flow. $y_{i,t}$ denotes stock-level characteristics (e.g., size, valuation, profitability, growth, recent returns), past extensive flows, and let z_t denote market variables (e.g., market returns and risk-free rates). The predictor set we use consists of contemporaneous and up to four quarterly lags of each characteristic, as well as lagged quarterly market returns and the risk-free rate. We estimate the forecasting model's parameters, θ, using a rolling LASSO over a 5-year (20-quarter) window. For each prediction quarter t, the model is trained on the most recent 20 quarters. The penalty parameter is selected via a simple time-series validation step that reserves the most recent year within the window and chooses the ℓ_1 penalty minimizing validation MSE. The LASSO is refit on the full training window with the selected penalty and used to generate cross-sectional forecasts $h_{\theta,t}(y_{i,t}(n), z_t)$ for all stocks in quarter t.

We allow investors to have heterogeneous expectations based on the statistically optimal forecasts. We assume their flow expectations follow an affine function of the forecasts. In period t, the investor i's flow expectation is

$$\tilde{m}_{i,t}^{\text{stat}}(n) = a_{i,t}^{\text{stat}} + b_{i,t}^{\text{stat}} \widehat{h}_{\widehat{\theta}}(y_{i,t}, z_t).$$

The heterogeneity in expectations is captured by the coefficient on statistically optimal forecasts, $b_{i,t}^{\text{stat}}$, and an intercept, $a_{i,t}^{\text{stat}}$. This specification provides a benchmark in which investors process public information efficiently through a stable forecasting technology.

Specification 2: Extrapolative expectations.

We now turn to cases where investors form subjective beliefs and potentially deviate from the rational benchmark. We first consider extrapolative belief, a psychologically founded belief formation widely studied and documented in the literature (Barberis et al., 2015, 2018; Da, Huang, and Jin, 2021). This specification assumes investors extrapolate from past realizations of extensive flows. Let $m_{t-\ell}(n)$, $\ell = 1, 2, \ldots$, be lagged extensive flows observed at the time of decision. We model expectations as a weighted average of past realizations, following a constant-gain learning rule (Malmendier and Nagel, 2016):

$$\tilde{m}_t^{\text{extrp, avg}} = \sum_{\ell=1}^{L} \omega_\ell m_{t-\ell}, \qquad \omega_\ell \geq 0, \quad \sum_{\ell=1}^{L} \omega_\ell = 1,$$

where

$$\omega_\ell = \nu(1-\nu)^{\ell-1} / \sum_{k=1}^{L} \nu(1-\nu)^{k-1}.$$

ν is the constant gain parameter. We set $\nu = 0.018$ and $L = 20$.[11] We interpret these expectations as the average expectations across investors at time t. This specification emphasizes persistence and trend-chasing in flows: recent inflows raise perceived contemporaneous inflows, with the sensitivity governed by ν.

Similarly, we allow investors to have heterogeneous expectations based on the average extrap-

[11] Malmendier and Nagel (2016) show that $\nu = 0.018$ for quarterly data matches the average belief in survey data of inflation expectations, and lies within the range of estimates from household investment decisions in Malmendier and Nagel (2011).

olative expectations.

$$\tilde{m}_{i,t}^{\text{extrp}} = a_{i,t}^{\text{extrp}} + b_{i,t}^{\text{extrp}} \tilde{m}_t^{\text{extrp, avg}}$$

The heterogeneity in expectations is captured by the sensitivity to the average extrapolative expectations, $b_{i,t}^{\text{extrp}}$, and an intercept, $a_{i,t}^{\text{extrp}}$.

Specification 3: Subjective expectations proxied by social media attention.

Finally, we use the social media attention index (Cookson et al., 2024; Liu and Yin, 2025) as a proxy for investors' subjective expectations about contemporaneous extensive flows. We use the attention index shared by Cookson et al. (2024), which has a mean of zero ad standard deviation of one. The idea is that extensive flow is driven by other investors' contemporaneous participation decisions. Social media attention indicates whether other investors are paying attention to the stock. Therefore, the attention index captures investors' overall participation and serves as a natural proxy for the expectation of extensive flows.

Operationally, we treat social media attention as an observable signal, $h_t^{\text{att}}(n)$, that is increasing in the perceived strength of contemporaneous inflows into stock n. We allow investors to have heterogeneous expectations based on the social media attention:

$$\tilde{m}_{i,t}^{\text{att}}(n) \equiv a_{i,t}^{\text{att}} + b_{i,t}^{\text{att}} h_t^{\text{att}}(n),$$

with $a_{i,t}^{\text{att}}$ and $b_{i,t}^{\text{att}}$ absorbing scale differences relative to the flow measures and capturing heterogeneity in investor expectations.

5.4 Estimating the Asset Demand System

We now derive the estimation equation by collecting terms and rearranging (41):

$$\begin{aligned} q_{i,t}(n) =& \gamma_{i,t}\mathbf{1}_{i,t}^{E}(n)\tilde{m}_{i,t}(n) + \left[\zeta_{0,i,t} + \zeta_{1,i,t}\mathbf{1}_{i,t}^{E}(n)\right] \cdot p_t(n) \\ &+ \left[\beta_{0,i,t} + \beta_{1,i,t}\mathbf{1}_{i,t}^{E}(n)\right] \cdot x_t(n) + \alpha_{i,t} + \alpha_t(n)\mathbf{1}_{i,t}^{E}(n) + u_{i,t}(n), \end{aligned} \tag{42}$$

where $\mathbf{1}_{i,t}^{E}(n)$ is a dummy variable which takes the value of one if this demand is at the extensive margins, and zero otherwise. $\tilde{m}_{i,t}(n)$ is investor i's expectation of extensive flow into stock n at time

t. $x_t(n)$ is the vector of characteristics of stock n at time t. In our cross-sectional estimation (and thus omitting subscript t), $\alpha_{i,t}$ represents the investor fixed effects and $\alpha_t(n)$ represents the stock fixed effects. $u_{i,t}(n)$ represents the uncorrelated residual demand shocks. They are assumed to be conditionally mean-independent across investors.

For the demand at the intensive margins ($\mathbf{1}^E_{i,t}(n) = 0$), the price elasticities reduce to $\zeta_{0,i,t}$. $\zeta_{1,i,t}$ captures the wedge in elasticities between demand at extensive and intensive margins. A $\zeta_{1,i,t}$ larger than zero means the price elasticities of extensive demand are larger than those of intensive demand. Similarly, the coefficients on stock characteristics reduce to $\beta_{0,i,t}$ for intensive demand, and $\beta_{1,i,t}$ shows the difference in the coefficients on stock characteristics across the two margins.

Granular instrumental variables (GIV). The main identification challenge to estimate the demand equation above is that demand and prices are jointly determined in equilibrium. Price $p_t(n)$ is endogenous and is correlated with investor i's demand shock. Therefore, we need an instrument for prices. This type of instrument is difficult to find. One workhorse approach in the literature is to search for idiosyncratic demand shocks from some investors to instrument for prices in another investor's demand equation (Gabaix and Koijen, 2024). We follow Chaudhry and Li (2025) to use the optimal granular instrumental variables (GIV) methodology in Chaudhary, Fu, and Zhou (2024) to estimate price elasticities and other demand coefficients.

Our identification assumption is the conditional independence of residual demand shocks across investors:

$$\mathbb{E}\left[u_{i,t}(n)u_{j,t}(n)\right] = 0, \quad \forall i \neq j, \forall t, \forall n,$$

$u_{i,t}(n)$ and $u_{j,t}(n)$ are residual demand shocks of different investors for stock n. Based on this assumption, we construct the GIV for prices using the combination of other investors' residual demand shocks: $z_{i,t}(n) = \sum_{j \neq i} u_{j,t}(n)$. Given this construction of GIV, we obtain the equivalent moment conditions:

$$\mathbb{E}\left[u_{i,t}(n)z_{i,t}(n)\right] = 0. \tag{43}$$

We estimate the demand system cross-section by cross-section. Investors' residual demand shocks are constructed based on the demand elasticities, which are the variables that need to be estimated. Following the procedure in Chaudhary, Fu, and Zhou (2024), we jointly estimated the residual demand shocks and the price elasticities.

Instead of filling zero holdings of all stocks in the investment universe defined above, we only focus on one specific type of "zero holdings": extensive out, i.e., stocks that were held in the previous period but are fully liquidated in the current period. This approach allows us to better capture investors' extensive margin decisions as discussed in the previous sections. Thus, we construct the estimation sample by including all positive holdings and extensive out observations for each investor in each quarter.[12]

Stock characteristics. The characteristics for each stock other than the market equity that we include in the estimation include: which includes log book equity, profitability, investment, dividend ratio, and market beta. Book equity represents firm size. We follow Fama and French (2015) to construct the characteristics of profitability and investment. Profitability is measured as operating profit scaled by book equity. Operating profit to equity is measured as revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense. Investment is calculated as the yearly change in total assets. Dividend ratio is measured as the sum of total dividends over the past year divided by book equity. Market beta is estimated from the 60-month rolling regression of excess stock returns on excess market returns.

6 Structural Estimation Results and Counterfactuals

We structurally estimate the demand system under different specifications of extensive flow expectations specified in Section 5.3. We show the estimated results and verify model predictions. Finally, we compute a counterfactual to quantify the impact of extensive demand on stock returns, volatility, and price informativeness.

[12]Koijen and Yogo (2019) compare linear GMM estimates (excluding zero holdings) to non-linear GMM estimates (including zero holdings) and document that the latter tend to yield more negative demand elasticities, especially for small institutions. Our notion of "extensive out" focuses on stocks that were held in the previous period but are fully exited in the current period, which is conceptually different from including all zero positions in the estimation sample. As a result, the level of our estimated elasticities is not directly comparable to their non-linear GMM estimates.

6.1 Estimated Price Elasticities

Table 6 reports descriptive statistics for the estimated price elasticities of extensive and intensive demand under different specifications of extensive flow expectations. Under the statistically optimal expectations specification, the mean and various percentiles of the price elasticity estimates for extensive demand are higher than those for intensive demand. The mean price elasticity for extensive demand is -0.34, compared with -0.43 for intensive demand. We find similar results for the two subjective expectations specifications. Under both extrapolative expectations and subjective expectations proxied by social media attention, the mean and all reported percentiles of the estimated price elasticities for extensive demand are higher than those for intensive demand. These findings confirm that price elasticities of extensive demand are higher (less negative) than those of intensive demand and that this pattern is robust across different specifications of extensive flow expectations.

[Insert Table 6 Here.]

Figure 7 Panel (a) plots the overall distributions of estimated price elasticities for extensive and intensive demand, under the statistically optimal expectations specification. The distribution for extensive demand is shifted to the right relative to the distribution for intensive demand. Another salient feature of the graph is the non-trivial mass of estimates with positive price elasticities. The fraction of positive elasticities is visibly larger for extensive demand than for intensive demand, indicating that upward-sloping demand is more prevalent along the extensive margin. Figure 7 Panel (b) shows the time series of yearly average elasticities, estimated under the same flow expectation specification. The panel reveals that the average extensive demand elasticity is consistently higher than the average intensive demand elasticity. This pattern holds almost universally across the entire sample period. The magnitude of the difference between the two elasticities varies over time. In several years, the gap in magnitudes is sizable.

[Insert Figure 7 Here.]

These findings are consistent with the implications of our model. Because investors have less information about assets at the extensive margin and therefore rely more heavily on prices, learning from prices attenuates the magnitude of their price elasticity. As a result, extensive demand is less downward sloping than intensive demand. Furthermore, the model predicts that extensive demand elasticities can become positive. This occurs when investors' subjective expectations about extensive flows decrease as prices rise. If this is the case, investors overreact to the perceived price signals, which can push the elasticity into the positive regime. Our estimates have a non-trivial mass for extensive demand elasticities, aligning well with this theoretical prediction. Our findings differ from Koijen and Yogo (2019) and Koijen, Richmond, and Yogo (2024), who restrict price elasticities to be non-positive during estimation. We show, both theoretically and empirically, that the price elasticity of extensive demand can be positive. As illustrated above, this occurs when investors believe the extensive flow is negatively correlated with prices.

Table 7 shows the descriptive statistics of estimated price elasticities by active share quintile. In general, more passive investors (those with lower active share) tend to have price elasticities closer to zero. For the most passive group (quintile 1), the extensive demand elasticity is −0.280, while the intensive demand elasticity is −0.219. However, the trend beyond the first quintile differs by demand type. For extensive demand, the elasticity becomes more negative after the first quintile. After this initial drop, the elasticity remains relatively flat, ranging from −0.340 to −0.376 across quintiles 2 through 5. This flatness suggests a similar level of information scarcity for investors who trade at the extensive margin. Conversely, the price elasticity of intensive demand decreases significantly and monotonically with activeness. The elasticity falls from −0.219 for the most passive group to −0.638 for the most active group. This result supports the interpretation that more active investors are likely to be more informed, leading to a more negative price elasticity for intensive demand. Most importantly, the wedge between extensive and intensive elasticities becomes more salient from passive to active investors (from roughly the same to a gap of 0.3). Such a pattern is consistent with our information-based mechanism—active investors actively trade on their information, amplifying the information gap between extensive and intensive demand.

[Insert Table 7 Here.]

6.2 Demand Sensitivity to Extensive Flow Expectations

Table 6 also presents the estimates of demand sensitivity to extensive flow expectations under different specifications. We begin with the rational benchmark in which investors form extensive flow expectations based on statistically optimal forecasts (Panel (a)). The mean of estimates is

−0.076, which indicates that if the extensive flow expectation goes up by one percentage point, investors reduce their demand by 0.076 percentage points. This negative relation is consistent with our mechanism of investor learning from prices. When investors believe that extensive inflows are driving prices up, they do not update their beliefs about fundamentals, or they update them downward. As a result, they reduce their demand for the stock. We find this negative relation for a large proportion of the estimates, but it also delivers a subset of sensitivity estimates with a positive sign. The 75% percentile of the estimates is 0.409. This pattern suggests that we may misspecify flow expectations—investors may deviate from the statistically optimal forecasting rule and instead form beliefs that are not fully rational.

Motivated by this evidence, we consider two alternative specifications of expectation formation. First, we draw on the behavioral finance literature and adopt the widely studied extrapolative beliefs. We find that estimation results based on extrapolative expectations are most consistent with our framework. The mean of the estimates is −0.812, indicating that a one-percentage-point increase in extensive flow expectations is associated with a 0.8-percentage-point decline in investor demand. The magnitude of the mean of the estimates is much larger than that under statistically optimal expectations. All reported percentiles are negative. Together, these findings suggest that investors overreact to recent realized flows when forming expectations.

We also take a data-driven approach and incorporate additional information plausibly relevant for belief formation. Specifically, we use social media attention as a proxy for investors' perceptions of market participation and interest (Cookson et al., 2024; Liu and Yin, 2025). The idea is that heightened discussion of a stock on social media may lead investors to infer that other market participants are entering the stock, generating extensive inflows. The estimation results under this specification are broadly consistent with our framework. The mean of the estimates is −0.719, which is close to that under extrapolative expectations. However, a non-trivial proportion of estimates remain positive. The 75% percentile of the estimates is 0.804. Overall, among the specifications we consider, extrapolative expectations appear most consistent with our framework.

6.3 Counterfactual: Reallocating Capital from Intensive to Extensive Investors

We conduct a counterfactual analysis to quantify the effects of an increase in investor demand at the extensive margin. Our counterfactual analysis is based on the estimation under the specification of statistical optimal flow expectations.

In the analysis, we reallocate capital from intensive to extensive investors. We first classify investors by their share of extensive demand. For each investor i at time t, we calculate their extensive demand proportion as the number of positive extensive positions divided by the total number of positions. We then average this proportion across time for each investor to generate a single, time-invariant measure. Based on this measure, we divide investors into five quintiles and label the top quintile as "extensive investors."

Next, we implement a budget-neutral reallocation. We increase the total AUM of extensive investors by 10% by reallocating capital from other investors. The reallocation is financed by an AUM haircut on investors in the other four quintiles. Specifically, we reduce the AUM of each other investor by a fixed percentage such that the total reduction in AUM across all other investors equals 10% of the total AUM of extensive investors. The reduced AUM from intensive investors is then proportionally allocated to extensive investors based on their initial AUM. This process captures a market-wide shift in capital from intensive to extensive demand.

After the reallocation, we recalculate the equilibrium prices based on market clearing conditions. We assume firm characteristics and shares outstanding do not change. We also hold investors' demand coefficients and extensive flow expectations unchanged. Finally, we compute changes in stock returns, volatility, and price informativeness using counterfactual prices.

We choose to use the statistically optimal flow expectation here because it provides a natural, rational, and robust benchmark that is transparent and easy to interpret. It shuts down the channel of biased beliefs, so that the counterfactual isolates the effect of reallocating capital toward investors with higher extensive demand, the focus of this counterfactual analysis. Using alternative behavioral expectation specifications is more difficult in this setting. A large reallocation could change the belief-formation process itself, so treating behavioral expectations as invariant under the counterfactual may be less realistic. Implementing such counterfactuals would require direct investor–stock-level data on flow expectations, which we typically do not observe. Furthermore, how beliefs adjust under the counterfactual is not straightforward to infer from the data. Nevertheless, we implement these specifications, and the results are available upon request.

6.3.1 Impact on Return and Volatility

Panel (a) of Table 8 shows results on return and volatility. Overall, reallocating capital from intensive to extensive demand leads to significant increases in average stock returns and in return volatility. The average annualized return increases by 3.3 percentage points. The average return volatility across stocks increases by 3.7 percentage points. These results are consistent with our empirical findings in Section 3.2 that the percentage composition of extensive flow is positively correlated with stock return and volatility.

[Insert Table 8 Here.]

6.3.2 Impact on Price Informativeness

We next examine the impact of the reallocation on price informativeness. We consider two measures of price informativeness: stock-specific price informativeness (Dávila and Parlatore, 2025) and market price informativeness (Bai, Philippon, and Savov, 2016).

Stock-specific price informativeness. We first calculate stock-specific price informativeness based on actual prices and counterfactual prices using equations (1) to (3). Panel (b) of Table 8 shows the comparison between actual and counterfactual price informativeness. Reallocating capital from intensive to extensive demand leads to a significant increase in price informativeness. The average price informativeness across stocks increases by 0.4 percentage points in the full sample. However, the reallocation has heterogeneous impacts across stocks. For stocks with already high price informativeness (above-median actual price informativeness), the reallocation leads to a further increase in price informativeness of 1.0 percentage points. In contrast, for stocks with low actual price informativeness (below-median), the reallocation leads to a decrease in price informativeness of 0.2 percentage points. These findings are consistent with our empirical results in Section 3.2, which show that extensive flow is positively correlated with price informativeness for stocks with high actual price informativeness but negatively correlated for stocks with low actual price informativeness.

This stock-specific price informativeness captures how changes in prices reflect changes in future earnings at the individual stock level. When prices are very informative, extensive investors can better learn information from prices. This learning limits the role of their subjective beliefs about extensive flow, makes their demand move more closely with future fundamentals, and thus increases price informativeness. By contrast, when prices are noisy, extensive investors place more weight on their subjective beliefs about extensive flow. This causes their demand to deviate further from future fundamentals and to distort prices more.

Market price informativeness. Following Bai, Philippon, and Savov (2016), we estimate the market-wide price informativeness by first running cross-sectional regressions of future earnings on current market prices:

$$\frac{E_{i,t+h}}{A_{i,t}} = a_{t,h} + b_{t,h} \log\left(\frac{M_{i,t}}{A_{i,t}}\right) + c_{t,h}\left(\frac{E_{i,t}}{A_{i,t}}\right) + d^s_{t,h}\mathbf{1}^s_{i,t} + \epsilon_{i,t,h}, \tag{44}$$

where h represents different horizons of future earnings, $\mathbf{1}^s_{i,t}$ is the indicator of sector (one-digit SIC code). The measure of price informativeness is calculated as

$$\tau_m = b_{t,h} \times \sigma_t(\log(M/A)), \tag{45}$$

where $\sigma_t(\log(M/A))$ is the cross-sectional standard deviation of $\log(M/A)$ at time t.

[Insert Figure 8 Here.]

We compute market price informativeness based on actual and counterfactual prices. Figure 8 plots the time series of actual and counterfactual market price informativeness for different horizons. The graphs show that reallocating capital from intensive to extensive demand leads to a slight decline in market price informativeness. However, this does not contradict our previous findings on stock-specific price informativeness. Market price informativeness measures, in the cross-section, how well the current price level predicts future earnings. Our findings suggest that increases in extensive demand cause distortions in prices that reduce their cross-sectional predictive power for future earnings.

7 Conclusion

This paper has shown that decomposing investor demand into its extensive and intensive margins is both conceptually important and empirically relevant. Using institutional and retail holdings, we document that extensive flows—entry and exit trades—are economically large, systematically shaped by attention and past returns, and carry distinct asset pricing implications relative to intensive margin adjustments within existing positions. A higher composition of extensive flows predicts higher contemporaneous returns and volatility, and its relation to price informativeness is state-dependent: it is informationally beneficial when prices are already informative and detrimental when they are not. Dynamic evidence from local projections further highlights a strong two-way interaction between extensive flows and returns, with large short-run effects that reverse over longer horizons. Together, these results point to flow composition, rather than flow levels alone, as a key factor in understanding the relation between investor demand and asset prices.

Guided by these facts, we develop a Grossman–Stiglitz–style model with unobserved extensive flows and subjective beliefs about them. We embed it in a characteristics-based demand system that allows for distinct elasticities at the extensive and intensive margins. Structural estimates confirm the model's core predictions: extensive demand is substantially less downward-sloping, and often upward-sloping, than intensive demand, reflecting investors' learning from prices with potentially inaccurate beliefs about unobserved extensive flows. Counterfactual reallocation of capital from intensive to extensive investors raises average returns and volatility and has heterogeneous effects on price informativeness across stocks, improving it where prices are already informative and degrading it where they are not. These findings suggest that who trades at which margin is central for market dynamics, and they open avenues for future work on how regulation, market microstructure, and information that alter the composition of extensive versus intensive investors feed back into volatility, comovement, and the informational efficiency of asset prices.

References

Ali, A., L.-S. Hwang, and M. A. Trombley. 2003. Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics* 69:355–73.

An, Y., Y. Su, and C. Wang. 2025. Quantity, risk, and return. *Available at SSRN 4098609* .

Audoly, R., R. McGee, S. Ocampo-Diaz, and G. Paz-Pardo. 2025. A practitioner's note on the shapley-owen-shorrocks decomposition. *FRB of New York Staff Report* .

Bai, J., T. Philippon, and A. Savov. 2016. Have financial markets become more informative? *Journal of Financial Economics* 122:625–54.

Banerjee, S., and B. Green. 2015. Signal or noise? uncertainty and learning about whether other traders are informed. *Journal of Financial Economics* 117:398–423.

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

———. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies* 21:785–818.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer. 2015. X-capm: An extrapolative capital asset pricing model. *Journal of financial economics* 115:1–24.

———. 2018. Extrapolation and bubbles. *Journal of Financial Economics* 129:203–27.

Bastianello, F., and P. Fontanier. 2025a. Expectations and learning from prices. *Review of Economic Studies* 92:1341–74.

———. 2025b. Partial equilibrium thinking, extrapolation, and bubbles .

Bianchi, F., S. C. Ludvigson, and S. Ma. 2022. Belief distortions and macroeconomic fluctuations. *American Economic Review* 112:2269–315.

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

Bretscher, L., L. Schmid, I. Sen, and V. Sharma. 2025. Institutional corporate bond pricing. *The Review of Financial Studies* hhaf067.

Campbell, J. Y., T. Ramadorai, and A. Schwartz. 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of financial economics* 92:66–91.

Chaudhary, M., Z. Fu, and H. Zhou. 2024. Anatomy of the treasury market: Who moves yields? *Available at SSRN* .

Chaudhry, A., and J. Li. 2025. Endogenous elasticities: Price multipliers are smaller for larger demand shocks. *Available at SSRN* .

Chen, H., G. Noronha, and V. Singal. 2004. The price response to s&p 500 index additions and deletions: Evidence of asymmetry and a new explanation. *The Journal of Finance* 59:1901–30.

Choi, J., X. Tian, Y. Wu, and M. Kargar. 2025. Investor demand, firm investment, and capital misallocation. *Journal of Financial Economics* 168:104039–.

Chordia, T., R. Roll, and A. Subrahmanyam. 2002. Order imbalance, liquidity, and market returns. *Journal of Financial Economics* 65:111–30.

Cookson, J. A., R. Lu, W. Mullins, and M. Niessner. 2024. The social signal. *Journal of Financial Economics* 158:103870–.

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

Cremers, K. M., and A. Petajisto. 2009. How active is your fund manager? a new measure that predicts performance. *The Review of Financial Studies* 22:3329–65.

Da, Z., X. Huang, and L. J. Jin. 2021. Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics* 140:175–96.

Darmouni, O., K. Siani, and K. Xiao. 2022. Nonbank fragility in credit markets: Evidence from a two-layer asset demand system. *Available at SSRN* 4288695.

Dávila, E., and C. Parlatore. 2023. Volatility and informativeness. *Journal of Financial Economics* 147:550–72.

———. 2025. Identifying price informativeness. *The Review of Financial Studies* hhaf051.

Davis, C., M. Kargar, and J. Li. 2025. Why do portfolio choice models predict inelastic demand? *Journal of Financial Economics* 172:104096–.

De Silva, T., and D. Thesmar. 2024. Noise in expectations: Evidence from analyst forecasts. *The Review of Financial Studies* 37:1494–537.

Dou, W. W., L. Kogan, and W. Wu. 2022. Common fund flows: Flow hedging and factor pricing. Working Paper, National Bureau of Economic Research.

Enke, B., and T. Graeber. 2023. Cognitive uncertainty. *The Quarterly Journal of Economics* 138:2021–67.

Eyster, E., M. Rabin, and D. Vayanos. 2019. Financial markets where traders neglect the informational content of prices. *The Journal of Finance* 74:371–99.

Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1–22.

Frazzini, A., and O. A. Lamont. 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics* 88:299–322.

Fuchs, W., S. Fukuda, and D. Neuhann. 2023. Demand-system asset pricing: Theoretical foundations. *Available at SSRN 4672473* .

Gabaix, X., and R. S. Koijen. 2021. In search of the origins of financial fluctuations: The inelastic markets hypothesis. Working Paper, National Bureau of Economic Research.

———. 2024. Granular instrumental variables. *Journal of Political Economy* 132:2274–303.

Gabaix, X., R. S. Koijen, F. Mainardi, S. Oh, and M. Yogo. 2025. Asset demand of us households. *Available at SSRN 4251972* .

Graves, D. 2025. What lies beneath zero: Censoring, demand estimation, and hidden beliefs. *Available at SSRN* .

Greenwood, R. 2005. Short-and long-term demand curves for stocks: theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics* 75:607–49.

Greenwood, R., and D. Thesmar. 2011. Stock price fragility. *Journal of Financial Economics* 102:471–90.

Grossman, S. J., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *The American Economic Review* 70:393–408.

Haddad, V., Z. He, P. Huebner, P. Kondor, and E. Loualiche. 2025. Causal inference for asset pricing. *Available at SSRN 5187305* .

Haddad, V., P. Huebner, and E. Loualiche. 2025. How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. *American Economic Review* 115:975–1018.

Harris, L., and E. Gurel. 1986. Price and volume effects associated with changes in the s&p 500 list: New evidence for the existence of price pressures. *the Journal of Finance* 41:815–29.

Hartzmark, S. M., and D. H. Solomon. 2025. Market-wide predictable price pressure. *American Economic Review* 115:3171–213.

Hartzmark, S. M., and A. B. Sussman. 2025. Price agnostic demand. *Available at SSRN* .

Hasbrouck, J. 1991. Measuring the information content of stock trades. *The Journal of Finance* 46:179–207.

He, Z., P. Kondor, and J. S. Li. 2025. Demand elasticity in dynamic asset pricing. *Available at SSRN 5297190* .

Hellwig, M. F. 1980. On the aggregation of information in competitive markets. *Journal of economic theory* 22:477–98.

Hendershott, T., and A. J. Menkveld. 2014. Price pressures. *Journal of Financial Economics* 114:405–23.

Huebner, P. 2023. The making of momentum: A demand-system perspective. In *Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting*.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp. 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84:571–626.

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

Kaul, A., V. Mehrotra, and R. Morck. 2000. Demand curves for stocks do slope down: New evidence from an index weights adjustment. *The Journal of Finance* 55:893–912.

Kelley, E. K., and P. C. Tetlock. 2013. How wise are crowds? insights from retail orders and stock returns. *The Journal of Finance* 68:1229–65.

Koijen, R. S., R. J. Richmond, and M. Yogo. 2024. Which investors matter for equity valuations and expected returns? *Review of Economic Studies* 91:2387–424.

Koijen, R. S., and M. Yogo. 2019. A demand system approach to asset pricing. *Journal of Political Economy* 127:1475–515.

———. 2025. On the theory and econometrics of (demand system) asset pricing. *Available at SSRN 5274709* .

Liu, Y., and X. Yin. 2025. Sentiments about others in the stock markets. *Available at SSRN 5760202* .

Llorente, G., R. Michaely, G. Saar, and J. Wang. 2002. Dynamic volume-return relation of individual stocks. *The Review of Financial Studies* 15:1005–47.

Lou, D. 2012. A flow-based explanation for return predictability. *The Review of Financial Studies* 25:3457–89.

Malmendier, U., and S. Nagel. 2011. Depression babies: do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics* 126:373–416.

———. 2016. Learning from inflation experiences. *The Quarterly Journal of Economics* 131:53–87.

Mendel, B., and A. Shleifer. 2012. Chasing noise. *Journal of financial economics* 104:303–20.

Mondria, J., X. Vives, and L. Yang. 2022. Costly interpretation of asset prices. *Management Science* 68:52–74.

Patton, A. J., and A. Timmermann. 2010. Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics* 57:803–20.

Pavlova, A., and T. Sikorskaya. 2023. Benchmarking intensity. *The Review of Financial Studies* 36:859–903.

Rendleman Jr, R. J., C. P. Jones, and H. A. Latane. 1982. Empirical anomalies based on unexpected earnings and the importance of risk adjustments. *Journal of Financial Economics* 10:269–87.

Schmidt-Engelbertz, P., and K. Vasudevan. 2025. Speculating on higher-order beliefs. *The Review of Financial Studies* hhaf019.

Shleifer, A. 1986. Do demand curves for stocks slope down? *The Journal of Finance* 41:579–90.

Shorrocks, A. F., et al. 2013. Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *Journal of Economic Inequality* 11:99–126.

van Binsbergen, J. H., B. David, and C. C. Opp. 2025. How (not) to identify demand elasticities in dynamic asset markets. *Available at SSRN 5362386* .

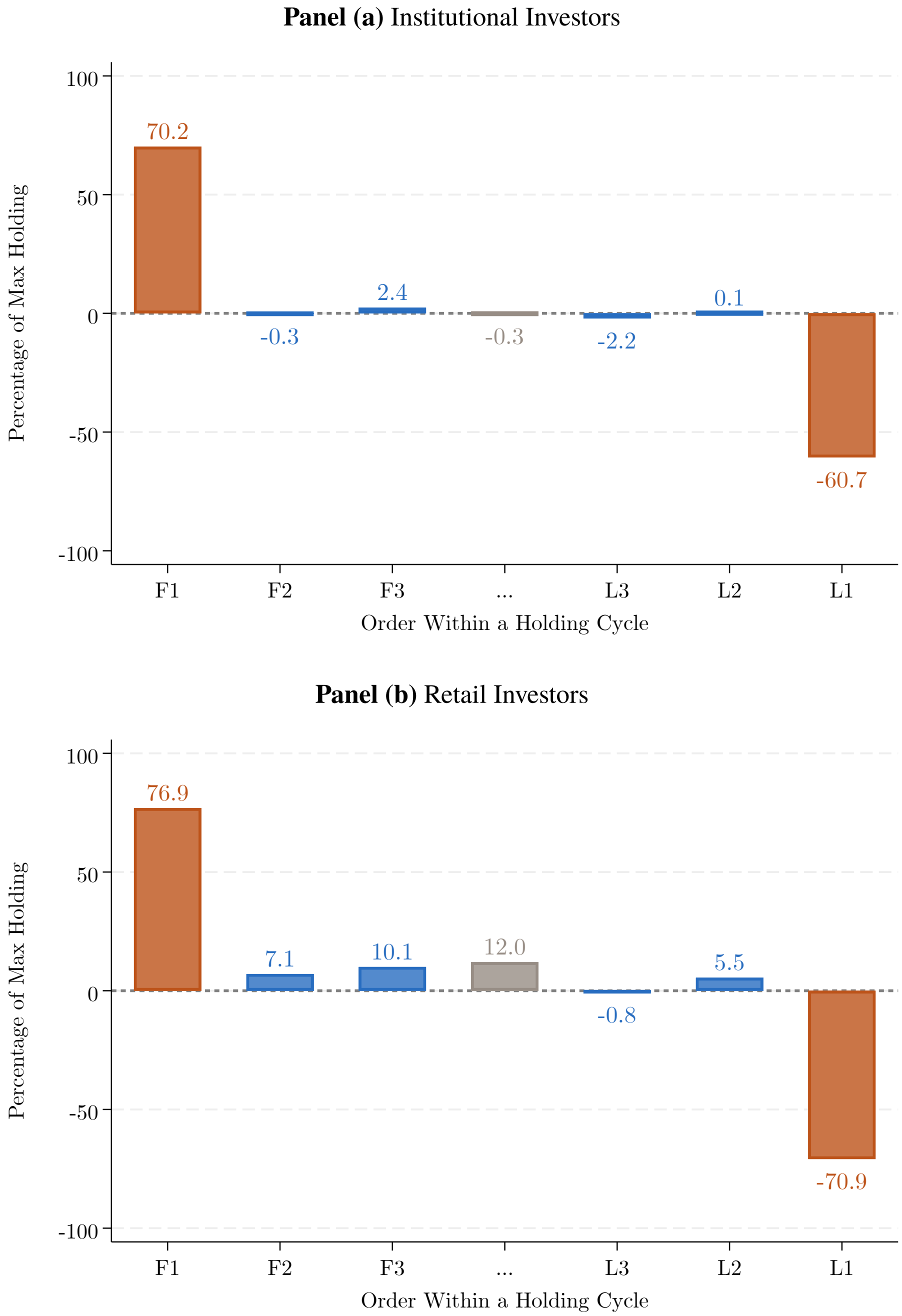
van der Beck, P. 2022. Short-versus long-run demand elasticities in asset pricing. *Asset Pricing (May 16, 2022). Swiss Finance Institute Research Paper* .

Van Nieuwerburgh, S., and L. Veldkamp. 2010. Information acquisition and under-diversification. *The Review of Economic Studies* 77:779–805.

Vayanos, D., and P. Woolley. 2011. Fund flows and asset prices: A baseline model .

Figures

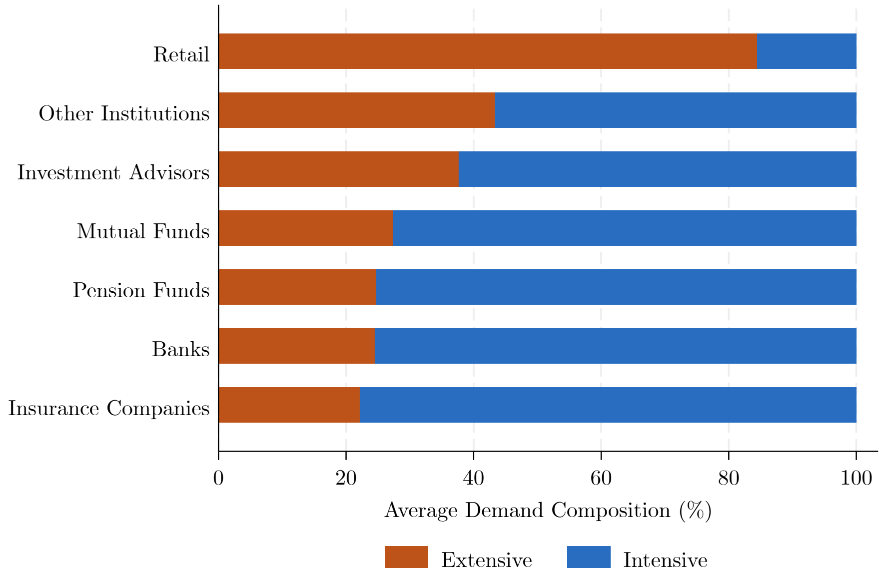
Figure 1. Flow Distribution Within a Holding Cycle



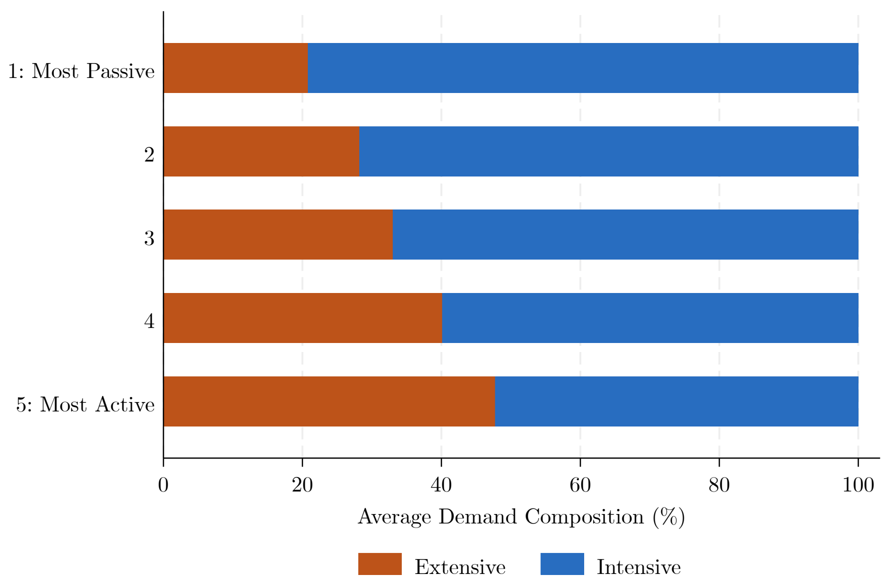
Notes. This figure plots the distribution of flows as a percentage of the maximum position within an average holding cycle, using the institutional investor sample in Panel (a) and the retail investor sample in Panel (b). F1–F3 denote the first, second, and third periods of the cycle, while L1–L3 correspond to the last, second-to-last, and third-to-last periods. Red bars indicate extensive flows, whereas blue and gray bars represent intensive flows.

Figure 2. Compositions of Extensive and Intensive Flows

Panel (a) By Investor Type



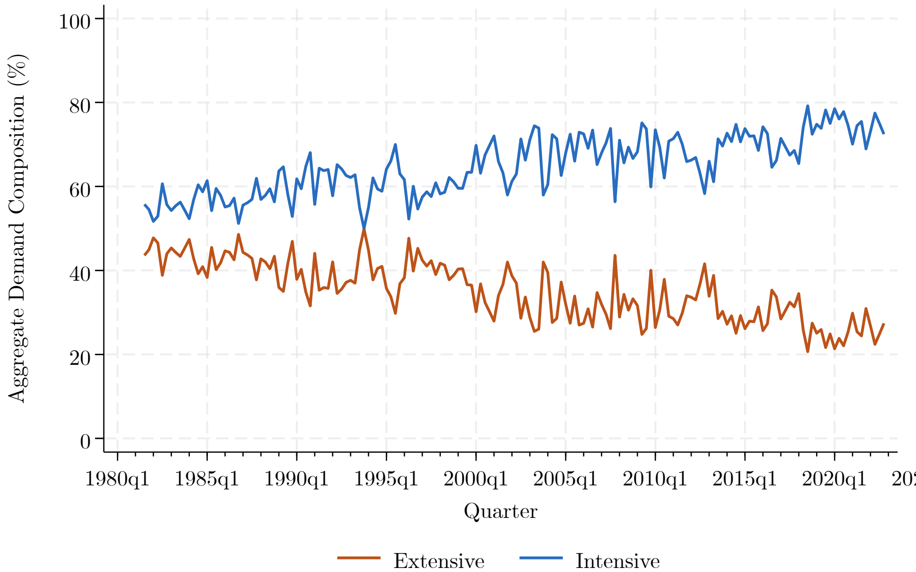
Panel (b) By Active Share Quintile



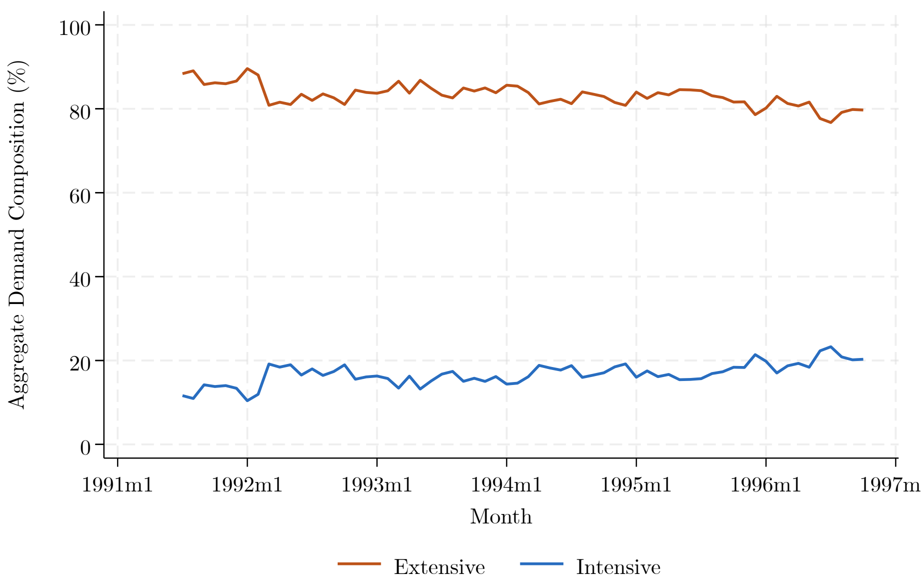
Notes. This figure plots the percentage composition of extensive flows (the sum of extensive in and out) and intensive flows (the sum of intensive in and out), by investor type in Panel (a) and by active-share quintile in Panel (b).

Figure 3. Compositions of Aggregate Flows

Panel (a) Institutional Investors



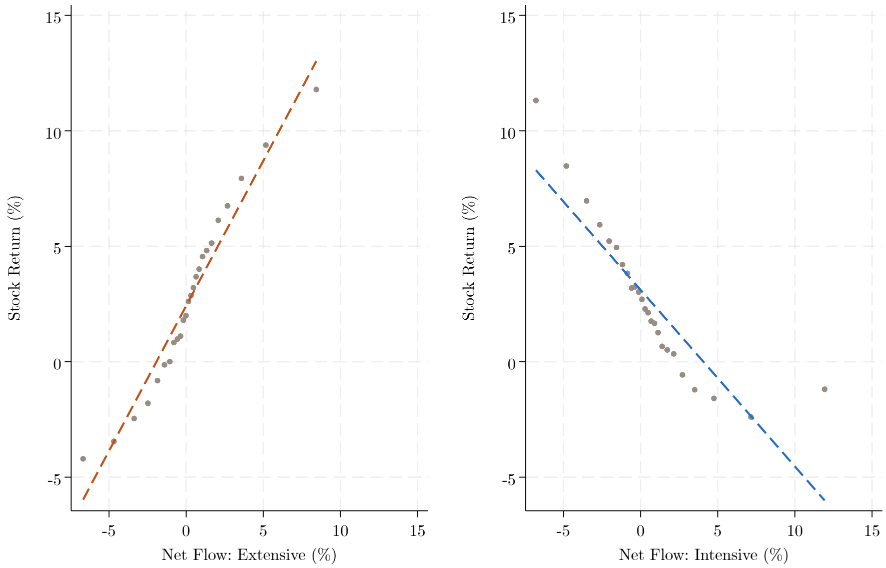
Panel (b) Retail Investors



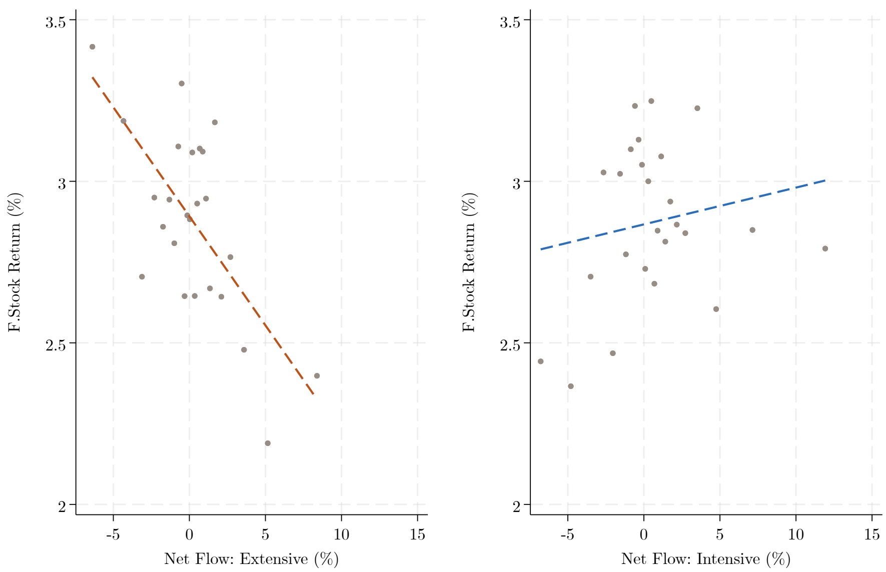
Notes. This figure plots the time series of the percentage composition of aggregate extensive flows (the sum of extensive in and out) and intensive flows (the sum of intensive in and out), using the institutional investor sample in Panel (a) and the retail investor sample in Panel (b).

Figure 4. Relationship Between Net Flows and Stock Returns

Panel (a) Contemporaneous Stock Return



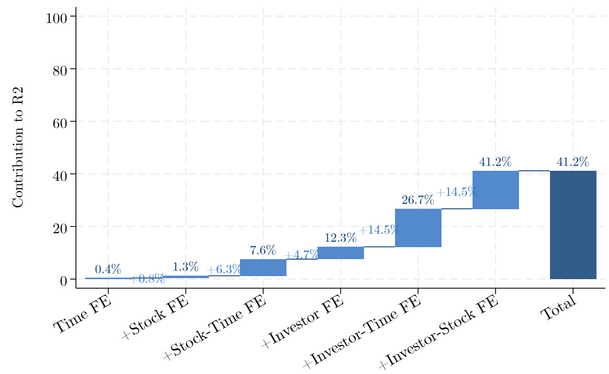
Panel (b) Future Stock Return



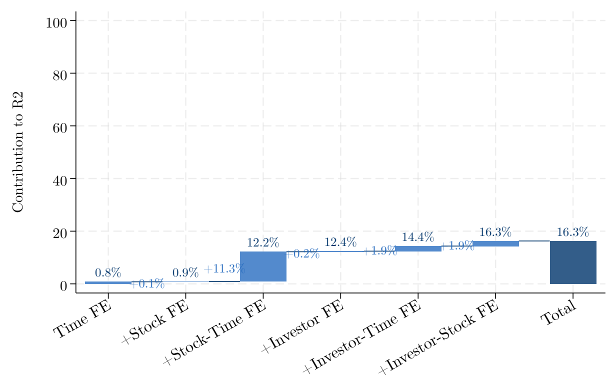
Notes. This figure plots binned scatterplots of the relationship between quarterly net flows and stock returns, showing contemporaneous returns in Panel (a) and next-quarter returns in Panel (b). The underlying regression specification (6) includes stock fixed effects, quarter fixed effects, as well as controls for standardized unexpected earnings (SUE), size, book-to-market ratio, lagged return, and profitability.

Figure 5. Decomposing Variations of Extensive and Intensive Flows

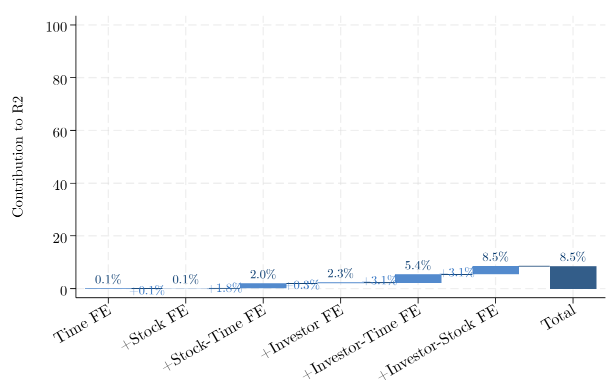
Panel (a) Type: Extensive or Intensive



Panel (b) Dollar Value: Extensive



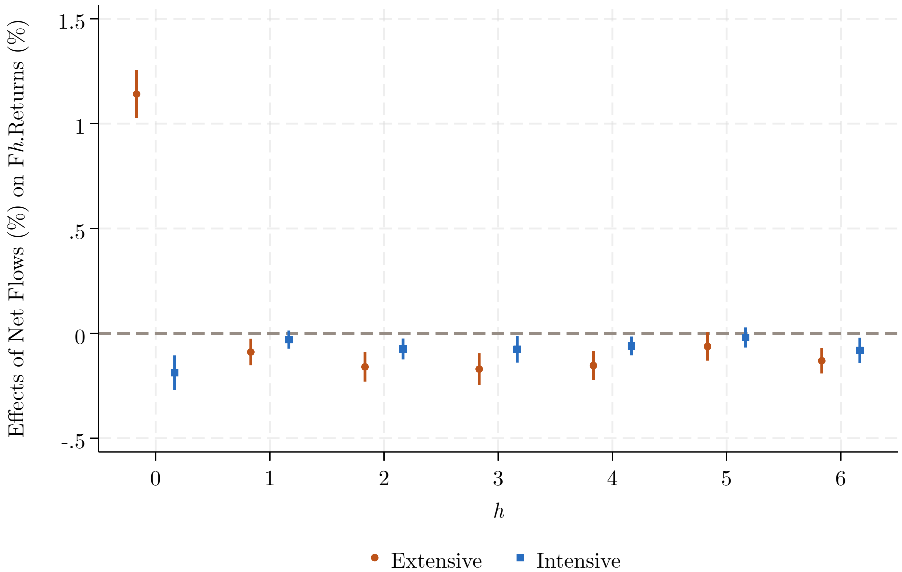
Panel (c) Dollar Value: Intensive



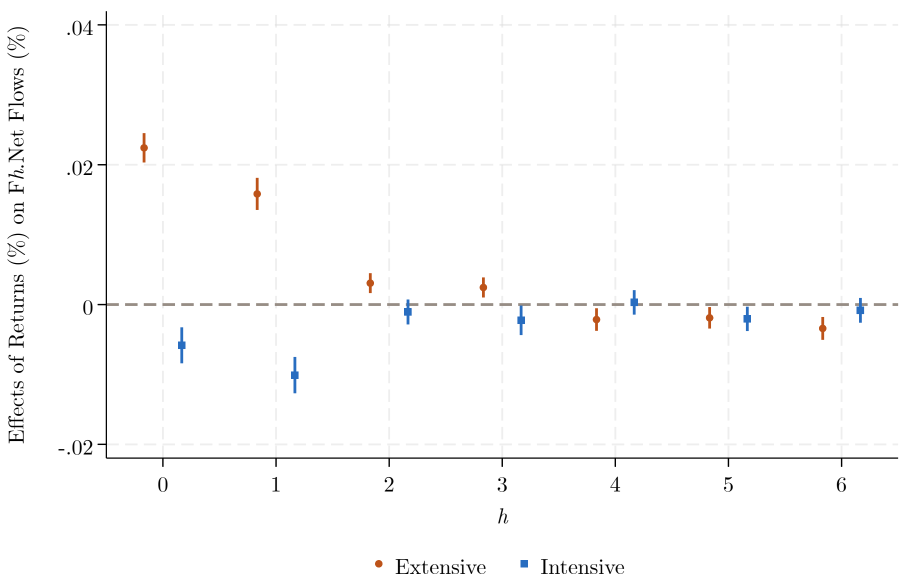
Notes. This figure plots the share of variation in flow type and dollar value explained by different fixed-effect specifications. Panel (a) decomposes the variation in whether a flow is extensive or intensive, while Panels (b) and (c) show the corresponding decompositions for the dollar value of extensive and intensive flows. The figure is based on a Shapley–Owen decomposition using an investor–stock–time panel, considering time, investor, stock, investor–time, stock–time, and investor–stock fixed effects.

Figure 6. Local Projection Orthogonalized IRFs

Panel (a) Net Flows$_t$ (%) → Returns$_{t+h}$ (%)



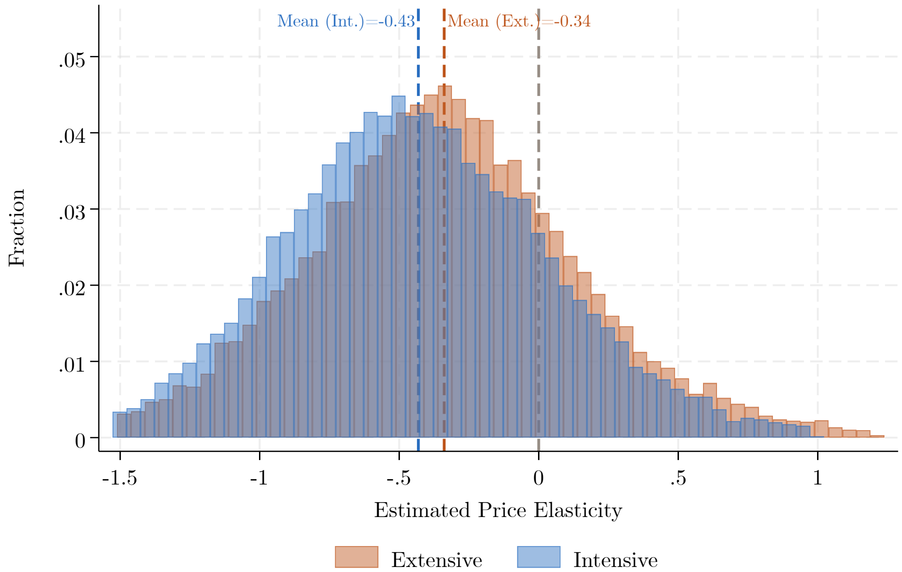
Panel (b) Returns$_t$ (%) → Net Flows$_{t+h}$ (%)



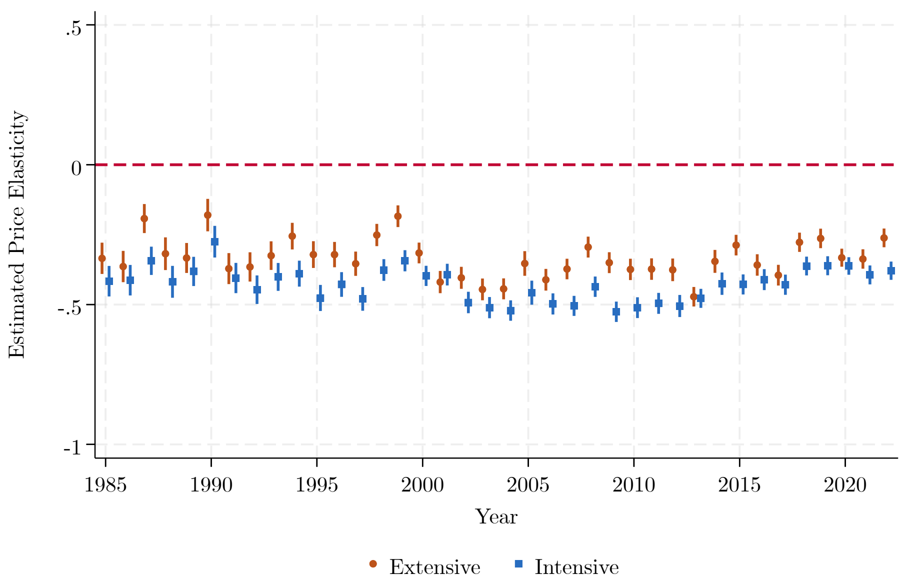
Notes. This figure plots orthogonalized impulse response functions (IRFs). Panel (a) shows the responses of stock returns to a one-percentage-point increase in net extensive or intensive flows, while Panel (b) shows the responses of net flows to a one-percentage-point increase in returns. The IRFs are estimated using local projections (8) with three lags, controlling for time fixed effects, stock fixed effects, lagged dependent and independent variables, standardized unexpected earnings (SUE), size, book-to-market ratio, and profitability.

Figure 7. Estimated Price Elasticities of Extensive and Intensive Demand

Panel (a) Overall Distribution



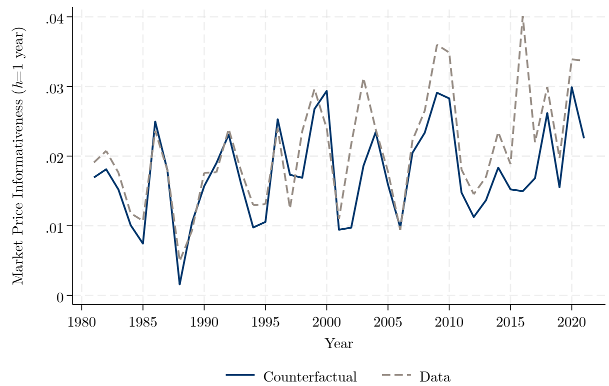
Panel (b) Time Series of Yearly Averages



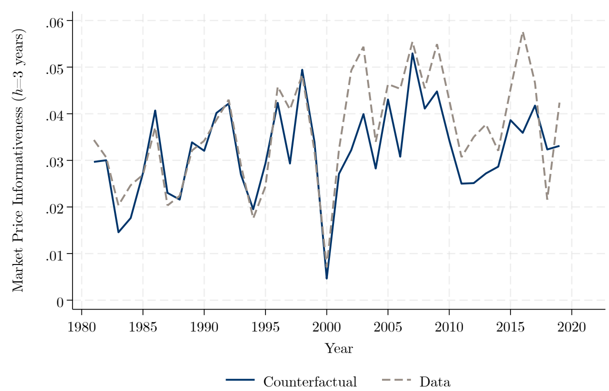
Notes. This figure plots the distribution and time series of the estimated price elasticities of extensive and intensive demand, under the specification of statistical optimal extensive flow expectations. Panel (a) shows the cross-sectional distribution, and Panel (b) reports the annual time series of mean elasticities with 95% confidence intervals.

Figure 8. Counterfactual: Market Price Informativeness

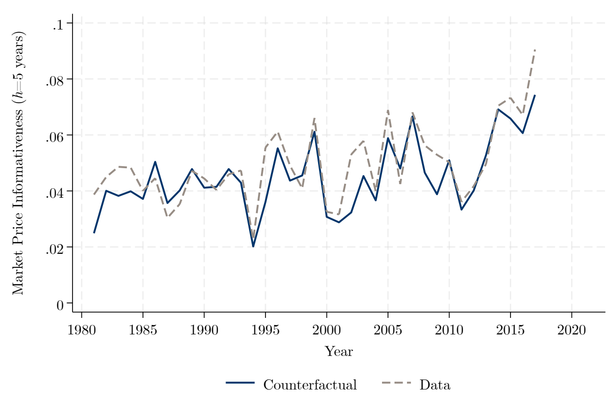
Panel (a) One-Year Horizon



Panel (b) Three-Year Horizon



Panel (c) Five-Year Horizon



Notes. This figure plots the time series of market price informativeness, comparing the data with the counterfactual. Following Bai, Philippon, and Savov (2016), market price informativeness is constructed from the cross-sectional regressions in (44) and computed using (45). Panels (a)–(c) present the results for the one-, three-, and five-year horizons, respectively. The blue solid lines depict the market informativeness of counterfactual prices, while the dashed gray lines depict that of actual prices.

Tables

Table 1. Descriptive Statistics

Panel (a) Investor-Stock-Quarter Panel

	N	Mean	SD	25%	50%	75%
1(Extensive) (%)	93,117,390	25.608	43.646	0.000	0.000	100.000
1(Intensive) (%)	93,117,390	55.048	49.745	0.000	100.000	100.000
Extensive Flow ($ M)	23,845,210	5.424	112.225	0.177	0.484	2.017
Intensive Flow ($ M)	51,259,456	5.377	106.079	0.055	0.277	1.562
Active Share (%)	93,117,390	46.352	16.203	36.678	47.020	57.661
Social Media: Attention (z)	11,427,940	0.000	1.000	−0.709	−0.346	0.326
Social Media: Sentiment (z)	11,427,940	0.000	1.000	−0.628	0.058	0.655
News Media: Coverage	45,648,742	227.623	539.468	43.000	106.000	208.000
News Media: Sentiment (z)	45,648,742	0.000	1.000	−0.439	0.107	0.596
Return (%)	93,117,390	2.809	17.255	−7.854	2.720	13.332
abs(Return) (%)	93,117,390	13.741	10.808	4.869	10.883	20.577
SUE (%)	45,582,021	0.003	0.937	−0.036	0.038	0.164
abs(SUE) (%)	45,582,021	0.391	0.851	0.037	0.113	0.320

Panel (b) Stock-Quarter Panel

	N	Mean	SD	25%	50%	75%
Net Flow: Extensive (%)	235,315	0.201	3.266	−1.118	0.000	1.332
Net Flow: Intensive (%)	235,315	0.578	4.031	−1.295	0.045	1.741
Composition: Extensive (%)	235,315	31.128	17.801	19.046	29.787	41.676
Composition: Intensive (%)	235,315	68.872	17.801	58.324	70.213	80.954
Return (%)	235,315	2.669	21.561	−9.957	1.990	14.319
Volatility (%)	235,191	47.099	28.864	26.445	38.516	58.182
Idiosyncratic Volatility (%)	235,191	41.433	27.349	21.991	33.115	51.278
SUE (%)	235,315	−0.212	4.426	−0.204	0.035	0.282
Size	235,315	6.230	2.120	4.731	6.191	7.639
B/M	235,315	0.679	0.524	0.297	0.543	0.891
Profitability (%)	235,315	−0.012	9.028	−3.058	−0.013	2.952
Stock-Specific Price Informativeness (%)	93,468	4.750	6.828	0.480	2.150	6.159

Notes. This table reports descriptive statistics for quarterly institutional investor holdings (an investor-stock-quarter panel) in Panel (a) and for the quarterly stock panel in Panel (b).

Table 2. Magnitudes of Extensive and Intensive Flow

Panel (a) By Investor Type

	Flow Type			
Investor Type	Extensive In	Extensive Out	Intensive In	Intensive Out
Retail ($K)	9.63	11.29	10.35	12.68
	[54.14%]	[30.35%]	[10.98%]	[4.53%]
Banks	4.36	4.96	3.33	3.27
	[12.74%]	[11.76%]	[37.78%]	[37.73%]
Insurance Companies	4.16	4.78	3.30	3.91
	[11.62%]	[10.53%]	[42.58%]	[35.27%]
Investment Advisors	3.91	3.85	3.07	3.20
	[20.09%]	[17.56%]	[31.41%]	[30.94%]
Mutual Funds	7.54	7.96	7.48	8.03
	[14.2%]	[13.15%]	[37.97%]	[34.68%]
Pension Funds	6.64	6.95	2.05	2.35
	[13.12%]	[11.59%]	[36.87%]	[38.42%]
Others	4.64	3.94	4.14	4.49
	[23.29%]	[20.03%]	[29.03%]	[27.65%]

Panel (b) By Active Share Quintile

	Flow Type			
Active Share Quintile	Extensive In	Extensive Out	Intensive In	Intensive Out
1: Most Passive	8.57	8.61	5.65	5.82
	[11.44%]	[9.36%]	[42.64%]	[36.57%]
2	3.90	4.14	3.02	3.42
	[15.01%]	[13.18%]	[35.62%]	[36.19%]
3	4.13	4.14	3.00	3.48
	[17.29%]	[15.71%]	[33.2%]	[33.8%]
4	4.29	4.11	3.00	3.54
	[20.95%]	[19.15%]	[29.92%]	[29.98%]
5: Most Active	5.14	4.87	2.92	3.39
	[24.79%]	[22.92%]	[25.45%]	[26.83%]

Notes. This table reports the average dollar value (in millions, except for retail investors, whose unit is thousands) and the corresponding percentage share (shown in square brackets below the dollar value) of extensive and intensive inflows and outflows, by investor type in Panel (a) and by active–share quintile in Panel (b).

Table 3. Origins of Extensive and Intensive Flows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1(Extensive) (%)				**1**(Intensive) (%)			
Social Media: Attention	0.678***				0.246			
	(0.246)				(0.238)			
Social Media: Sentiment	−0.194*				0.034			
	(0.099)				(0.087)			
News Media: Coverage		0.243***				0.922***		
		(0.084)				(0.124)		
News Media: Sentiment		−0.805***				0.406***		
		(0.046)				(0.039)		
L.abs(Return) (%)			0.087***				−0.036***	
			(0.004)				(0.004)	
L.abs(SUE) (%)				−1.685***				1.320***
				(0.066)				(0.061)
Investor-Stock FE	Y	Y	Y	Y	Y	Y	Y	Y
Investor-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	11,427,940	45,648,742	93,117,390	45,582,021	11,427,940	45,648,742	93,117,390	45,582,021
No. of Investors	7,467	8,705	12,158	8,704	7,467	8,705	12,158	8,704
R^2	0.529	0.502	0.485	0.503	0.506	0.476	0.452	0.476
Mean of Dep. Var.	20.396	23.438	25.512	23.440	59.046	58.315	55.134	58.323

Notes. This table examines the determinants of investor flow type (extensive or intensive) by estimating specification (5) on an investor (i)-stock (n)-quarter (t) panel. The dependent variables are indicator variables (in percentage terms) for whether a holding is extensive or intensive. Independent variables include a stock's standardized social media attention and sentiment indices from Cookson et al. (2024), news coverage, standardized news sentiment from RavenPack, and the lagged absolute values of returns and standardized unexpected earnings (SUE). News coverage is measured as the log of the number of news articles covering a stock within a quarter. Standard errors are double-clustered by stock and time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4. Impacts of Extensive and Intensive Flows

Panel (a) Relationship to Return and Volatility

	(1)	(2)	(3)
	Return (%)	Total Volatility (%)	Idiosyncratic Volatility (%)
Composition: Extensive (%)	0.038***	0.050***	0.051***
	(0.004)	(0.005)	(0.004)
Stock FE	Y	Y	Y
Time FE	Y	Y	Y
Controls	Y	Y	Y
Observations	234,334	234,230	234,230
No. of Stocks	7,080	7,076	7,076
R^2	0.342	0.697	0.704
Mean of Dep. Var.	2.676	46.963	41.285

Panel (b) Relationship to Price Informativeness

	(1)	(2)	(3)
	Stock-Specific Price Informativeness (%)		
	Full Sample	High	Low
Composition: Extensive (%)	0.209*	0.304***	−0.031***
	(0.123)	(0.063)	(0.008)
Stock FE	Y	Y	Y
Time FE	Y	Y	Y
Controls	Y	Y	Y
Observations	93,468	46,734	46,734
No. of Stocks	3,116	2,575	2,583
R^2	0.496	0.552	0.229
Mean of Dep. Var.	4.752	8.843	0.672

Notes. This table examines the impact of flow composition on stock returns and volatility in Panel (a), and on stock-specific price informativeness in Panel (b), following Dávila and Parlatore (2023, 2025). The specification (6) is estimated on a stock (n)-quarter (t) panel, where the dependent variable is the percentage composition of extensive flows. The intensive-flow composition is omitted because it is perfectly collinear with the extensive-flow measure. In Panel (a), independent variables include stock return, total volatility, and idiosyncratic volatility. In Panel (b), the same specification is estimated for the full sample, for stocks with above-median price informativeness ("High"), and for those with below-median informativeness ("Low"). All regressions include time and stock fixed effects, lagged dependent and independent variables, standardized unexpected earnings (SUE), size, book-to-market ratio, and profitability. Standard errors are double-clustered by stock and time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5. Flow Composition as Stock Heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)
	Return (%)		Total Volatility (%)		Idiosyncratic Volatility (%)	
SUE (%)	0.313***	0.231***	−0.154***	−0.087**	−0.162***	−0.095***
	(0.025)	(0.026)	(0.028)	(0.033)	(0.031)	(0.035)
Composition: Extensive (%)		0.040***		0.049***		0.050***
		(0.004)		(0.005)		(0.004)
Composition: Extensive (%) × SUE (%)		0.003***		−0.002***		−0.002***
		(0.000)		(0.001)		(0.001)
Stock FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Observations	234,334	234,334	234,230	234,230	234,230	234,230
No. of Stocks	7,080	7,080	7.076	7.076	7,076	7,076
R^2	0.340	0.342	0.697	0.697	0.704	0.704
Mean of Dep. Var.	2.676	2.676	46.963	46.963	41.285	41.285

Notes. This table examines how a stock's flow composition affects the relation between stock prices and fundamentals, estimating specification (7) on a stock (n)-quarter (t) panel. Independent variables include stock return, total volatility, and idiosyncratic volatility, while dependent variables include standardized unexpected earnings (SUE), the percentage composition of extensive flows, and their interaction term. The intensive-flow composition is omitted because it is perfectly collinear with the extensive-flow measure. All regressions include time and stock fixed effects, lagged dependent and independent variables, size, book-to-market ratio, and profitability. Standard errors are double-clustered by stock and time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6. Price Elasticities and Demand Sensitivity to Extensive Flow Expectations

Panel (a) Statistically Optimal Expectations

	N	Mean	SD	25%	50%	75%
Price Elasticity						
Extensive	19,445	−0.339	0.475	−0.656	−0.351	−0.040
Intensive	19,445	−0.431	0.462	−0.753	−0.451	−0.122
Pooled	38,890	−0.385	0.471	−0.709	−0.400	−0.079
Demand Sensitivity to						
Flow Expectation ($\gamma_{i,t}$)	19,445	−0.076	0.791	−0.659	−0.209	0.409

Panel (b) Extrapolative Expectations

	N	Mean	SD	25%	50%	75%
Price Elasticity						
Extensive	20152	−0.309	0.473	−0.626	−0.322	−0.013
Intensive	20152	−0.412	0.456	−0.731	−0.431	−0.111
Pooled	40304	−0.360	0.467	−0.682	−0.376	−0.061
Demand Sensitivity to						
Flow Expectation ($\gamma_{i,t}$)	20152	−0.812	0.388	−1.072	−0.810	−0.550

Panel (c) Expectations Proxied by Social Media Attention

	N	Mean	SD	25%	50%	75%
Price Elasticity						
Extensive	5842	−0.280	0.460	−0.578	−0.294	0.002
Intensive	5842	−0.384	0.471	−0.712	−0.387	−0.070
Pooled	11684	−0.332	0.468	−0.647	−0.336	−0.033
Demand Sensitivity to						
Flow Expectation ($\gamma_{i,t}$)	5842	−0.719	2.520	−2.303	−0.733	0.804

Notes. This table reports the estimated price elasticities of extensive and intensive demand, as well as the estimated sensitivity of demand to expected extensive flows, $\gamma_{i,t}$. Each panel presents results under a different belief specification. Within each panel, estimates are reported separately for extensive demand, intensive demand, and the full sample, along with the sensitivity to expected extensive flows. The panels report the mean, standard deviation, and the 25th, 50th, and 75th percentiles of the estimates. Standard errors are shown in square brackets.

Table 7. Price Elasticities of Extensive and Intensive Demand by Active Share Quintile

	Demand Type	
Active Share Quintile	Extensive	Intensive
1: Most Passive	−0.280	−0.219
	[0.466]	[0.411]
2	−0.376	−0.362
	[0.467]	[0.440]
3	−0.345	−0.409
	[0.489]	[0.452]
4	−0.340	−0.509
	[0.480]	[0.456]
5: Most Active	−0.349	−0.638
	[0.466]	[0.441]

Notes. This table reports the estimated price elasticities of extensive and intensive demand across quintiles of active share, under the specification of statistical optimal extensive flow expectations. The first quintile corresponds to the most passive investors and the fifth to the most active. Standard deviations are shown in square brackets.

Table 8. Counterfactual: Return, Volatility, and Stock-Specific Price Informativeness

Panel (a) Return and Volatility (%)

	Data	Counterfactual	Difference
Return (%)	18.790	22.127	3.337***
	(0.136)	(0.146)	(0.041)
Volatility (%)	48.781	52.434	3.653***
	(0.195)	(0.219)	(0.070)

Panel (b) Stock-Specific Price Informativeness

	Data	Counterfactual	Difference
Full Sample	4.520	4.941	0.421***
	(0.015)	(0.012)	(0.013)
High	8.388	9.406	1.018***
	(0.023)	(0.029)	(0.022)
Low	0.640	0.468	−0.172***
	(0.002)	(0.004)	(0.012)

Notes. This table compares return, volatility, and stock-specific price informativeness between the data and the counterfactual. Panel (a) reports annualized returns and volatilities, while Panel (b) reports stock-specific price informativeness following Dávila and Parlatore (2023, 2025). Price informativeness is reported in percentages. The first row shows estimates for the full sample, and the second and third rows show estimates for subsamples with above-median ("High") and below-median ("Low") price informativeness, respectively. Standard errors are reported in brackets.

Internet Appendix

Inelastic Demand at the Extensive and Intensive Margins

Table of Contents

A Stylized Model Proofs

A.1 Proof of Proposition 1

At $t = 1$, all investors share the same prior for each asset: $d_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ for $n \in \{A, B\}$. Under mean-variance preferences, for any investor facing price $p_{n,1}$, the optimal demand is

$$q_{n,1} = \frac{\mu_n - p_{n,1}}{\alpha\sigma_n^2}.$$

Because information is common at $t = 1$, intensive investors and extensive investors (conditional on holding that asset) choose the same per-capita demand.

Market clearing for A and B implies:

$$(1-\phi)q_{A,1} + \phi\theta_1 q_{A,1} = 1, \qquad (1-\phi)q_{B,1} + \phi(1-\theta_1)q_{B,1} = 1,$$

so

$$q_{A,1} = \frac{1}{(1-\phi)+\phi\theta_1}, \qquad q_{B,1} = \frac{1}{(1-\phi)+\phi(1-\theta_1)}.$$

Substituting $q_{n,1} = (\mu_n - p_{n,1})/(\alpha\sigma_n^2)$ yields the price formulas:

$$p_{A,1} = \mu_A - \frac{\alpha\sigma_A^2}{(1-\phi)+\phi\theta_1}, \qquad p_{B,1} = \mu_B - \frac{\alpha\sigma_B^2}{(1-\phi)+\phi(1-\theta_1)}.$$

Finally, extensive investors are indifferent between holding A and B at $t = 1$:

$$V_1(A) = V_1(B).$$

Under CARA-normal, the certainty equivalent for holding asset n with optimal position is

$$CE_n = \frac{(\mu_n - p_{n,1})^2}{2\alpha\sigma_n^2}.$$

Indifference implies

$$\frac{(\mu_A - p_{A,1})^2}{\sigma_A^2} = \frac{(\mu_B - p_{B,1})^2}{\sigma_B^2}.$$

Using the market-clearing expressions $\mu_n - p_{n,1} = \alpha\sigma_n^2/((1-\phi) + \phi \cdot \text{mass}_n)$ gives

$$\frac{\sigma_A^2}{((1-\phi)+\phi\theta_1)^2} = \frac{\sigma_B^2}{((1-\phi)+\phi(1-\theta_1))^2}.$$

Taking square roots and solving yields

$$\theta_1^\star = \frac{\sigma_A - (1-\phi)\sigma_B}{(\sigma_A+\sigma_B)\phi}.$$

The regularity condition in the proposition ensures $\theta_1^\star \in (0,1)$. □

A.2 Proof of Proposition 2

At $t = 2$, investors who observe η_n update beliefs about d_n. Since

$$\eta_n = d_n + e_n, \quad d_n = \mu_n + \varepsilon_n, \quad \varepsilon_n \sim \mathcal{N}(0,\sigma_n^2),\ e_n \sim \mathcal{N}(0, v_n^2),$$

the posterior is Gaussian with

$$\widehat{\mu}_n = \mu_n + g_n(\eta_n - \mu_n), \qquad g_n = \frac{\sigma_n^2}{\sigma_n^2 + v_n^2}, \qquad \widehat{\sigma}_n^2 = \frac{\sigma_n^2 v_n^2}{\sigma_n^2 + v_n^2}.$$

Given posterior $(\widehat{\mu}_n, \widehat{\sigma}_n^2)$ and price $p_{n,2}$, mean-variance optimal demand is

$$q_{n,2} = \frac{\widehat{\mu}_n - p_{n,2}}{\alpha\widehat{\sigma}_n^2} = \frac{\mu_n + g_n(\eta_n - \mu_n) - p_{n,2}}{\alpha\widehat{\sigma}_n^2}.$$

If the investor does not observe η_n and instead uses an inferred signal $\tilde{\eta}_n$, the same derivation applies with η_n replaced by $\tilde{\eta}_n$. □

A.3 Proof of Corollary 1

Fix asset A at $t = 2$. Consider an investor i who does not observe η_A and forms a belief $\tilde{m}_{A,i}$ about the (net) extensive inflow into A as in (19), where $\tilde{m}_{A,i}$ is interpreted as the net buy *quantity* (shares) by unobserved extensive investors.

From her perspective, the investors who directly condition on η_A consist of intensive investors (mass $1-\phi$) and extensive investors who held A at $t = 1$ (mass $\phi\theta_1^\star$), after accounting for the random exit. Let

$$\bar{\theta}_A \equiv \left[(1-\phi) + \phi\theta_1^\star\right](1-\delta)$$

denote the mass of such investors. Since $\tilde{m}_{A,i}$ units of supply are absorbed by the unobserved extensive inflow, the residual supply held by the η_A-informed group is $1 - \tilde{m}_{A,i}$. Therefore, the perceived CARA-normal market-clearing mapping implies

$$p_{A,2} = \widehat{\mu}_A - \frac{\alpha\widehat{\sigma}_A^2}{\bar{\theta}_A}\left(1 - \tilde{m}_{A,i}\right).$$

Rearranging yields

$$\widehat{\mu}_A = p_{A,2} + \frac{\alpha\widehat{\sigma}_A^2}{\bar{\theta}_A}\left(1 - \tilde{m}_{A,i}\right).$$

Using $\widehat{\mu}_A = \mu_A + g_A(\tilde{\eta}_{A,i} - \mu_A)$ (the same Bayesian mapping but applied to the inferred signal) implies

$$\mu_A + g_A(\tilde{\eta}_{A,i} - \mu_A) = p_{A,2} + \frac{\alpha\widehat{\sigma}_A^2}{\bar{\theta}_A}\left(1 - \tilde{m}_{A,i}\right),$$

so

$$\tilde{\eta}_{A,i} = \frac{1}{g_A}\left(p_{A,2} + \frac{\alpha\widehat{\sigma}_A^2}{\bar{\theta}_A}\left(1 - \tilde{m}_{A,i}\right) - \mu_A\right) + \mu_A,$$

which is (24). The expression for $\tilde{\eta}_{B,i}$ follows identically. □

A.4 SFE system at $t = 2$

For completeness, we collect the equilibrium system underlying Definition 2.

Let θ_2 be the equilibrium fraction of extensive investors holding A at $t = 2$. Rather than parameterizing unobserved extensive reallocation by a *mass* of switching investors, we summarize it by a (net) *quantity* order imbalance, analogous to the traditional noise-trader order.

Specifically, define $\tilde{m}_{A,i}$ ($\tilde{m}_{B,i}$) as investor i's perceived (net) extensive inflow into A (B) at $t = 2$ (in shares), i.e., the quantity of A (B) absorbed by extensive investors who did not hold A (B) at $t = 1$:

$$\tilde{m}_{A,n} = f_A(p_{A,2}) + \tilde{\epsilon}_{m,A,n}, \qquad \tilde{m}_{B,n} = f_B(p_{B,2}) + \tilde{\epsilon}_{m,B,n}.$$

Non-holders infer signals from prices using Corollary 1. Given perceived signals, demands are as in Proposition 2. Market clearing for A and B can be written as:

$$(1-\phi)q^{I}_{A,2}(\eta_A,\eta_B,p_{A,2},p_{B,2}) \;+\; \phi\theta^{\star}_{1}\, q_{A,2}(\eta_A,p_{A,2}) \;+\; m_A(\tilde{\eta}_A(\tilde{m}_A,p_{A,2}),p_{A,2}) = 1,$$
$$(1-\phi)q^{I}_{B,2}(\eta_A,\eta_B,p_{A,2},p_{B,2}) \;+\; \phi(1-\theta^{\star}_{1})\, q_{B,2}(\eta_B,p_{B,2}) \;+\; m_B(\tilde{\eta}_B(\tilde{m}_B,p_{B,2}),p_{B,2}) = 1,$$

where $m_A(\cdot)$ and $m_B(\cdot)$ denote the *actual* equilibrium quantities absorbed by extensive entrants/switchers (the extensive order imbalances), and the dependence on $\tilde{\eta}_n(\tilde{m}_n,p_{n,2})$ emphasizes that these quantities are determined jointly with perceived signal extraction.

Finally, extensive investors are indifferent between holding A and B at $t=2$. This indifference pins down θ_2 jointly with prices and the realized extensive order imbalances (m_A,m_B), where investors evaluate A and B using the relevant perceived signals (true if they observe it; inferred otherwise).

This system highlights the central identification problem: prices mix fundamentals and unobserved extensive reallocation (here summarized by quantity order imbalances m), and subjective beliefs $\tilde{m}$ (including the noise term) determine how non-holders invert prices into perceived signals.

B Structural Model Proofs

B.1 Proof of Proposition 3

Belief updating. Intensive investors form posterior beliefs based on their private signal. Their posterior belief about the idiosyncratic shock $\varepsilon(n)$ is:

$$E[\varepsilon(n)|\mathcal{I}_i] = \hat{\varepsilon}_i(n) = \hat{\sigma}^2_{\varepsilon i}(n)\sigma^{-2}_{\eta}\eta_i(n)$$
$$\mathrm{Var}(\varepsilon(n)|\mathcal{I}_i) = \hat{\sigma}^2_{\varepsilon i}(n) = (\sigma^{-2}_{\varepsilon}+\sigma^{-2}_{\eta})^{-1}$$

Suppose the extensive investor i believes the extensive flow comes from the proportion of $\phi_i(n)$ investors. Let $\phi_i = \left[\phi_i(1),\phi_i(2),\ldots,\phi_i(n)\right]^{\top}$ denotes the vectors of this belief. Then, the perceived price signal by extensive investor i is η_{p_i}.

The posterior belief is:

$$E[\varepsilon(n)|\mathcal{I}_i] = \hat{\varepsilon}_i(n) = \hat{\sigma}^2_{\varepsilon i}(n)\sigma^{-2}_{p_i}\eta_{p_i}(n) \tag{A1}$$

$$\text{Var}(\varepsilon(n)|\mathcal{I}_i) = \hat{\sigma}^2_{\varepsilon i}(n) = (\sigma^{-2}_{\varepsilon} + \sigma^{-2}_{p_i})^{-1} \tag{A2}$$

Combining the two cases together, in vector-matrix form, the posterior beliefs for any investor i are:

$$E[\varepsilon|\mathcal{I}_i] = \hat{\varepsilon}_i = \hat{\Sigma}_{\varepsilon i}\Sigma^{-1}_{\bar{\eta}_i}\bar{\eta}_i \tag{A3}$$

$$\text{Var}(\varepsilon|\mathcal{I}_i) = \hat{\Sigma}_{\varepsilon i} = (\Sigma^{-1}_{\varepsilon} + \Sigma^{-1}_{\bar{\eta}_i})^{-1} \tag{A4}$$

where

$$\begin{aligned}\bar{\eta}_i &= s^I_i\eta + s^E_i\eta_{p_i},\\ \Sigma_{\bar{\eta}_i} &= s^I_i\Sigma_\eta + s^E_i\Sigma_{p_i}s^E_i\\ &= \Sigma_\eta + \underbrace{s^E_i\Sigma_{\eta p}s^E_i}_{\Sigma^E_{\eta p}}.\end{aligned}$$

$\Sigma_{\bar{\eta}_i}$ can be written as the sum of a diagonal matrix, Σ_η and a block matrix, $\Sigma^E_{\eta p}$. The only non-zero block in $\Sigma^E_{\eta p}$ is its bottom-right block.

First order conditions. Solving the first-order condition, the optimal demand for investor i is:

$$q_i = \frac{1}{\gamma_i}\hat{\Sigma}^{-1}_{di}(\hat{\mu}_{di} - pR_f) \tag{A5}$$

Subjective market clearing conditions. We then pin down the coefficients of the optimal demand for each investor based on their subjective beliefs about the composition of the market. Specifically, the subjective market-clearing condition for investor i is that the perceived aggregate demand equals supply:

$$\int_0^1 \hat{q}_j dj = \tilde{m}_i + \tilde{\epsilon}_{m,i} + \int_{\phi_i}^1 \gamma_j^{-1}\hat{\Sigma}^{-1}_{dj}(\hat{\mu}_{dj} - pR_f)dj = \mathbf{1}$$

where ϕ_i represents the starting point of the continuum of intensive investors, and define

$$\int_{\phi_i}^{1} f(j)dj = \begin{bmatrix} \int_{\phi_i(1)}^{1} f(j)dj \\ \vdots \\ \int_{\phi_i(n)}^{1} f(j)dj \end{bmatrix}$$

Denote the beliefs for "pure" intensive investors (who invest in all the assets and therefore $s_i = \mathbf{1}$). Their posterior belief about the idiosyncratic shock is $\hat{\varepsilon}_I = \hat{\Sigma}_{\varepsilon I}\Sigma_\eta^{-1}\eta$, with posterior variance $\hat{\Sigma}_{\varepsilon I} = (\Sigma_\varepsilon^{-1} + \Sigma_\eta^{-1})^{-1}$. The conditional variance of payoffs is thus $\hat{\Sigma}_{dI} = \rho\rho' + \hat{\Sigma}_{\varepsilon I}$. We define the aggregate risk tolerance of the intensive investors as $\bar{\gamma}_{iI}^{-1} = \int_{\phi_i}^{1} \gamma_j^{-1} dj$, and the aggregate risk aversion $\bar{\gamma}_{iI}$ as the (element-wise) reciprocal of the aggregate risk tolerance. The aggregate demand from this group is then:

$$\int_{\phi_i}^{1} q_j dj = \bar{\gamma}_{iI}^{-1}\hat{\Sigma}_{dI}^{-1}(\mu + \hat{\varepsilon}_I - pR_f)$$

Substituting this into the subjective market clearing condition and solving for p yields:

$$p = \frac{1}{R_f}\left[\mu + \hat{\Sigma}_{\varepsilon I}\Sigma_\eta^{-1}\eta - \bar{\gamma}_{iI}\hat{\Sigma}_{dI}(\mathbf{1} - \tilde{m}_i - \tilde{\epsilon}_{m,i})\right].$$

This equation represents the equilibrium price as perceived by an investor with subjective beliefs m_i and ϕ_i. It is a linear function of the expected fundamentals (μ) and the aggregated information on idiosyncratic shocks(η), adjusted by a term that depends on the perceived mass of extensive traders. From this, an extensive investor can back out the implied signal from the price, η_{p_i} as:

$$\eta_{p_i} = \Sigma_\eta\hat{\Sigma}_{\varepsilon I}^{-1}\left[pR_f - \mu + \bar{\gamma}_{iI}\hat{\Sigma}_{dI}(\mathbf{1} - \tilde{m}_i)\right]. \tag{A6}$$

Therefore, the perceived covariance matrix of the price signal is

$$\Sigma_{p_i} = \Sigma_\eta + \underbrace{\bar{\gamma}_{iI}^2\left(\Sigma_\eta\hat{\Sigma}_{\varepsilon I}^{-1}\hat{\Sigma}_{dI}\right)\Sigma_m\left(\Sigma_\eta\hat{\Sigma}_{\varepsilon I}^{-1}\hat{\Sigma}_{dI}\right)'}_{\Sigma_{\eta p}}. \tag{A7}$$

Rewriting optimal demand. Substituting all posterior beliefs into the solution of the first-order condition (A5) yields:

$$q_i = \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\left(\left(\mathbf{1}-\delta_i\right)\mu + g(\tilde{m}_i) + \hat{\Sigma}_{\varepsilon i}^{-1}\Sigma_{\bar{\eta}_i}^{-1}s^I\eta - \left(\mathbf{1}-\delta_i\right)R_f p\right),$$

where $\delta_i = \hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}i}^{-1}s^E\Sigma_\eta\hat{\Sigma}_{\varepsilon I}^{-1}$, and $g(\tilde{m}_i) = \delta_i\bar{\gamma_{iI}}\hat{\Sigma}_{dI}(\mathbf{1}-\tilde{m}_i)$. δ_i is a block matrix in which only the bottom-right block is nonzero.

B.2 Proof of Proposition 4

Objective market clearing condition. The objective market clearing condition can be written as:

$$\int q_i di = \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}(\hat{\mu}_{di} - pR_f)di = \mathbf{1} \tag{A8}$$

Define a matrix Ω as the risk-tolerance-weighted average posterior precision of all investors:

$$\Omega = \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}di$$

The term Ω captures the aggregate sensitivity of investors' static demands to expected excess returns, holding their information sets and subjective beliefs fixed. The aggregate demand can then be written as:

$$\int q_i di = \Omega(\mu - pR_f) + \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\hat{\varepsilon}_i di \tag{A9}$$

The second term captures the demand component driven by the posterior about the idiosyncratic shocks.

Substituting the posterior beliefs (32) and (33) into the second term of (A10) yields:

$$\begin{aligned}
&\int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\hat{\varepsilon}_i \mathrm{d}i \\
&= \int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}_i}^{-1}\left(s_i^I\eta + s_i^E\eta_{pi}\right)\mathrm{d}i \\
&= \underbrace{\int r_i^{-1}\hat{\Sigma}_{di}^{-1}\hat{\Sigma}_{\varepsilon i}\Sigma_{\eta_i}^{-1}\Sigma_{\bar{\eta}_i}^{-1}s_i^I d_i}_{\Omega_\eta}\cdot\eta \\
&+ \underbrace{\int \gamma_i^{-1}\hat{\Sigma}_{di}^{-1}\hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}_i}^{-1}s_i^E\Sigma_\eta\hat{\Sigma}_{\varepsilon I}^{-1}}_{\Omega_p}\left[pR_f - \mu + \bar{\gamma}_{iI}\hat{\Sigma}_{dI}\left(\mathbf{1}-\tilde{m}_i\right)\right] d_i \\
&= \Omega_\eta\eta + \Omega_p\left(pR_f - \mu\right) + g(\phi, m),
\end{aligned} \tag{A10}$$

where ϕ and m are the matrix that combines all the subjective beliefs about extensive flow across investors. $g(\phi, m)$ is then a function of aggregating subjective beliefs.

Substituting (A10) back into the objective market clearing condition (A8) and collecting terms, we obtain:

$$p = \frac{1}{R_f}\left[\mu - (\Omega - \Omega_p)^{-1}\left(\mathbf{1} - \Omega_m(\bar{m})\right) + (\Omega - \Omega_p)^{-1}\Omega_\eta\eta\right]$$

B.3 Proof of Corollary 2

We rewrite investors' demand into a form that separates own and cross-asset substitutions and obtain the Corollary of Proposition 3.

Corollary 3. *Investors' optimal demand can be rewritten as*

$$q_i = \gamma_i^{-1}\left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}_{\varepsilon i}^{E,d}\right)^{-1}\left(-\left(I - \delta_i^{E,d}\right)R_f p + \left(I - \delta_i^{E,d}\right)\mu + \delta_i^{E,d}\bar{\gamma}_{iI}\hat{\Sigma}_{dI}(\mathbf{1} - \tilde{m}_i) + \Sigma_{\varepsilon I}\hat{\Sigma}_\eta^{-1}s^I\eta - \zeta_i\right) \tag{A11}$$

We separate the posterior variance of idiosyncratic shocks, $\hat{\Sigma}_{\varepsilon i}$, into a diagonal matrix, $\hat{\Sigma}_{\varepsilon i}^{E,d}$, and a block matrix with only the bottom-right block being non-zero. The non-zero part is a rank-1 matrix and captures the posterior covariance among assets induced by learning from prices. We obtain a similar separation of δ_i and the diagonal matrix is $\delta_i^{E,d}$. ζ_i is an investor-specific vector that captures the cross-asset substitution induced by investors' posterior beliefs.

Proof of Corollary 3

Proof. First, we separate the posterior variance about idiosyncratic shocks into one diagonal matrix and one rank-1 matrix with only the bottom right block is nonzero. We want to simplify $\hat{\Sigma}_{\varepsilon i}^{-1} = \Sigma_{\varepsilon}^{-1} + \Sigma_{\bar{\eta}_i}^{-1}$. Since we have

$$\Sigma_{\bar{\eta}_i} = \Sigma_{\eta} + \underbrace{s^E \Sigma_{\eta p} s^E}_{\Sigma_{\eta p}^E},$$

using the Woodbury matrix identity yields:

$$\Sigma_{\bar{\eta}_i}^{-1} = (\Sigma_{\eta} + \Sigma_{\eta p}^E)^{-1} = \Sigma_{\eta}^{-1} - \underbrace{\Sigma_{\eta}^{-1}\Sigma_{\eta p}^E (I + \Sigma_{\eta}^{-1}\Sigma_{\eta p}^E)^{-1}\Sigma_{\eta}^{-1}}_{\Lambda_{\bar{\eta}_i}^E}$$

Next, we substitute this result into definition for $\hat{\Sigma}_{\varepsilon i}^{-1}$:

$$\hat{\Sigma}_{\varepsilon i}^{-1} = \Sigma_{\varepsilon}^{-1} + \Sigma_{\bar{\eta}_i}^{-1}.$$

Substituting the expression from Step 1:

$$\hat{\Sigma}_{\varepsilon i}^{-1} = \Sigma_{\varepsilon}^{-1} + \left[\Sigma_{\eta}^{-1} - \Sigma_{\eta}^{-1}\Sigma_{\eta p}^E (I + \Sigma_{\eta}^{-1}\Sigma_{\eta p}^E)^{-1}\Sigma_{\eta}^{-1}\right].$$

We can group the first two terms:

$$\begin{aligned}\hat{\Sigma}_{\varepsilon i}^{-1} &= (\Sigma_{\varepsilon}^{-1} + \Sigma_{\eta}^{-1}) - \Sigma_{\eta}^{-1}\Sigma_{\eta p}^E (I + \Sigma_{\eta}^{-1}\Sigma_{\eta p}^E)^{-1}\Sigma_{\eta}^{-1} \\ &= \hat{\Sigma}_{\varepsilon I}^{-1} - \Lambda_{\bar{\eta}_i}^E\end{aligned}$$

Applying Woodbury identity again yields:

$$\hat{\Sigma}_{\varepsilon i} = \hat{\Sigma}_{\varepsilon I} + \underbrace{\left(\hat{\Sigma}_{\varepsilon I}\Sigma_{\eta}^{-1}\Sigma_{\eta p}^E\right)\left[(I + \Sigma_{\eta}^{-1}\Sigma_{\eta p}^E) - \Sigma_{\eta}^{-1}\hat{\Sigma}_{\varepsilon I}\Sigma_{\eta}^{-1}\Sigma_{\eta p}^E\right]^{-1}\left(\Sigma_{\eta}^{-1}\hat{\Sigma}_{\varepsilon I}\right)}_{\hat{\Sigma}_{\varepsilon i}^E}.$$

From $\hat{\Sigma}_{\varepsilon i}^{-1} = \Sigma_{\varepsilon}^{-1} + \Sigma_{\bar{\eta}_i}^{-1}$, we have $\hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}_i}^{-1} = I - \hat{\Sigma}_{\varepsilon i}\Sigma_{\varepsilon}^{-1}$. Substituting $\hat{\Sigma}_{\varepsilon i} = \hat{\Sigma}_{\varepsilon I} + \Sigma_{2i}^E$ inside yields

$\hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}_i}^{-1} = I - \hat{\Sigma}_{\varepsilon I}\Sigma_{\varepsilon}^{-1} - \Sigma_{2i}^{E}\Sigma_{\varepsilon}^{-1}$. Using the definition of $\hat{\Sigma}_{\varepsilon I}$, we have $I - \hat{\Sigma}_{\varepsilon I}\Sigma_{\varepsilon}^{-1} = \hat{\Sigma}_{\varepsilon I}\Sigma_{\eta}^{-1}$. We can then derive:

$$\hat{\Sigma}_{\varepsilon i}\Sigma_{\bar{\eta}_i}^{-1} = \hat{\Sigma}_{\varepsilon I}\Sigma_{\eta}^{-1} - \Sigma_{2i}^{E}\Sigma_{\varepsilon}^{-1}.$$

Next, we want to further simplify $\hat{\Sigma}_{\varepsilon i}^{E}$ to separate it to a diagonal matrix and a rank-1 matrix: Substituting the expressions of $\Sigma_{\eta p} = \bar{\gamma}_{iI}\Sigma_{\eta}\hat{\Sigma}_{\varepsilon I}^{-}1\hat{\Sigma}_{dI}$, $\hat{\Sigma}_{\varepsilon I} = (\Sigma_{\varepsilon}^{-1} + \Sigma_{\eta}^{-1})^{-1}$, and $\hat{\Sigma}_{dI} = \rho\rho' + \hat{\Sigma}_{\varepsilon I}$ into $\hat{\Sigma}_{\varepsilon i}^{E}$ and we can simplify the expression as:

$$\hat{\Sigma}_{\varepsilon i}^{E} = \underbrace{\bar{\gamma}_{iI}s^{E}\hat{\Sigma}_{\varepsilon I}\left[I + \bar{\gamma}_{iI}s^{E}\Sigma_{\varepsilon}^{-1}\hat{\Sigma}_{\varepsilon I}s^{E}\right]^{-1}\Sigma_{\eta}^{-1}\hat{\Sigma}_{\varepsilon I}s^{E}}_{\hat{\Sigma}_{\varepsilon i}^{E,d}}$$
$$+ \underbrace{k_i v_{i1} v_{i2}'}_{\hat{\Sigma}_{\varepsilon i}^{E,r}},$$

where

$$k_i = \frac{\bar{\gamma}_{iI}}{1 + \bar{\gamma}_{iI}\rho' s^{E}\left[I + \bar{\gamma}_{iI}s^{E}\Sigma_{\varepsilon}^{-1}\hat{\Sigma}_{\varepsilon I}s^{E}\right]^{-1}\Sigma_{\varepsilon}^{-1}\rho},$$
$$v_{i1} = s^{E}\left(I - \bar{\gamma}_{iI}\hat{\Sigma}_{\varepsilon I}\left[I + \bar{\gamma}_{iI}s^{E}\Sigma_{\varepsilon}^{-1}\hat{\Sigma}_{\varepsilon I}s^{E}\right]^{-1}\Sigma_{\varepsilon}^{-1}\right)\rho,$$
$$v_{i2}' = \rho' s^{E}\left[I + \bar{\gamma}_{iI}s^{E}\Sigma_{\varepsilon}^{-1}\hat{\Sigma}_{\varepsilon I}s^{E}\right]^{-1}\Sigma_{\eta}^{-1}\hat{\Sigma}_{\varepsilon I}.$$

k_i is a scalar, and v_{i1} and v_{i2} are two N-by-1 vectors. Next, we rewrite $\hat{\Sigma}_{di}$. Substituting $\hat{\Sigma}_{\varepsilon i}$ inside and regrouping, we obtain:

$$\hat{\Sigma}_{di} = \underbrace{\left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}_{\varepsilon i}^{E,d}\right)}_{\text{Combined Diagonal Part}} + \underbrace{\left(\rho\rho' + k_i v_{i1} v_{i2}'\right)}_{\text{Combined Rank-2 Part}}$$

Using the Woodbury matrix identity again yields:

$$\hat{\Sigma}_{di}^{-1} = \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}_{\varepsilon i}^{E,d}\right)^{-1} + \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}_{\varepsilon i}^{E,d}\right)^{-1}\begin{pmatrix}\rho & v_{i1}\end{pmatrix} M_i,$$

where M_i is a 2-by-N matrix that is fixed for investor i and doesn't vary across assets.

$$M_i = \left[\begin{pmatrix} 1 & 0 \\ 0 & 1/k_i \end{pmatrix} + \begin{pmatrix} \rho' \\ v'_{i2} \end{pmatrix} \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1} \begin{pmatrix} \rho & v_{i1} \end{pmatrix}\right]^{-1} \begin{pmatrix} \rho' \\ v'_{i2} \end{pmatrix} \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1}$$

Expand and rewrite $\begin{pmatrix} \rho & v_{i1} \end{pmatrix} M_i$:

$$\begin{aligned}
\begin{pmatrix} \rho & v_{i1} \end{pmatrix} M_i &= \rho R_{i1} + v_{i1} R_{i2}, \\
\text{where } R_{i1} &= 1/\det(M_i) \left[(1 + v'_2 \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1} v_1)\rho' - (\rho' \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1} v_1) v'_2\right] \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1}, \\
R_{i2} &= 1/\det(M_i) \left[-(v'_2 \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1} \rho)\rho' + (1 + \rho' \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1} \rho) v'_2\right] \left(\hat{\Sigma}_{\varepsilon I} + \hat{\Sigma}^{E,d}_{\varepsilon i}\right)^{-1}.
\end{aligned}$$

Substituting the expression of $\hat{\Sigma}^E_{\varepsilon i}$ into δ_i yields:

$$\begin{aligned}
\delta^E_i(s_i, \phi_i) &= \hat{\Sigma}_{\varepsilon i} \Sigma^{-1}_{pi} \Sigma_\eta \hat{\Sigma}^{-1}_{\varepsilon I} \\
&= \hat{\Sigma}_{\varepsilon I} \Sigma^{-1}_\eta - \Sigma^E_{2i} \Sigma^{-1}_\varepsilon s^E \Sigma_\eta \hat{\Sigma}^{-1}_{\varepsilon I} \\
&= \underbrace{s^E \left(I - \Sigma^{-1}_\varepsilon \Sigma_\eta \hat{\Sigma}^{-1}_{\varepsilon I} \hat{\Sigma}^{E,d}_{\varepsilon i}\right)}_{\delta^{E,d}_i} - \underbrace{s^E \Sigma^{-1}_\varepsilon \Sigma_\eta \hat{\Sigma}^{-1}_{\varepsilon I} \hat{\Sigma}^{E,r}_{\varepsilon i}}_{\delta^{E,r}_i}
\end{aligned}$$

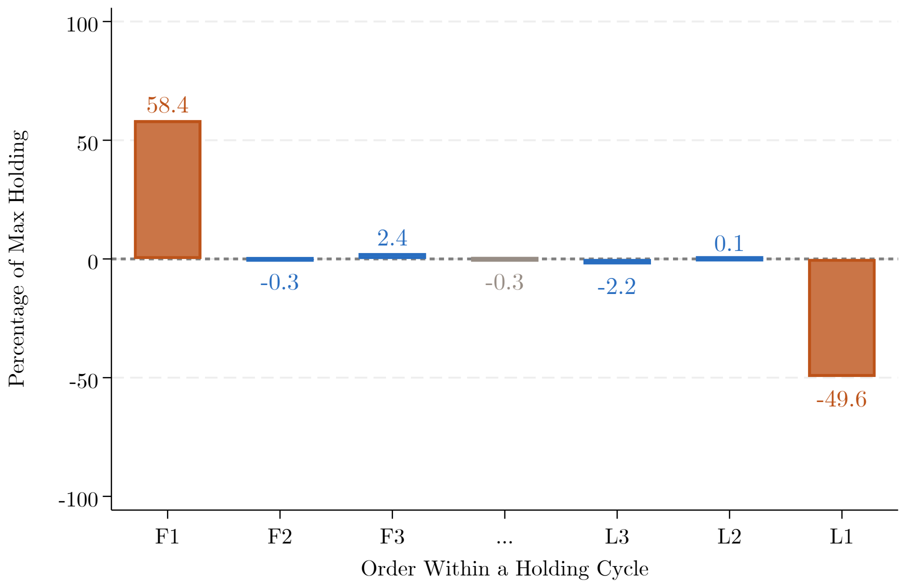
Substituting the above expressions into the optimal demand in Proposition 3, gives the results.

□

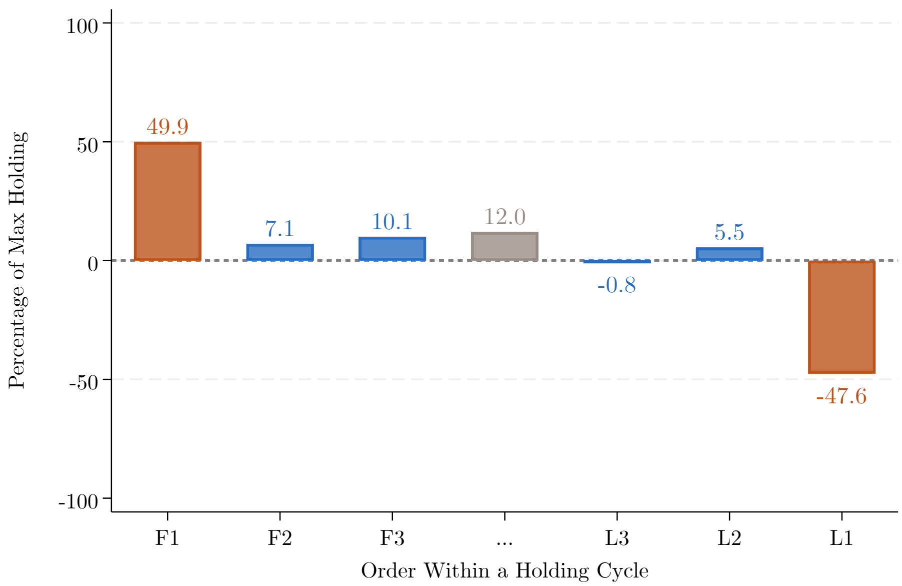
C Appendix Figures

Figure A1. Flow Distribution Within a Holding Cycle (Excluding One-Buy-One-Sell)

Panel (a) Institutional Investors



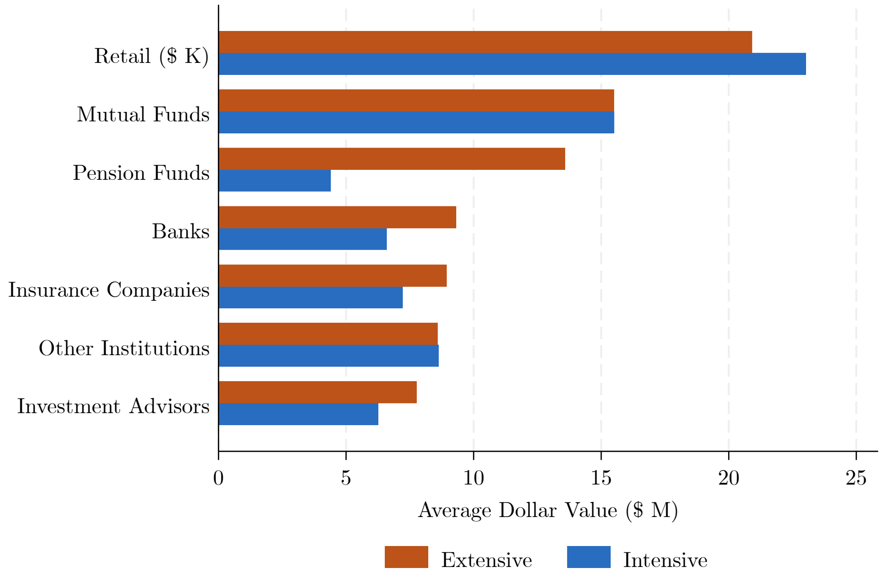
Panel (b) Retail Investors



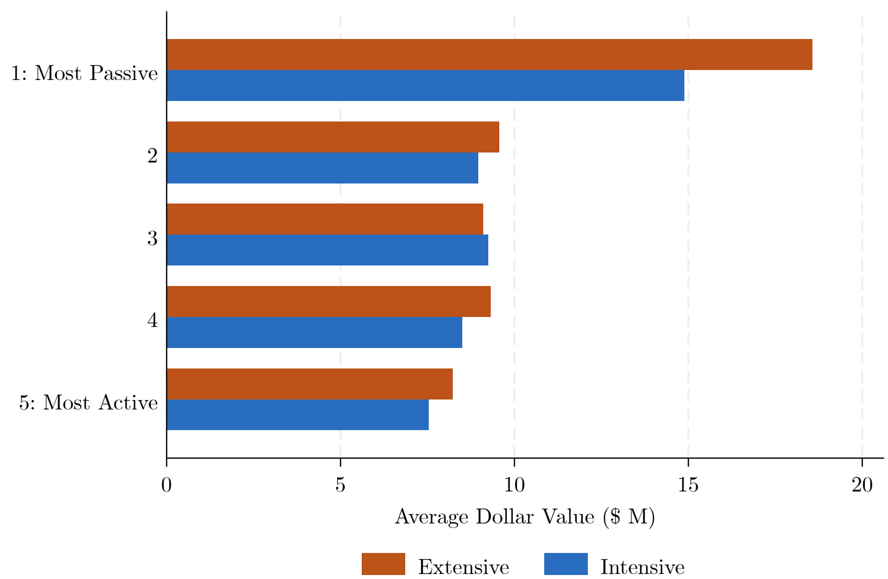
Notes. This figure plots the distribution of flows as a percentage of the maximum position within an average holding cycle, using the institutional investor sample in Panel (a) and the retail investor sample in Panel (b). Both samples exclude holding cycles that contain only one buy and one sell observation. F1–F3 denote the first, second, and third periods of the cycle, while L1–L3 correspond to the last, second-to-last, and third-to-last periods. Red bars indicate extensive flows, whereas blue and gray bars represent intensive flows.

Figure A2. Dollar Values of Extensive and Intensive Flows

Panel (a) By Investor Type



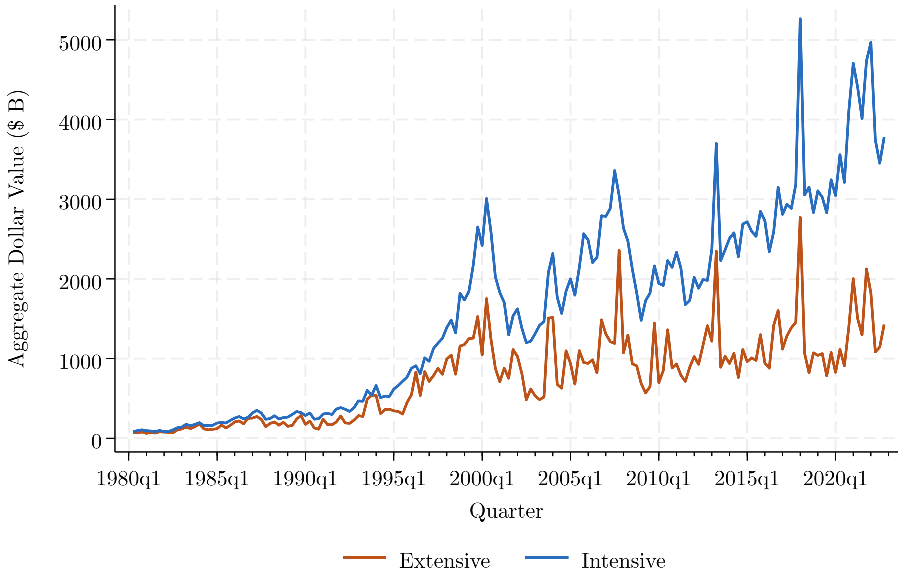
Panel (b) By Active Share Quintile



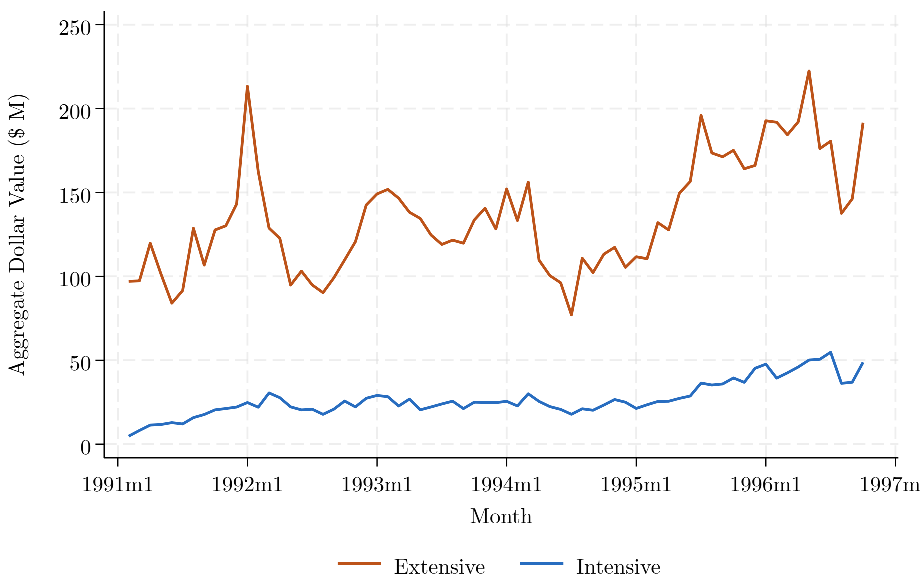
Notes. This figure plots the average dollar value of extensive flows (the sum of extensive in and out) and intensive flows (the sum of intensive in and out), by investor type in Panel (a) and by active–share quintile in Panel (b).

Figure A3. Dollar Values of Aggregate Flows

Panel (a) Institutional Investors



Panel (b) Retail Investors



Notes. This figure plots the time series of the aggregate dollar value of extensive flows (the sum of extensive in and out) and intensive flows (the sum of intensive in and out), using the institutional investor sample in Panel (a) and the retail investor sample in Panel (b).