Passive Flows and the Limits to Arbitrage

Yao Deng* Marco Sammon†

August 2026

Abstract

We show that passive flow-induced trading (PFIT) acts as a limit to arbitrage. Across a comprehensive sample of 146 anomalies, we find that long-short strategies earn 5.01% per year among stocks with low realized passive inflows during the holding quarter, but only 0.60% among stocks with high inflows. This difference is highly significant and is positive for about 70% of individual anomalies. While passive inflows raise returns in both legs, they raise returns by more in the short leg, even though the two legs receive nearly identical PFIT on average. Our proposed mechanism is that this asymmetry reflects larger demand multipliers in anomaly short legs. Consistent with this mechanism, estimated multipliers are larger for short-leg portfolios, and anomalies with larger short-minus-long multiplier gaps exhibit greater return compression. One amplification channel is short covering: among heavily shorted stocks, passive-flow shocks are followed by larger declines in short positions.

Keywords: passive investing; anomalies; flow-induced trading; inelastic demand; limits to arbitrage; short selling.

JEL Classification: G11, G12, G14, G23.

*University of Connecticut; yao.deng@uconn.edu.

†Harvard Business School, Harvard University; mcsammon@gmail.com.

We thank Adi Sunderam and seminar participants at Fordham University for helpful comments.

1 Introduction

As index mutual funds and ETFs have grown, so has the importance of mechanical demand associated with their portfolio allocation rules. When money flows into or out of an index fund, the fund will scale its positions up or down in proportion to what it held before the flow arrived [Sammon and Shim, 2026b]. And, the information content and predictability of these trades will affect how prices respond to passive demand. One possibility is that these flows reflect information at the *fund level*. Investors may use sector funds to trade on sector news or style funds to execute factor-timing strategies. The resulting pro-rata scaling by index funds, however, need not be the optimal portfolio adjustment implied by all available information for each underlying stock. For example, an investor who receives a signal about a new technology may want to buy a tech-focused ETF, but that does not imply that all stocks in the fund will benefit equally from the innovation. Thus, even if the fund-level flow is informative on average about the tech sector's prospects, the resulting *stock-level* trades may not allocate capital exactly to the expected winners and losers.

A second possibility is that index fund flows are not informed. Many investors allocate to index funds mechanically through retirement accounts and other savings vehicles. In these cases, fund flows may reflect time-varying savings decisions or institutional allocation mandates, such as rebalancing to maintain a 60-40 target stock-bond allocation, rather than information about expected returns. If so, the induced stock-level trades will look even more like inelastic demand shocks: the index fund scales its positions because money enters or exits the fund, not because investors have learned something about the securities it will buy or sell. Under either view, passive flow-induced trading likely contains a large mechanical component at the stock level, which may affect asset prices and therefore the opportunities available to active managers.

In this paper, we ask whether these mechanical demand shocks affect the realized returns to long-short anomaly portfolios. Some practitioners have argued that the growth of passive investing makes it harder for active managers to make money because passive flows

push prices away from fundamental values [Sorkin, 2017]. But long-short anomaly strategies are typically convergence trades: they buy stocks that appear cheap and short stocks that appear expensive. Their profitability, therefore, depends on *relative* price movements. If passive flows create similar price pressure on both the long and short legs of these portfolios, anomaly spreads may be unaffected. In fact, passive flows should compress long-short anomaly strategy returns only if (1) there is more passive demand in the short leg than the long leg, and/or (2) the price pressure per unit of passive demand is stronger in the short leg than in the long leg.

To answer this question, we focus on passive flow-induced trading, or PFIT. We construct PFIT by applying fund-level flows to each index fund's lagged stock holdings, which gives the implied flow-induced trade in each stock [Sammon and Shim, 2026b]. We then aggregate these implied trades across all index funds holding the stock and scale the total by lagged shares outstanding. Our PFIT measure, therefore, is a stock-quarter-level demand shock, measured as a fraction of each stock's shares outstanding.

Importantly, there is substantial stock-level variation in PFIT each quarter. This variation is driven by the fact that there are thousands of index funds, each with different investment universes, portfolio weights, and flows [Sammon and Shim, 2026a]. When thinking about passive investing, many of the most salient products are broad-market funds such as Vanguard's Total Stock Market ETF (Ticker: VTI) or the SPDR S&P 500 ETF Trust (Ticker: SPY). But the passive universe also includes thousands of sector, style, and factor funds with much smaller investment universes. Aggregating flow-induced trades across funds with differing portfolios generates substantial cross-sectional variation in PFIT, both across stocks and over time. In our sample, PFIT averages 0.13% of shares outstanding per quarter, with a cross-sectional standard deviation of 0.43%, and it is only weakly correlated with size and essentially uncorrelated with active flow-induced trading.

Two properties of passive flows make PFIT particularly useful for identifying demand *shocks*. The first is that passive flows are hard to predict. As discussed in Lou [2012], active

fund flows are strongly predicted by past flows and past performance. Passive flows are predictable to some degree, but much less so: in our sample, our forecast of passive flows has an out-of-sample R^2 of 2.1% and a correlation of 0.14 with realized flows, compared with 14.7% and 0.38 for active flows. The second property is that passive flows do not appear to be information-driven. Specifically, we find that current passive flows do not forecast subsequent fund returns across a wide variety of possible specifications.

These two features of passive flows matter because if markets are forward looking, prices should incorporate *expected* future demand when it becomes predictable rather than when the trades are executed. Further, if passive flows were driven by information, they could be correlated with changes in prices for reasons other than a response to the inelastic demand itself. Because passive flows are neither predictable nor informative, realized PFIT is close to an unanticipated, uninformed demand shock.

We study the asset pricing implications of these demand shocks across a broad anomaly universe. Rather than pre-selecting a handful of signals, we begin from the full set of predictors in the Chen and Zimmermann [2022] open-source asset pricing database. We then filter to signals for which the double-sorted portfolios used in our tests are well populated, leaving us with 146 anomalies. Following Muravyev, Pearson, and Pollet [2025], we report cross-anomaly averages together with the distribution of anomaly-level estimates in order to show that the results are not driven by a small number of signals with extreme returns.

Our main finding is that, in the cross section, passive inflows significantly reduce realized long-short anomaly returns. To show this, we restrict to stock-quarters with positive PFIT and double sort stocks every quarter on PFIT and the anomaly signal. Specifically, we first sort stocks into low-, middle-, and high-PFIT terciles using realized PFIT during the holding quarter. Within each PFIT tercile, we then sort stocks into anomaly deciles using the signal available at portfolio formation (i.e., the last day before the holding quarter begins). Averaging across the 146 anomalies, the long-short anomaly strategy return falls monotonically from 5.01% among low-PFIT stocks to 2.67% in the middle tercile and finally

to 0.60%, statistically indistinguishable from zero, among high-PFIT stocks. The low-minus-high difference is 4.41% with a cross-anomaly t-statistic of 7.15, and the corresponding CAPM-alpha difference is 4.56%.

This positive difference in returns between low and high PFIT stocks is not just concentrated in a few anomaly signals. Ranking the 146 anomaly-level estimates, we find that 69.9% have a positive low-minus-high difference in long-short returns. There is, however, significant heterogeneity across anomalies, with anomaly-level differences range from roughly -9% to $+38\%$. These results are robust to alternative weighting and PFIT samples. While the baseline portfolios are equal-weighted following Muravyev et al. [2025], we find that with value-weighted portfolios, the long-short return falls from 3.59% in the low-PFIT tercile to -0.71% in the high-PFIT tercile, a difference of 4.30% ($t = 6.34$) that is positive for 69.2% of anomalies. The online appendix shows that the conclusion also holds when we include stock-quarters with negative or zero PFIT.

One possible explanation for these results is that the short leg of each anomaly portfolio has a higher *level* of PFIT than the long leg. We find that this is not the case. By construction, our double sorts compare long and short legs within the same PFIT tercile. Further, in single-sorted portfolios, we find that average PFIT is 0.121% of shares outstanding in the short leg and 0.118% in the long leg. What differs across long and short legs, however, is the return response to a given amount of PFIT. Moving from the low- to the high-PFIT tercile raises short-leg returns by 14.77% but long-leg returns by only 10.36%. In other words, passive inflows appear to push up prices in both legs, but they affect stocks in the short leg relatively more.

The heterogeneous effects of PFIT on the long and short legs of anomaly portfolios is generally consistent with the idea that stocks in anomaly short legs are generally harder to arbitrage. Prior work shows that short-leg stocks tend to be more costly to arbitrage, more heavily shorted, and more exposed to short-sale constraints [Muravyev et al., 2025]. Our single sorts show the same long- and short-leg asymmetry: relative to the long leg,

the average short leg has a higher short-interest ratio, higher securities-lending utilization, and higher borrowing fees. In this paper, our contribution is to show that this asymmetry in limits to arbitrage across legs matters for the transmission of passive demand shocks to prices, and ultimately the strategies used by active managers.

A natural concern is that PFIT is correlated with other stock characteristics that predict returns. We address this concern using stock-level Fama-MacBeth regressions that interact PFIT with each anomaly signal's rank and additionally control for size, the level of passive ownership, active flow-induced trading, short interest, liquidity, and market beta. We find that the PFIT interaction term remains negative and significant even with all these additional controls. In the fully saturated specification, a 0.10% increase in PFIT reduces the annualized return spread between the top and bottom anomaly decile by about 0.57% ($t = -4.04$).

Our results on longer-horizon returns provide additional support for a price-pressure interpretation rather than one based on differences in fundamentals across high and low PFIT stocks. In both the short and long legs of each anomaly, high-PFIT portfolios earn especially high returns during the quarter in which passive flow-induced trading occurs, while high- and low-PFIT portfolio returns are similar afterward. This pattern is perhaps not surprising given that, as discussed above, PFIT itself is not persistent. The absence of a sharp reversal is consistent with incomplete pass-through of demand shocks into future returns [Davis, Han, Sokolinski, and Tamoni, 2026]. More importantly for our interpretation, the PFIT sort does not appear to capture persistent differences in fundamentals as, if it did, we would expect the return gap to persist beyond the first holding quarter.

We then estimate demand multipliers directly. Within each anomaly decile, we regress stock-level quarterly returns on realized PFIT, controlling for the Fama-French five factors. The coefficient on PFIT, therefore, is our estimate of the demand multiplier. Across our sample of anomalies, we find that the average short-leg multiplier is 2.77 and the average long-leg multiplier is 2.57. The mean gap of 0.21 is positive with a cross-anomaly t-statistic of 2.02, providing evidence that the same demand shock moves short-leg prices more.

This evidence is suggestive of our proposed mechanism, but it is not the whole story. We can learn more by exploiting heterogeneity in return and multiplier gaps across anomalies. Specifically, we combine each anomaly's leg-specific multipliers with the average demand shocks in its own double-sorted portfolios to calculate a PFIT-implied low-minus-high return spread. This is the return spread we would expect if differences in passive demand and demand multipliers were the *only* forces generating the observed compression in anomaly returns.

Regressing the realized spread on this multiplier/demand-implied spread yields a slope of 1.31 ($t = 5.69$) with an R^2 of 19.3%. In other words, anomalies for which passive demand and estimated multipliers imply more compression do in fact display more compression. As noted above, the long and short legs receive very similar demand shocks on average, so this variation is driven primarily by differences in multipliers across the two legs. A decomposition of the implied spread confirms this: multiplier differences account for 14.1% of the realized cross-anomaly variation, while demand differences account for only 5.4%.

Finally, we provide evidence on one amplification mechanism for our results: short covering. In the spirit of De Long, Shleifer, Summers, and Waldmann [1990], we argue that passive inflows act like noise-trader demand shocks because they can push prices up for reasons unrelated to stock-level fundamentals. When these shocks hit heavily shorted stocks, they move prices against short sellers. Some short sellers may then buy to cover, adding further buying pressure and amplifying the original passive-flow-induced trade.

Consistent with this channel, we find that securities-lending utilization and borrowing fees fall after passive inflows. These declines are larger in anomaly short legs when the short leg has either a higher demand multiplier or more short interest than the long leg. The two measures are also related: across anomalies, the correlation between the multiplier gap and the shorting-intensity gap exceeds 0.3 and is statistically significant. This suggests that heavier shorting may help explain why demand multipliers are larger in anomaly short legs.

Overall, we argue that passive flows compress anomaly returns because short legs have

larger demand multipliers, not because they receive larger demand shocks. This creates a limit to arbitrage for active investors implementing these strategies: even when anomaly signals remain informative about future returns, passive demand during the holding period can substantially reduce their realized long-short performance.

1.1 Related Literature

The paper contributes to three related strands of literature. First, our paper contributes to the large literature studying the effects of passive investing on financial markets. Prior work shows that passive ownership affects corporate governance, volatility, return comovement, price informativeness, securities lending, and the elasticity of stock demand [Appel, Gormley, and Keim, 2016, Ben-David, Franzoni, and Moussawi, 2018, Da and Shive, 2018, Israeli, Lee, and Sridharan, 2017, Glosten, Nallareddy, and Zou, 2021, Coles, Heath, and Ringgenberg, 2022, Sammon, 2025, Palia and Sokolinski, 2019, Von Beschwitz, Honkanen, and Schmidt, 2025, Behmaram and Davis, 2026]. Other recent work studies the mechanics of passive investing itself, including the measurement of passive ownership, who trades against index-fund demand, and how index funds rebalance [Chinco and Sammon, 2024, Pavlova and Sikorskaya, 2023, Sammon and Shim, 2026a,b]. We build on this work by focusing on passive flow-induced *trading*, rather than the *level* of passive ownership. The closest paper to ours is Filippou, He, Li, and Zhou [2023], who study the level of ETF ownership and anomaly returns. We instead show that passive flow-induced trading compresses anomaly returns because the same demand shock has a larger price impact in short-leg stocks.

Second, our paper builds on the literature on flow-induced trading, mechanical demand shocks, and demand-based asset pricing. Coval and Stafford [2007], Lou [2012], and Frazzini and Lamont [2008] use mutual fund flows and other sources of forced trading to study the price effects of inelastic demand, while Shleifer [1986], Harris and Gurel [1986], Kaul, Mehrotra, and Morck [2000], Greenwood [2005], and Chang, Hong, and Liskovich [2015] use index events and related rebalancing shocks to show that demand curves for stocks appear

to slope down. More recent papers on demand-system asset pricing generalize this logic by estimating elasticities and multipliers across broader sets of assets and investors [Koijen and Yogo, 2019, Gabaix and Koijen, 2021, van der Beck, 2022, Haddad, Huebner, and Loualiche, 2025, Gabaix and Koijen, 2024, Davis et al., 2026]. Relative to this literature, we believe PFIT provides a useful empirical demand shock because passive funds mechanically scale lagged holdings in response to flows and because passive flows are difficult to predict and do not forecast fund returns.

Third, the paper connects the anomalies literature with work on limits to arbitrage and short-sale frictions. Anomaly strategies are natural convergence trades: they buy firms that appear cheap and short firms that appear expensive [Fama and French, 2008b,a, Chen and Zimmermann, 2022]. Prior work shows that the short side of these strategies is especially exposed to limits to arbitrage, including short-sale constraints, financing frictions, and costly arbitrage [De Long et al., 1990, Nagel, 2005, Stambaugh, Yu, and Yuan, 2012, 2015, Drechsler and Drechsler, 2014, Hanson and Sunderam, 2014, Muravyev et al., 2025]. More broadly, Lewellen [2011] shows that institutional investors as a group hold portfolios close to the market and show little tendency to bet aggressively on the main characteristics known to predict returns. We add a mechanism through which limits to arbitrage affect the realized payoffs to these strategies: passive inflows create demand shocks during the holding period, and these shocks have especially large price effects in anomaly short legs.

2 Conceptual Framework

This section provides a simple framework for thinking about how passive flow-induced trading can affect the realized returns to long-short anomaly strategies. The framework decomposes the return on each leg of an anomaly portfolio into a component associated with the anomaly itself and a component associated with passive-flow-induced price pressure. This decomposition highlights two channels. Passive flows can affect anomaly returns because the long

and short legs receive different amounts of passive flow-induced trading, and/or because the same amount of passive flow-induced trading has a different price impact across the two legs.

2.1 Anomaly Portfolio Return Decomposition

Consider anomaly strategy a, with long-leg return $R^L_{a,t+1}$ and short-leg return $R^S_{a,t+1}$. We write returns on the two legs as

$$\begin{aligned} R^L_{a,t+1} &= FactorRet^L_{a,t+1} + PFIT^L_{a,t+1} \times M^L_a, \\ R^S_{a,t+1} &= FactorRet^S_{a,t+1} + PFIT^S_{a,t+1} \times M^S_a. \end{aligned} \tag{1}$$

where $FactorRet$ denotes the non-PFIT component of returns. The intended interpretation is that, absent price pressure from passive flows, the long-short difference in $FactorRet$ measures the original or true anomaly return, regardless of whether the underlying mispricing later resolves or builds up over time. However, it should be interpreted broadly, because it is a residual component and therefore also includes all other return components not captured by passive flow-induced trading. $PFIT$ is passive flow-induced trading, measured as a fraction of lagged shares outstanding, and M is the demand multiplier that maps passive flow-induced trading into price changes. The long-short anomaly payoff is therefore

$$R^{LS}_{a,t+1} = R^L_{a,t+1} - R^S_{a,t+1}. \tag{2}$$

Substituting the leg-level returns into the long-short portfolio return gives

$$\begin{aligned} R^{LS}_{a,t+1} = &\left(FactorRet^L_{a,t+1} - FactorRet^S_{a,t+1}\right) \\ &+ \left(PFIT^L_{a,t+1} \times M^L_a - PFIT^S_{a,t+1} \times M^S_a\right). \end{aligned} \tag{3}$$

This expression makes clear that passive flows can affect anomaly profits through two

channels. The first is an exposure channel: the long and short legs may experience different levels of passive flow-induced trading, $PFIT^L \neq PFIT^S$. If passive inflows are systematically larger in the short leg, then even if multipliers are similar across the long and short legs, passive flows could compress long-short anomaly returns. The second is a multiplier channel: the same passive demand shock may have different price impact across the two legs, $M^L \neq M^S$. If the short leg has a larger multiplier, then passive inflows could compress anomaly returns even if the long and short legs have similar PFIT exposure.

2.2 Benchmark Cases

This framework is useful because it nests several natural null hypotheses. The first is the frictionless benchmark in which demand curves for individual stocks are flat [Scholes, 1972]. In this case with perfectly elastic demand, $M = 0$ for all stocks and passive flow-induced trading has no effect on prices. In this case, passive flows should not affect anomaly returns, and long-short spreads are determined only by the factor component of returns.

A second benchmark allows demand curves to slope down, but assumes that prices are fully forward looking. In this case, predictable components of passive demand should be incorporated into prices *when they become predictable*, not when the trades are implemented. A dramatic example of this logic is apparent in the case of index additions: once an addition becomes predictable or is announced, prices move before the effective date, even though index funds do not complete the associated trading until the implementation date [Pegoraro, Sammon, and Shim, 2026]. Realized PFIT may therefore have little contemporaneous return effect if the associated demand was already anticipated.

The third benchmark combines downward-sloping demand with limited arbitrage or imperfect anticipation of the demand shocks [Shleifer, 1986]. With balance-sheet constraints, short-sale frictions, disagreement, and other limits to arbitrage, passive demand can generate contemporaneous price pressure and affect realized anomaly returns during the holding period. These effects should be largest where arbitrage is most costly. For long-short anomaly

strategies, this case is especially relevant because the short leg is often where arbitrage is most constrained [Muravyev et al., 2025].

2.3 Empirical Tests

The framework guides three sets of empirical tests. First, we ask whether anomaly returns are lower in stocks with high passive flow-induced trading. This test establishes whether passive demand shocks are associated with lower realized long-short returns, without distinguishing whether it is driven by the exposure channel or the multiplier channel.

Second, we ask whether the long and short legs differ in their exposure to passive flow-induced trading, i.e., we compare $PFIT^L$ and $PFIT^S$. If passive flows compress anomaly returns because the short leg receives more passive demand, then we should observe larger PFIT in the short leg than in the long leg.

Third, we estimate whether the long and short legs differ in their demand multipliers, i.e., we compare M^L and M^S. If passive flows compress anomaly returns because the same demand shock has larger price impact on the short leg, then the short-leg multiplier should be larger than the long-leg multiplier. Because we estimate these objects for each of 146 anomalies, we can also ask whether the anomalies with the largest estimated multiplier gaps are the ones whose realized spreads compress the most.

Our main tests use realized PFIT during the holding period of each anomaly as our empirical measure of the unanticipated inelastic demand shock. This object is useful because, as we show below, passive flow-induced trading is difficult to predict out of sample and does not appear to forecast fund returns. Realized PFIT is therefore close to an innovation in demand, which is what the third benchmark case requires.

3 Data and Variable Construction

This section describes the data sources and the anomaly sample, explains the construction of passive flow-induced trading, summarizes its cross-sectional properties, and studies the predictability and informativeness of passive fund flows.

3.1 Data Sources

Our sample combines stock returns and shares outstanding from CRSP, fund returns and total net assets from the CRSP mutual fund database, fund-stock holdings from Thomson Reuters S12, anomaly signals from the Chen and Zimmermann [2022] open-source asset pricing database, factor returns from Ken French's data library, firm characteristics from the WRDS Financial Ratios Suite, and short-sale information from Compustat and Markit. The sample period for the main portfolio tests is 2002Q2 through 2023Q4, which yields 86 holding quarters after the required merges and lags. We restrict the stock sample to CRSP common stocks (share codes 10 and 11) traded on major exchanges (exchange codes 1, 2, and 3) with nonmissing next-quarter returns and positive market equity. Daily returns are adjusted for delisting and compounded to the quarterly frequency. We classify funds as passive or active following the procedure in Appel et al. [2016], compute fund flows using total net assets and fund returns from CRSP, and clean the S12 holdings data following Sammon and Shim [2023]. We merge fund flows to lagged fund-stock holdings to construct flow-induced trading measures, as described in the next section. To be included in our sample, a stock must be held by at least one index fund before the anomaly holding period. Given the breadth of total-market index funds such as Vanguard's Total Stock Market Index Fund (Ticker: VTI), this restriction mainly excludes very small or low-float stocks that account for a tiny share of aggregate market capitalization.

The anomaly signals come from the Chen and Zimmermann [2022] open-source asset pricing database, which provides signed signals so that higher values point toward higher

expected returns. Rather than pre-selecting a small number of signals, we work with a broad anomaly universe. We start from the full set of signed predictors in Chen and Zimmermann [2022] and then impose two coverage requirements that are specific to our research design, because our main tests require portfolios formed on the interaction of passive demand and anomaly signals. First, all thirty positive-PFIT-tercile by anomaly-decile cells must exist and must contain at least twenty stocks on average across the quarters in which they are available. Second, within each of the three positive-PFIT terciles, the top and bottom anomaly deciles must both have nonmissing equal- and value-weighted returns in at least twenty common quarters. These filters leave 146 anomaly signals, which is the universe used throughout the paper.

Each quarter, stocks are sorted into anomaly deciles using the signal available at the end of the formation quarter. Throughout the paper, decile 1 (D1) is the short leg and decile 10 (D10) is the long leg, so the long-short anomaly return is D10 minus D1. We also compute the fractional rank of each signal, $z^a_{i,t-1} \in [0, 1]$, for use in the stock-level regressions. Following Chen and Zimmermann [2022], our baseline portfolio returns are equal-weighted; value-weighted results are reported in the online appendix. Unless otherwise noted, portfolio returns and alphas are annualized percentage points, equal to 400 times a quarterly decimal return.

3.2 Passive Flow-Induced Trading

We measure passive demand shocks using passive flow-induced trading, or PFIT. The measure maps fund-level flows into stock-level demand shocks under the assumption that index funds mechanically scale their lagged holdings in response to inflows and outflows [Sammon and Shim, 2026b]. We compute fund flows as a percentage of total net assets (TNA) following Chevalier and Ellison [1997]:

$$Flow_{f,t} = \frac{TNA_{f,t} - TNA_{f,t-1}(1 + r_{f,t})}{TNA_{f,t-1}}. \tag{4}$$

Under this definition, $Flow_{f,t}$ is the change in TNA that cannot be explained by fund returns $r_{f,t}$. Under proportional scaling, the implied flow-induced change in shares held for fund f in stock i is

$$\Delta Shares^{FIT}_{f,i,t} = Flow_{f,t} \times Shares_{f,i,t-1}. \tag{5}$$

In other words, if an index fund experiences a 1% inflow (relative to TNA), it is expected to scale up all existing positions in the fund by 1% *in share terms*. This scaling *must* hold for index funds with very strict index-tracking mandates, e.g., SPY, which is a unit investment trust and therefore has no flexibility in how it replicates the S&P 500. As shown by Sammon and Shim [2026b], however, this proportional scaling in share terms is also a good description of how non-value-weighted index funds adjust their positions in response to flows.

We then aggregate flow-induced trading across all index funds holding stock i to compute a stock-quarter-level flow-induced demand shock

$$\Delta Shares^{PFIT}_{i,t} = \sum_{f \in Passive} \Delta Shares^{FIT}_{f,i,t}. \tag{6}$$

Finally, to make this quantity comparable across stocks and across time, we scale it by lagged shares outstanding:

$$PFIT_{i,t} = 100 \times \frac{\Delta Shares^{PFIT}_{i,t}}{SharesOut_{i,t-1}}. \tag{7}$$

PFIT is therefore a stock-quarter demand shock measured as a percentage of shares outstanding. Positive values correspond to passive inflow-induced buying pressure, while negative values correspond to passive outflow-induced selling pressure. Shares held and shares outstanding are adjusted using CRSP's cumulative share adjustment factor before lagging.

To better understand how PFIT is constructed, consider the simple example in Table 1. Recall from Equation 6 that f indexes funds and i indexes stocks. In this example, the stock i is AAPL, and it is held by two index funds, SPY and VTI. For each fund-stock

pair, we multiply the fund's flow, measured as a percentage of lagged TNA, by the fund's lagged shares held of the stock. This gives the fund's implied flow-induced trade in shares, $\Delta Shares^{FIT}_{f,i,t}$. We then aggregate these implied trades across all index funds holding the stock to obtain total passive flow-induced trading in shares, $\Delta Shares^{PFIT}_{i,t}$, and scale by the stock's lagged shares outstanding to obtain $PFIT_{i,t}$.

In the example, SPY has a 5% inflow and holds 50,000 shares of AAPL at the end of 2026Q1, implying a flow-induced purchase of $0.05 \times 50{,}000 = 2{,}500$ shares in 2026Q2. VTI has a 10% inflow and holds 100,000 shares of AAPL at the end of 2026Q1, implying a flow-induced purchase of $0.10 \times 100{,}000 = 10{,}000$ shares. Aggregating across the two funds gives 12,500 shares of passive flow-induced trading. Because AAPL has 1,000,000 lagged shares outstanding, $PFIT_{i,2026Q2} = 100 \times 12{,}500/1{,}000{,}000 = 1.25$ percent of shares outstanding.

Table 1. Example Construction of Passive Flow-Induced Trading

Fund	Stock	$SharesOut_{i,2026Q1}$	$Shares_{f,i,2026Q1}$	$Flow_{f,2026Q2}$	$\Delta Shares^{FIT}_{f,i,2026Q2}$
SPY	AAPL	1,000,000	50,000	5%	2,500
VTI	AAPL	1,000,000	100,000	10%	10,000
				$\Delta Shares^{PFIT}_{i,2026Q2}$	12,500
				$PFIT_{i,2026Q2}$	1.25%

This table gives a simple example of how passive flow-induced trading is constructed for one stock, indexed by i, held by two index funds, indexed by f. For each fund-stock pair, the implied flow-induced trade equals the fund's flow multiplied by its lagged shares held. Aggregating across funds gives total passive flow-induced trading in shares. PFIT scales this quantity by lagged shares outstanding, so $PFIT_{i,2026Q2} = 100 \times 12{,}500/1{,}000{,}000 = 1.25$ percent.

PFIT isolates flow-induced scaling of existing holdings. Of course, passive fund trades can also arise from extensive-margin rebalancing (i.e., when stocks enter or leave an index) and from intensive-margin rebalancing (i.e., when firms issue/repurchase equity or index providers change float adjustment factors) [Sammon and Shim, 2026b]. These rebalancing trades are important, and empirically explain the majority of position changes for index funds [Sammon and Shim, 2026a]. These rebalancing trades, however, are conceptually distinct from the flow-induced trading we study. Rebalancing trades may be mechanically related

to characteristics that predict returns. For example, index funds mechanically rebalance toward firms that issue equity, and equity issuance has historically predicted low returns [Pontiff and Woodgate, 2008].

Importantly, there is substantial stock-quarter-level variation in PFIT. When people think of index funds, they often have in mind very broad-based funds such as total-market funds or S&P 500 index funds. Even broad-based index funds, however, do not hold identical portfolios. A total-market fund, an S&P 500 fund, and a Russell 1000 fund have different investment universes and different float-adjustment factors applied to each security. As a result, the same dollar of investor flow can generate different stock-level demand depending on which broad-based fund receives the flow.

This heterogeneity is even larger once we include the full passive fund universe. Many index funds track sector, style, factor, thematic, or specialized benchmarks. These funds hold more concentrated portfolios and often have more volatile flows. Flow-induced trading from these funds, therefore, can generate large variation in flow-induced trading demand. This variation is central to our empirical design because it allows us to compare anomaly returns across stocks that receive different amounts of passive flow-induced trading during the holding period.

We also construct active flow-induced trading, or AFIT, using the same procedure. Specifically, we compute active fund flows, apply those flows to each active fund's lagged stock holdings, aggregate the implied trades across all active funds that hold the stock, and scale by lagged shares outstanding. AFIT is a useful comparison because flow-induced trading from active mutual funds is widely used to identify demand shocks and estimate price pressure [Frazzini and Lamont, 2008, Lou, 2012]. The interpretation, however, is cleaner for PFIT than AFIT. Index funds largely must scale existing positions in response to flows, while active managers have discretion over how to invest inflows and meet redemptions. As a result, predicted active flow-induced trading based on lagged holdings can differ from actual active trading. Consistent with this distinction, Lou [2012] show that active funds scale

existing positions nearly one-for-one in response to outflows, but only partially in response to inflows. This wedge is important because active managers' discretionary trades may be correlated with expected returns.

Table 2 reports summary statistics for PFIT. The sample contains 322,276 stock-quarter observations with nonzero PFIT for stocks held by at least one index fund.[1] The average PFIT observation is 0.13% of shares outstanding, with a standard deviation of 0.43%, and 70.4% of stock-quarters have positive passive flow-induced trading. The distribution is right-skewed: the fifth percentile is −0.19% and the ninety-fifth percentile is 0.65%, more than five times the median of 0.04%.

Panel B reports time-series averages of cross-sectional correlations between PFIT and other variables. We separately examine positive PFIT, corresponding to passive inflows, and the absolute value of negative PFIT, corresponding to passive outflow magnitude. PFIT is modestly positively correlated with firm size: larger firms experience somewhat more passive inflow- and outflow-induced trading. Passive inflows are essentially uncorrelated with active flow-induced trading (0.002), while passive outflow magnitude has only a weak negative correlation with AFIT (−0.084). The low correlation with AFIT is important for estimating demand multipliers. If passive and active flow-induced trading were strongly correlated, the return response to PFIT could partly reflect omitted active-fund demand. For example, if investors simultaneously allocated to active and passive value funds, stocks held by both groups would experience correlated active and passive buying pressure. A regression of returns on PFIT alone would then overstate the demand multiplier by attributing the effect of correlated active-fund trading to passive flows.

The level of passive ownership is positively correlated with both passive inflows (0.53) and passive outflows (0.44). The correlation between PFIT and passive share is perhaps unsurprising. Stocks with higher passive ownership are held more heavily by index funds, so a given pattern of index-fund flows mechanically generates more passive flow-induced trading

[1] In unreported results, we find that including zero-PFIT observations strengthens our main portfolio sort results. The online appendix reports portfolio results for the zero-PFIT group separately.

Table 2. PFIT Summary Statistics, Correlations, and Firm Characteristics

Panel A: PFIT summary statistics

Observations	322,276	Mean	0.130
Std. dev.	0.434	Median	0.039
p5	−0.192	p95	0.654
Negative observations	95,403 (29.6%)	Positive observations	226,873 (70.4%)

Panel B: Average quarterly cross-sectional correlations

Flow measure	Log(ME)	Passive ownership	AFIT
Passive inflows	0.212	0.527	0.002
	[11.98]	[25.29]	[0.13]
Passive outflow magnitude	0.128	0.440	−0.084
	[5.83]	[26.20]	[−8.22]

Panel C: Firm characteristics across positive-PFIT terciles

Characteristic	Low	Middle	High
PFIT	0.039	0.153	0.412
Cyclically adjusted P/E	8.993	14.956	18.829
Gross profit/assets	0.230	0.268	0.278
Return on assets	−0.046	0.048	0.084
Debt/assets	0.209	0.228	0.233
Log market equity	19.273	20.613	21.135
AFIT	−0.075	−0.069	−0.010
Amihud illiquidity ($\times 10^6$)	4.740	1.001	0.347

This table reports summary statistics, correlations, and firm characteristics for passive flow-induced trading. PFIT is stock-level passive flow-induced trading, scaled by lagged shares outstanding and expressed in percentage points. The sample contains stock-quarter observations with nonzero PFIT from 2002Q2 through 2023Q4. Panel A reports the distribution of PFIT. Panel B reports time-series averages of quarterly cross-sectional Pearson correlations, computed separately for positive PFIT (passive inflows) and for the absolute value of negative PFIT (passive outflow magnitude). Log(ME) is log market capitalization, passive ownership is the fraction of shares outstanding held by index funds, and AFIT is active flow-induced trading. Newey-West t-statistics with four lags, based on the time series of quarterly correlations, are reported in brackets. Panel C reports characteristics of stocks in each positive-PFIT tercile. Each characteristic is time-series average of mean portfolio characteristic.

in those stocks. It is nevertheless important to distinguish the two objects, both conceptually and empirically. Passive ownership is a stock-level state variable that changes slowly and has been linked to a wide range of outcomes, including governance, comovement, volatility, lending-market conditions, and price informativeness [Appel et al., 2016, Ben-David et al., 2018, Von Beschwitz et al., 2025]. PFIT is a flow: it is the marginal, time-varying demand that arrives during a particular quarter. Two stocks with identical passive ownership can have very different PFIT if the index funds that hold them receive different flows, and a stock's PFIT varies substantially from quarter to quarter while its passive share does not. Because our mechanism operates through price pressure realized *during* the anomaly holding period, PFIT is the more direct measure. The distinction also matters for identification: a research design based on passive ownership levels would compare permanently different stocks, whereas our design compares the same kinds of stocks across quarters in which they happen to receive more or less mechanical passive buying. We nevertheless control directly for passive ownership in the stock-level regressions in Section 4.3, and Section 4.1 shows that the long and short legs of the average anomaly have nearly identical passive ownership.

Panel C describes what kinds of firms are most exposed to passive-flow-induced trading. Moving from the low- to the high-PFIT tercile, average PFIT rises from 0.04 to 0.41% of shares outstanding. High-PFIT firms are larger, more profitable, more expensive on a cyclically adjusted earnings basis, slightly more levered, and substantially more liquid: average Amihud illiquidity falls by more than an order of magnitude from the low to the high tercile. Active flow-induced trading is mildly negative in all three groups and does not vary much across them. These are descriptive differences, and they motivate the stock-level controls in Section 4.3.

3.3 Passive Flow Predictability and Informativeness

Our conceptual framework treats realized PFIT as an unanticipated demand shock that is unrelated to stock-level fundamentals. Two properties of passive fund flows are needed for

this interpretation, and each has a distinct asset-pricing implication.

The first is *predictability*. If passive flows were easy to forecast, then in a forward-looking market the associated demand would be impounded in prices when it became predictable rather than when the index funds actually traded, as in the second benchmark case of Section 2. Realized PFIT would then be a poor proxy for a contemporaneous demand shock, and a contemporaneous return-on-PFIT regression would confound the price impact of the current trade with revisions in expected future demand. In that case, we would *overstate* the demand multiplier.

The second is *informativeness*. If passive flows carried information about the securities index funds buy and sell, then the associated price movements would not represent inelastic demand at all. Under that interpretation, higher returns on high-PFIT stocks would reflect information and not price pressure thus would not be able to cleanly identify demand multiplier.

We evaluate both properties directly. To assess flow predictability, we follow the logic of Lou [2012] and forecast fund flows using lagged flows and lagged fund returns. Specifically, for each fund type and quarter we estimate the cross-sectional regression

$$Flow_{f,t} = \alpha_t + \sum_{j=1}^{3} \rho_{j,t} Flow_{f,t-j} + \sum_{j=1}^{3} \gamma_{j,t} R^{fund}_{f,t-j} + \varepsilon_{f,t}, \tag{8}$$

separately for passive and active funds, requiring at least fifty complete funds per quarter. The lagged-flow terms capture persistence in investor flows, which may arise from gradual investor reallocation or from the inclusion of a particular fund on retirement-plan menus. The lagged-return terms capture the well-known flow-performance relation for mutual funds. Panel A of Table 3 reports time-series averages of the quarterly coefficients with Newey-West t-statistics. Passive flows are statistically related to their own lags and to lagged fund returns, but the explanatory power is modest: the average in-sample R^2 is 0.089. Active flows are considerably more persistent and more predictable, with an in-sample R^2 of 0.193 and much larger first-lag coefficients.

Table 3. Predictability of Passive and Active Fund Flows

	Passive funds	Active funds
Panel A: Quarterly Fama-MacBeth forecasting coefficients		
Flow, $t-1$	0.057	0.249
	(3.10)	(34.25)
Flow, $t-2$	0.098	0.137
	(7.43)	(21.27)
Flow, $t-3$	0.078	0.100
	(8.25)	(19.61)
Fund return, $t-1$	0.233	0.313
	(4.46)	(12.28)
Fund return, $t-2$	0.018	0.129
	(0.41)	(6.44)
Fund return, $t-3$	−0.065	0.052
	(−1.14)	(2.83)
Constant	0.022	−0.003
	(4.58)	(−1.11)
In-sample R^2	0.089	0.193
Observations	106,260	389,240
Panel B: Pooled out-of-sample calibration		
Out-of-sample R^2	0.021	0.147
Pearson correlation	0.144	0.383
Rank correlation	0.197	0.444

This table reports fund-quarter flow forecasting results. The dependent variable is fund flow, measured as a fraction of lagged total net assets. Independent variables include three lags of fund flow and three lags of quarterly value-weighted fund returns. Panel A reports time-series means of quarterly cross-sectional coefficients estimated separately for passive and active funds, with Newey-West t-statistics with four lags in parentheses. Panel B reports the pooled out-of-sample calibration regression of realized flows on predicted flows, where predicted flows for quarter t use the average of the coefficients estimated in quarters $t-8$ through $t-1$. The sample is from 2002Q2 to 2023Q4.

In-sample fit overstates what an investor could have anticipated in real time, so we also evaluate the forecasts out of sample. For each target quarter, we estimate Equation (8) separately in each of the prior eight quarters, average the coefficients across those quarters within fund type, and apply the averaged coefficients to each fund's lagged flows and returns. Estimating the rolling coefficients separately for passive and active funds allows the determinants of flows to differ across the two groups. We then run the pooled calibration regression of realized flows on predicted flows. Panel B of Table 3 reports the results, and Figure 1 plots them. For passive funds, predicted flows explain very little of realized flows: the out-of-sample R^2 is 2.1%, the Pearson correlation is 0.144, and the rank correlation is 0.197. For active funds, predictability is much higher, with an out-of-sample R^2 of 14.7% and a Pearson correlation of 0.383. The calibration slope is also much flatter for passive funds (0.55) than for active funds (0.77), indicating that even the ranking that the model does produce is attenuated in levels. Figure OA5 in the online appendix shows that the same conclusion holds when fund-quarters are weighted by total net assets, so this is not a small-fund phenomenon.

We have also considered richer forecasting models that include holdings-weighted characteristics of each fund's portfolio, such as size, book-to-market, momentum, profitability, investment, and beta, in the spirit of Van der Beck [2021]. These variables do not improve out-of-sample fit in our setting. In fact, the out-of-sample R^2 declines when we add the full set of holdings characteristics, suggesting that these models overfit in sample. We therefore use the simpler lagged-flow and lagged-return model throughout.

We next turn to informativeness. The natural fund-level test is whether flows are timed ahead of fund performance. If investors direct money to index funds because they know something about the securities those funds hold, flows should forecast subsequent fund returns. We therefore regress fund f's return in quarter $t+1$ on current and four lags of fund flows and returns, separately for passive and active funds, in equal- and value-weighted specifications, clustering by fund and quarter. Table OA1 in the online appendix reports

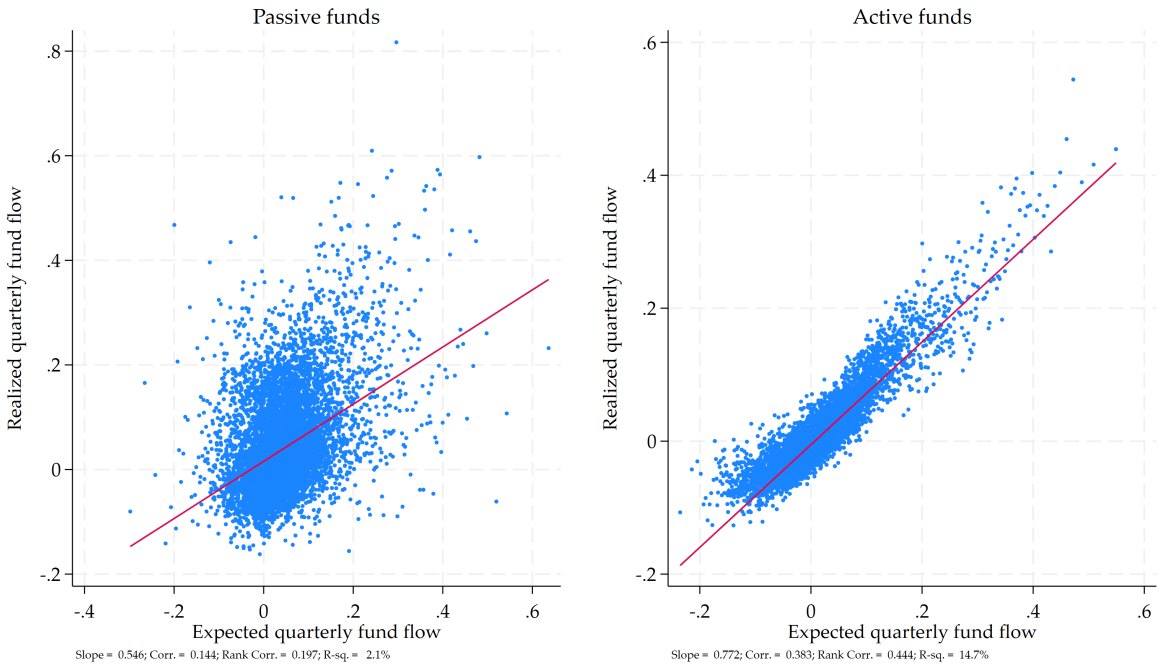


Figure 1. Predicted and realized fund flows: out-of-sample fit. This figure plots realized quarterly fund flows against out-of-sample predicted flows, separately for passive funds (left panel) and active funds (right panel). Predicted flows for quarter t are constructed by estimating the cross-sectional forecasting regression in Equation (8) separately in each of the eight preceding quarters, averaging the coefficients, and applying them to each fund's three lagged flows and three lagged returns. Plotted points are equal-count bins used for rendering only; the fitted line, slope, correlations, and R^2 reported below each panel are estimated on the full unbinned sample of fund-quarters with heteroskedasticity-robust standard errors. Fund flows are expressed as a fraction of lagged total net assets. The sample is from 2002Q2 to 2023Q4.

the results. Current flows do not significantly predict next-quarter fund returns in any of the four specifications. A few isolated lag coefficients are statistically significant, but their signs are not stable across weighting choices, and the regressions explain almost none of the variation in subsequent fund returns.

We read this evidence as saying that passive flows are close to the demand shocks our framework has in mind. They are difficult to anticipate, so prices are unlikely to have impounded them before index funds trade, and they do not appear to contain information about subsequent performance, so the associated price movements are unlikely to be a repricing of fundamentals. This is the sense in which PFIT is a useful proxy for an inelastic demand shock: it isolates buying and selling pressure that arrives for reasons largely unrelated to the expected returns of the stocks being traded.

4 Effect of PFIT on Long-Short Anomaly Returns

This section presents the main evidence that passive flow-induced trading acts as a limit to arbitrage for long-short anomaly strategies. We first describe the unconditional properties of our 146 anomalies and show that the long and short legs of the average anomaly receive nearly identical passive demand. We then show that long-short returns are substantially lower among stocks that experience high passive inflow-induced trading during the holding period, and that this pattern holds across the vast majority of individual anomalies. We then use the conceptual framework in Section 2 to distinguish between the exposure and multiplier channels. The evidence points to a multiplier channel: passive inflows raise returns in both legs, but the price pressure is larger in the short leg, where arbitrage is more constrained. Finally, Fama-MacBeth regressions show that the result survives stock-level controls, and longer-horizon tests show that the return effect is concentrated in the holding quarter.

4.1 Results for Single-Sorted Anomaly Portfolios

We begin with single-sort anomaly portfolios. At the end of each quarter, we sort stocks into deciles based on the anomaly signal available at that time, without conditioning on passive flows. We hold the portfolios for one quarter and compute equal-weighted returns during the holding quarter. We label the low-signal decile D1 and the high-signal decile D10. Throughout the paper, D1 is the short leg, D10 is the long leg, and the long-short anomaly return is D10 minus D1. Because we work with 146 anomalies, we follow Muravyev et al. [2025] and summarize the evidence with cross-anomaly averages: for each anomaly we estimate the object of interest in the time series, and we then report the mean of those estimates across anomalies together with a robust cross-anomaly t-statistic.

Table 4 reports the results. Panel A serves two purposes. First, it verifies that the anomaly universe earns economically meaningful long-short returns over our sample. The average short-leg excess return is 7.42% and the average long-leg excess return is 11.70%, so the average long-short spread is 4.28% with a t-statistic of 10.20. Risk adjustment strengthens the spread slightly: the average CAPM alpha of the long-short portfolio is 4.95% ($t = 8.24$). These magnitudes are smaller than those typically reported in the original anomaly papers, which is expected given that our sample is recent and that most of these signals were published well before 2002 [Muravyev et al., 2025]. They are nevertheless large enough to make the strategies worth implementing, which is what our tests are about.

Second, and more importantly, Panel B shows that the two legs of the average anomaly are almost indistinguishable in their exposure to passive investing. Average realized PFIT is 0.121% of shares outstanding in the short leg and 0.118% in the long leg. The difference of 0.004% is economically trivial and not significant. Average passive ownership is 7.23% in the short leg and 7.22% in the long leg, a difference that is statistically indistinguishable from zero. Active flow-induced trading is also similar across the two legs. This is the first piece of evidence against a simple exposure explanation: if passive flows compressed anomaly spreads because index funds simply buy more of the stocks anomaly strategies short, we should see

Table 4. Single-Sorted Anomaly Portfolios and Their Characteristics

Panel A: Cross-anomaly mean portfolio outcomes

Outcome	D1	D10	D10−D1	N anomalies
Mean excess return	7.424	11.703	4.279	146
	(23.40)	(47.66)	(10.20)	
CAPM alpha	−7.786	−2.841	4.945	146
	(−18.04)	(−9.73)	(8.24)	

Panel B: Average anomaly-decile profiles

Variable	D1	D10	D1−D10	t(difference)
PFIT	0.121	0.118	0.004	1.74
Passive ownership	7.231	7.218	0.013	0.24
Log market equity	19.788	19.622	0.166	5.63
AFIT	−0.078	−0.096	0.019	0.44
Compustat short-interest ratio	0.064	0.051	0.013	5.32
Markit utilization	22.623	17.971	4.652	29.53
Markit indicative fee	0.067	0.053	0.014	10.02
Amihud illiquidity ($\times 10^6$)	4.528	6.474	−1.947	−4.37

This table reports single-sorted anomaly portfolio outcomes and portfolio characteristics averaged across the 146 anomalies in our sample. At the end of quarter t, stocks are sorted into deciles based on the anomaly signal available at that time, and equal-weighted returns are measured during quarter $t + 1$. D1 is the short leg and D10 is the long leg. Panel A reports cross-anomaly means of anomaly-level average excess returns and CAPM alphas, expressed in annualized percentage points; anomaly-level estimates use heteroskedasticity-robust standard errors, and the reported t-statistics in parentheses test the cross-anomaly mean. Panel B reports average characteristics of the two legs, computed by averaging within quarter, anomaly, and decile, then across anomalies, and finally over time. PFIT passive ownership, and AFIT are expressed in percentage points. Compustat short-interest ratio is short interest scaled by shares outstanding, Markit utilization is the share of lendable shares on loan expressed in percentage points, and the Markit indicative fee is Markit's estimate of the borrow cost faced by a hedge fund. Newey-West t-statistics with four lags are reported in parentheses. Markit variables are available beginning in the 2010Q2 holding quarter. The sample is from 2002Q2 to 2023Q4.

it here, and we do not.

The two legs do differ along the dimensions emphasized by the limits-to-arbitrage literature. Short-leg stocks are more heavily shorted and harder to short: the average short-interest ratio is 0.064 in D1 versus 0.051 in D10, Markit utilization is 22.6% versus 18.0%, and indicative borrow fees are correspondingly higher. Short-leg stocks are also slightly larger and, on the Amihud measure, somewhat more liquid than long-leg stocks. Thus, the salient asymmetry between the legs is not how much passive demand they receive, nor how easily they trade, but how constrained the short side is. This is exactly the asymmetry that the multiplier channel requires.

4.2 Results for Double-Sorted Anomaly Portfolios

To ask whether anomaly returns depend on realized passive demand, we use double sorts. Each quarter, we restrict to stocks with positive PFIT, because we want to separate passive inflow-induced buying pressure from passive outflow-induced selling pressure and because passive inflows account for about 70% of observations. We first sort stocks into low-, middle-, and high-PFIT terciles using realized PFIT during the holding quarter, denoted P1, P2, and P3. Within each PFIT tercile, we then sort stocks into anomaly deciles using the signal available at portfolio formation. We compute equal-weighted portfolio returns during the holding quarter, estimate anomaly-level means and CAPM alphas, and then report cross-anomaly means.

Table 5 reports our main empirical finding. Averaging across the 146 anomalies, the long-short return declines monotonically across PFIT terciles, from 5.01% in P1, to 2.67% in P2, to 0.60% in P3. The spread in the high-PFIT tercile is statistically indistinguishable from zero at conventional levels ($t = 1.20$), while the spread in the low-PFIT tercile is large and highly significant. The low-minus-high difference is 4.41% with a cross-anomaly t-statistic of 7.15. Risk adjustment does not change the picture: CAPM alphas fall from 5.63% in P1 to 1.07% in P3, a difference of 4.56% ($t = 8.36$). To put the magnitude in perspective, the

compression of 4.41% is essentially as large as the average unconditional long-short spread of 4.28% reported in Table 4. Anomaly strategies earn most of their average return in the stock-quarters in which passive demand happens to be low.

Table 5. PFIT and Anomaly Double-Sorted Portfolio Returns

Outcome	Portfolio	Low	Middle	High	Low−High
Return	Short leg	9.655	13.970	24.422	−14.767
		(25.08)	(49.19)	(68.25)	(−33.80)
Return	Long leg	14.668	16.638	25.023	−10.355
		(37.26)	(60.38)	(82.86)	(−27.41)
Return	Long-short	5.013	2.668	0.600	4.412
		(8.60)	(6.67)	(1.20)	(7.15)
CAPM alpha	Short leg	−5.345	−0.316	9.141	−14.487
		(−14.25)	(−1.24)	(36.42)	(−38.51)
CAPM alpha	Long leg	0.284	3.040	10.211	−9.927
		(0.79)	(12.60)	(55.98)	(−26.32)
CAPM alpha	Long-short	5.629	3.356	1.070	4.559
		(9.65)	(9.32)	(3.30)	(8.36)

This table reports equal-weighted returns and CAPM alphas for portfolios conditionally double sorted on realized PFIT and anomaly signals. Each quarter, stocks with positive realized PFIT during the holding quarter are first sorted into terciles based on PFIT, and then, within each tercile, into deciles based on the anomaly signal available at the end of the previous quarter. The short leg is D1, the long leg is D10, and the long-short portfolio is D10 minus D1. For each anomaly and PFIT tercile, mean returns and CAPM alphas are estimated in the time series with heteroskedasticity-robust standard errors. The table reports robust cross-anomaly means, with cross-anomaly t-statistics in parentheses. Returns and alphas are annualized percentage points. The sample is from 2002Q2 to 2023Q4.

Table 5 also shows *why* the spreads compress, and Figure 2 displays the same information graphically. Passive inflows raise returns in both legs. Moving from P1 to P3, the average short-leg return rises from 9.66% to 24.42%, an increase of 14.77%, while the average long-leg return rises from 14.67% to 25.02%, an increase of 10.36%. Both legs are pushed up, so passive inflows are not simply a signal about which stocks are good investments. But the short leg is pushed up more, and the difference between the two increases is exactly the 4.41% compression in the long-short spread reported in the same table. Recall from Table 4 that the two legs receive nearly identical passive demand, and that within each tercile of this

double sort they are compared at similar levels of realized PFIT. The asymmetry therefore has to come from the price response per unit of demand rather than from the amount of demand. This is the multiplier channel, which we estimate directly in Section 5.1.

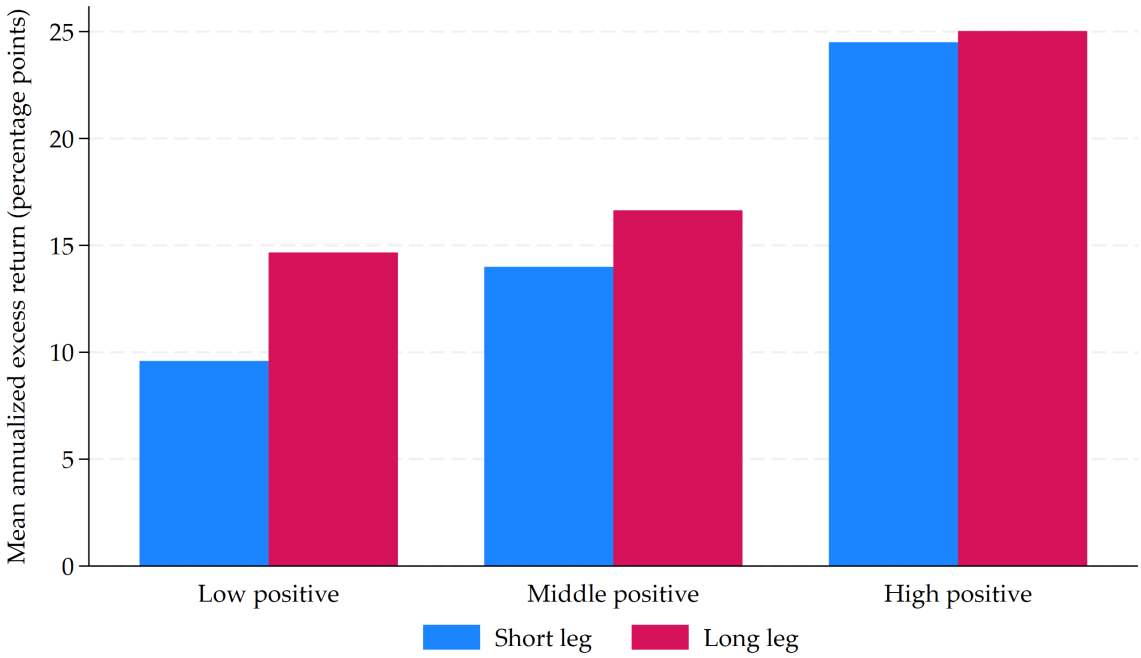


Figure 2. PFIT and anomaly double-sorted portfolio returns. This figure plots average short-leg and long-leg portfolio returns across realized positive-PFIT terciles. Each quarter, stocks with positive realized PFIT during the holding quarter are sorted into PFIT terciles and then, within each tercile, into anomaly deciles using the signal available at portfolio formation. The short leg is D1 and the long leg is D10. Bars report cross-anomaly averages of anomaly-level equal-weighted excess returns during the holding quarter, expressed in annualized percentage points. The sample is from 2002Q2 to 2023Q4.

Because we estimate this comparison for 146 separate signals, we can also ask whether the average masks substantial disagreement across anomalies. Figure 3 plots the 146 anomaly-level low-minus-high differences in ascending order. The result is broad-based: 69.9% of anomalies have a positive difference, and the cross-anomaly mean of 4.41% is highly significant. The figure also shows real heterogeneity. Anomaly-level differences range from about -9% to $+38\%$, with a long right tail: a modest number of signals display compression far larger than the average, while roughly thirty percent display none. This dispersion is in-

formative rather than a nuisance. If passive flows compress spreads through a multiplier channel, the anomalies whose short legs are most sensitive to passive demand should be the ones whose spreads compress the most. We use exactly this cross-anomaly variation in Section 5.1.

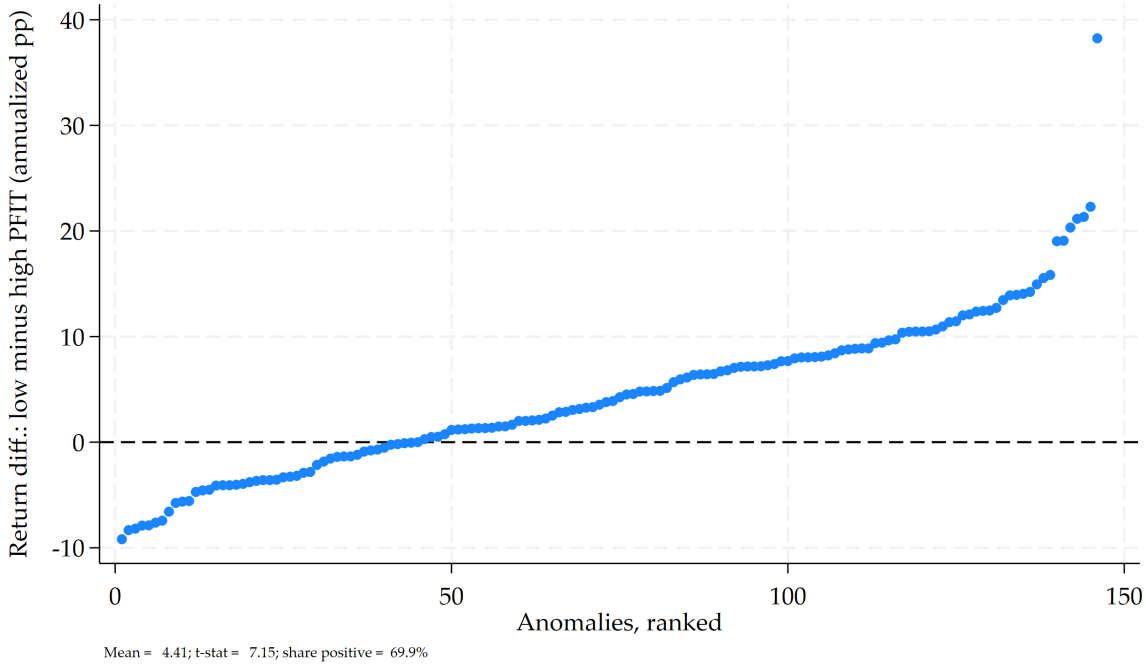


Figure 3. Low minus high PFIT spread return distribution. This figure ranks the 146 anomaly-level estimates of the difference between the long-short return in the low-positive-PFIT tercile and the long-short return in the high-positive-PFIT tercile. Portfolios are formed as in Table 5 using equal-weighted returns, and each point is one anomaly. Differences are expressed in annualized percentage points. The reported mean, cross-anomaly t-statistic, and share of positive differences are computed across the 146 anomalies. The sample is from 2002Q2 to 2023Q4.

The main result is not an artifact of equal weighting or of the restriction to positive-PFIT observations, and the online appendix reports the corresponding robustness checks. Value weighting delivers a very similar conclusion: as shown in online appendix Table OA2 and Figures OA1 and OA2, the value-weighted long-short return falls from 3.59% in the low-PFIT tercile to -0.71% in the high-PFIT tercile, a low-minus-high difference of 4.30% ($t = 6.34$) that is positive for 69.2% of anomalies. The value-weighted CAPM-alpha difference is 3.72%

($t = 6.03$). Moving from low to high PFIT raises value-weighted short-leg returns by 20.18% versus 15.88% for the long leg, so the leg-level effects of passive demand are larger among the large stocks that dominate value-weighted portfolios, while the difference between the two legs is again exactly the compression of the spread.

Online appendix Table OA3 adds the negative-PFIT and exactly-zero-PFIT groups. Because $PFIT^L$ and $PFIT^S$ enter Equation (3) with opposite signs, a larger short-leg multiplier implies an asymmetric prediction for outflows: it should widen the long-short spread when PFIT is negative, just as it compresses the spread when PFIT is positive. Consistent with this logic, long-short returns in both the negative- and zero-PFIT groups are well above the high-PFIT group: the negative-minus-high difference is 3.76% (t=6.15) and the zero-minus-high difference is 9.84% (t=5.62) with equal weighting, and 4.04% (t=5.86) and 6.68% (t=3.38) with value weighting. Reading across the five groups, long-short returns are highest where measured passive demand is zero, high where it is negative, and lowest where inflow-induced buying is largest, which is the ordering the price-pressure interpretation implies.

4.3 Fama-MacBeth Regressions

The double sorts show that anomaly spreads are lower among high-PFIT stocks. A natural concern, however, is that PFIT may be correlated with other stock characteristics that also predict returns. As Panel C of Table 2 shows, high-PFIT stocks are larger, more profitable, and more liquid, and they have higher levels of passive ownership. If so, the double-sort results could reflect these correlated characteristics rather than passive flow-induced trading itself.

We address this concern using stock-level Fama-MacBeth regressions, which allow us to control directly for stock characteristics while testing whether the relation between anomaly signals and returns varies with PFIT. For each anomaly a and quarter t, we estimate

$$r^e_{i,t} = \alpha_{a,t} + \beta^P_{a,t} PFIT_{i,t} + \beta^Z_{a,t} z^a_{i,t-1} + \beta^{PZ}_{a,t} \left(PFIT_{i,t} \times z^a_{i,t-1} \right) + \Gamma'_{a,t} X_{i,t} + \varepsilon_{i,t}, \tag{9}$$

where $r^e_{i,t}$ is the quarterly excess return during the holding quarter, $PFIT_{i,t}$ is realized passive flow-induced trading in that quarter expressed as a fraction of lagged shares outstanding, and $z^a_{i,t-1} \in [0, 1]$ is the stock's fractional rank on anomaly signal a at portfolio formation. All variables are winsorized within quarter at the 1st and 99th percentiles, and each regression requires at least 100 stocks. We average the coefficients across anomalies within each quarter and then over the 86 quarters, reporting Newey-West t-statistics with four lags. The coefficient of interest is β^{PZ}. A negative β^{PZ} implies that the return spread associated with the anomaly signal is weaker when passive flow-induced trading is high.

Table 6 reports the results. PFIT is strongly positively related to contemporaneous returns, consistent with passive inflow-induced price pressure: with full controls, a 0.10% increase in PFIT is associated with a return higher by roughly 3.0%. The interaction between PFIT and the anomaly rank is negative and statistically significant in both specifications. In the base specification, which controls only for size, the interaction is -1.526 with a t-statistic of -3.39. Adding passive ownership, active flow-induced trading, short interest, Amihud illiquidity, the bid-ask spread, and market beta leaves the interaction essentially unchanged at -1.432, with a t-statistic of -4.04. Because PFIT is measured as a fraction of shares outstanding and the rank runs from zero to one, this estimate implies that a 0.10% increase in PFIT reduces the return spread between the top and bottom anomaly deciles by about 0.57%. Thus, the relation between passive flow-induced trading and weaker anomaly returns is not explained by size, passive ownership levels, correlated active-fund demand, illiquidity, trading costs, market risk, or the level of short interest.

The controls also help separate the multiplier channel from several alternative explanations. Controlling for passive ownership distinguishes passive flow-induced trading from the

Table 6. Fama-MacBeth Regressions

	Base controls	Full controls
PFIT	7.685	7.463
	(8.97)	(8.07)
Signal fractional rank	0.010	0.007
	(3.57)	(3.95)
PFIT × signal rank	−1.526	−1.432
	(−3.39)	(−4.04)
Log market equity	−0.000	−0.002
	(−0.06)	(−1.27)
Passive ownership		0.027
		(0.69)
AFIT		1.112
		(3.30)
Short interest		−0.169
		(−4.52)
Amihud illiquidity ($\times 10^6$)		−0.004
		(−1.76)
Bid-ask spread		−1.327
		(−2.80)
Beta		0.010
		(1.57)
Constant	0.013	0.052
	(0.23)	(1.81)
Quarters	86	86
Mean stocks per anomaly-quarter regression	3,079	2,817
Mean cross-sectional R^2	0.032	0.070

This table reports stock-level Fama-MacBeth regressions of quarterly excess returns on realized PFIT, the fractional anomaly signal rank, and their interaction, as in Equation (9). Each anomaly-quarter regression is estimated with heteroskedasticity-robust standard errors and requires at least 100 stocks; coefficients are averaged first across the 146 anomalies within each quarter and then over the 86 quarters. Newey-West t-statistics with four lags are reported in parentheses. The sample is from 2002Q2 to 2023Q4.

level of passive ownership, which prior work links to many stock-level outcomes, including governance, lending-market conditions, comovement, and price informativeness [Appel et al., 2016, Von Beschwitz et al., 2025, Qin and Singal, 2015, Buss and Sundaresan, 2020]. Consistent with our argument in Section 3.3 that the flow and the stock are distinct objects, passive ownership itself carries essentially no return effect in these regressions once PFIT is included. Controlling for active flow-induced trading addresses the concern that PFIT is capturing correlated active mutual-fund demand. AFIT does enter positively and significantly, which is consistent with the price-pressure evidence in Lou [2012], but it does not subsume PFIT. Finally, controlling for short interest, illiquidity, and the bid-ask spread addresses the possibility that the result is simply a short-sale constraint or trading-cost effect [Daniel, Klos, and Rottke, 2023] rather than a differential response to passive demand shocks. The negative interaction remains large and significant after adding all these controls, which is consistent with the interpretation that passive demand shocks have larger price effects in stocks where arbitrage is more constrained.

4.4 Longer-Horizon Evidence

One concern with the double-sort results is selection. High-PFIT stocks may differ from low-PFIT stocks in ways that are not fully captured by the controls, and these differences could be related to fundamentals and therefore to expected returns. We address this concern by examining returns over longer horizons, holding the original portfolio sort fixed. If high-PFIT portfolios simply selected stocks with persistently different fundamentals, then the return differences between high- and low-PFIT stocks should continue beyond the holding quarter. Specifically, we fix the anomaly-decile and PFIT-tercile assignments formed in the original holding quarter and follow each portfolio for eight additional quarters without re-sorting, computing average annualized excess returns at each horizon and then averaging across the 146 anomalies.

That is not what we find. Figure 4 plots short-leg and long-leg portfolio returns across

holding horizons for the low- and high-positive-PFIT groups. In both the short and long legs, high-inflow portfolios earn especially high returns during the quarter in which passive flow-induced trading occurs: the gap is 14.75% in the short leg (24.07 versus 9.32) and 10.13% in the long leg (24.64 versus 14.51), consistent with a large part of these return differences reflecting passive-flow price pressure. After that, however, the return paths of high- and low-inflow stocks are similar in both legs: by horizon eight, they are only 0.36% in the short leg and 0.15% in the long leg. This pattern indicates that passive-flow price pressure is concentrated in the trading quarter rather than persisting over longer horizons. Online appendix Figure OA3 reports the value-weighted counterpart and shows the same concentration at horizon zero.

The absence of a sharp reversal is itself useful for interpretation. If the high-PFIT return effect were purely temporary price pressure that fully unwound, then high-PFIT stocks should *underperform* in the following quarters, holding the original sort fixed. Instead, subsequent returns are broadly similar across the two groups. This is consistent with incomplete pass-through of demand shocks into future returns, as discussed in Davis et al. [2026].

The main implication of Figure 4 is that the PFIT sort does not appear to be picking up unobserved differences in fundamentals or expected returns. If high-PFIT stocks were fundamentally different from low-PFIT stocks in a persistent way, we would expect the return gap to continue after the holding quarter. Instead, the gap is concentrated when the passive flow-induced trading occurs, and it is largest precisely in the stocks that the anomaly strategies are short.

5 Mechanism: Heterogeneity in Demand Multipliers and Short Covering

This section studies the mechanism behind the main portfolio results. Section 4 shows that anomaly spreads are compressed among stocks that experience high passive inflows, that this

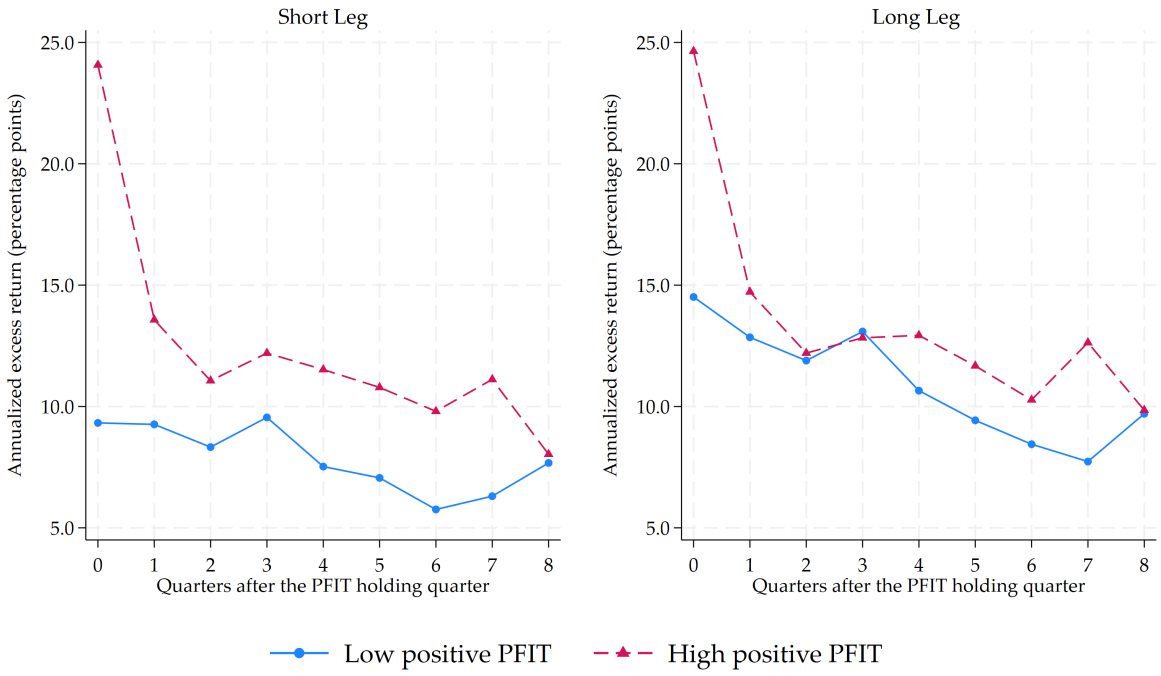


Figure 4. Long- and short-leg returns across holding horizons. This figure plots portfolio returns across holding horizons for the low- and high-positive-PFIT groups, separately for the short leg (left panel) and the long leg (right panel). Stocks are sorted each quarter into positive-PFIT terciles and anomaly deciles as in Table 5. The anomaly-decile and PFIT-tercile assignments are then held fixed and portfolio returns are measured for eight additional quarters without re-sorting. Horizon zero is the quarter in which the realized PFIT shock occurs. Returns are equal-weighted annualized excess returns at each horizon, averaged across the 146 anomalies. The sample is from 2002Q2 to 2023Q4.

compression is not explained by higher passive flow-induced trading in the short leg, and that it does not appear to reflect persistent differences in fundamentals or other observable stock characteristics. In light of the framework in Section 2, this points to the multiplier channel: the same passive demand shock must have a larger price impact in the short leg than in the long leg. We first estimate demand multipliers separately in the long and short legs of each anomaly, and then ask whether the cross-anomaly variation in estimated multipliers lines up with the cross-anomaly variation in realized compression documented in Figure 3. We then study one directly measurable amplification channel that can generate larger short-leg multipliers: short covering.

5.1 Multiplier Estimation

Given the evidence against the pure exposure channel in Section 4, we estimate the multiplier channel directly. For each anomaly a and signal decile d, we pool firm-quarters and estimate

$$r^e_{i,t+1} = \alpha_{a,d} + M_{a,d} PFIT_{i,t+1} + \delta'_{a,d} FF5_{t+1} + \varepsilon_{i,t+1}, \tag{10}$$

where $r^e_{i,t+1}$ is the stock's quarterly excess return during the holding quarter, $PFIT_{i,t+1}$ is realized passive flow-induced trading during the same quarter, including both passive inflows and outflows, and $FF5$ contains the contemporaneous market, size, value, profitability, and investment factors.[2] We exclude observations with exactly zero PFIT. Because both the return and PFIT are quarterly decimals, $M_{a,d}$ is a unit-free price-impact multiplier: a value of one means that a demand shock equal to one percent of shares outstanding moves the quarterly return by one percent. We focus on the comparison between the short leg, D1, and the long leg, D10. This specification is essentially a reduced-form estimate of a linearized demand system in the spirit of van der Beck [2022]. Because we run it for every anomaly, we obtain a distribution of 146 short-leg multipliers, 146 long-leg multipliers, and 146 multiplier

[2]The estimates are similar if we replace the factor controls with quarter fixed effects, which absorb common quarter-level return variation nonparametrically.

gaps.

The key identifying assumption is that the measured demand shock and the price response occur at the same time. This is why the evidence in Section 3.3 matters. If current flow-induced trading were predictable in advance, or if it revealed a longer stream of future mechanical demand, then part of the contemporaneous price response would reflect earlier or current revisions about future demand rather than the effect of the current trade itself, and we would *overstate* the multiplier. Similarly, if passive flows were informed, the price response would reflect information rather than demand. Neither appears to be the case: realized PFIT is close to an innovation in demand, and passive flows do not forecast future fund returns.

Table 7 reports the estimates. Across the 146 anomalies, the average short-leg multiplier is 2.77 and the average long-leg multiplier is 2.57. Both are estimated precisely and both are far from zero, which is itself informative: passive demand shocks move prices in both legs, so the demand curve for individual stocks slopes down throughout the anomaly cross section. The average gap between the two is 0.205, with a cross-anomaly t-statistic of 2.02. The median gap, 0.219, is very close to the mean, so the positive average is not driven by a few outliers. This is modest but direct evidence for the multiplier channel: the same demand shock has a larger price effect in the stocks that anomaly strategies short.

Table 7. Multiplier Estimation for the Long and Short Legs

Measure	Mean	Median	Std. dev.	t-stat.	N anomalies
Short-leg multiplier, M_S	2.774	2.618	0.907	36.94	146
Long-leg multiplier, M_L	2.569	2.449	0.871	35.65	146
Gap, $M_S - M_L$	0.205	0.219	1.228	2.02	146

This table reports the cross-anomaly distribution of estimated demand multipliers. For each anomaly decile, the multiplier is the slope on realized PFIT in the pooled stock-quarter regression in Equation (10), which controls for the contemporaneous Fama-French five factors. The short-leg multiplier M_S is estimated in D1, the long-leg multiplier M_L is estimated in D10, and the gap is their difference within anomaly. Observations with exactly zero PFIT are excluded. The reported t-statistics come from robust constant-only regressions across the 146 anomaly-level estimates. The sample is from 2002Q2 to 2023Q4.

Table 7 also shows why the average gap is only modestly significant: the cross-anomaly standard deviation of the gap, 1.228, is roughly six times its mean. Multiplier heterogeneity across anomalies is large relative to the average asymmetry between legs. This dispersion parallels the dispersion in realized compression in Figure 3, which suggests a sharper test than the average gap. If the multiplier channel is the right description of the mechanism, then the anomalies with the largest estimated multiplier gaps and the largest leg differences in realized demand should be exactly the anomalies whose long-short spreads compress the most.

We implement this test by constructing, for each anomaly, the return compression that its own estimated multipliers and its own realized demand shocks imply. Let $D_{S,a,p}$ and $D_{L,a,p}$ denote the average realized PFIT in the short and long legs of positive-PFIT tercile p, taken from the same double-sorted portfolios used in Table 5, and let $M_{S,a}$ and $M_{L,a}$ be the multipliers from Table 7. Following Equation (3), the PFIT contribution to the long-short return in tercile p is

$$I_{a,p} = M_{L,a} D_{L,a,p} - M_{S,a} D_{S,a,p}, \tag{11}$$

and the implied low-minus-high compression is $I_a^{P1-P3} = 400\,(I_{a,1} - I_{a,3})$ in annualized percentage points. We then regress each anomaly's realized low-minus-high spread difference, the object plotted in Figure 3, on its implied counterpart. This cross-sectional test asks the question: do the anomalies that *should* compress more actually compress more?

They do. Figure 5 plots the relation. The slope is 1.313 with a heteroskedasticity-robust t-statistic of 5.69, and the R^2 is 19.3%. A slope near one is what the accounting identity in Equation (3) predicts if the estimated multipliers and demand shocks are measured without systematic bias, although the standard error is wide enough that we would not want to lean on the point estimate. Online appendix Figure OA4 reports the value-weighted version, where the slope is 0.830 ($t = 2.86$) with an R^2 of 8.2%. The relation is therefore positive and significant under both weighting schemes.

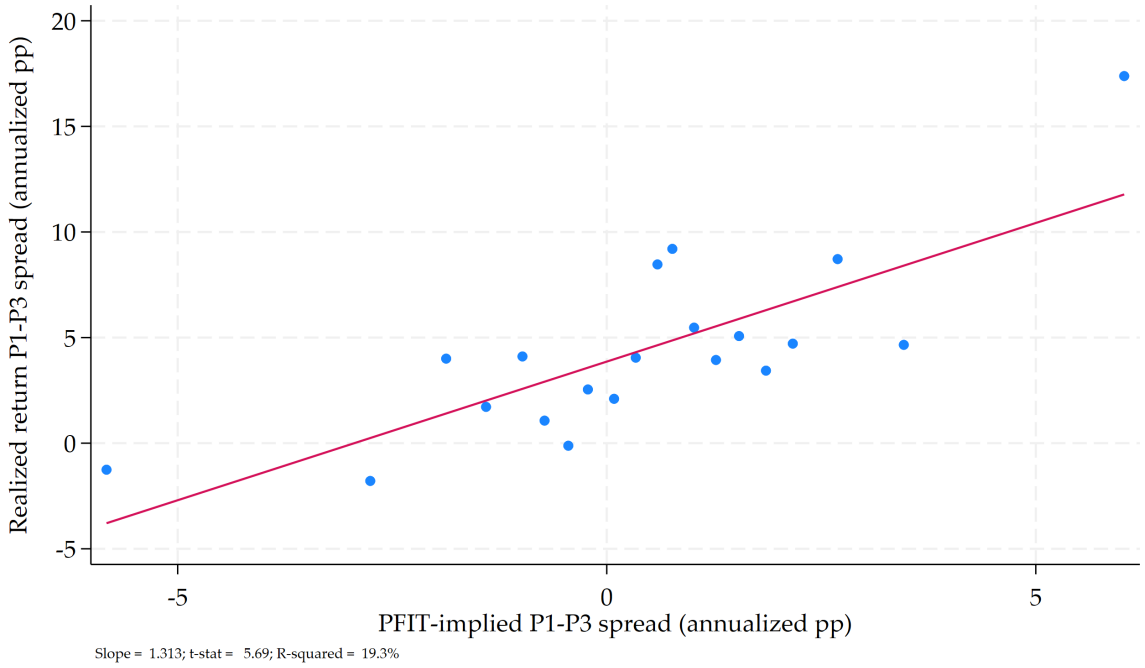


Figure 5. Implied versus realized low minus high PFIT spread return. This figure plots each anomaly's realized low-minus-high PFIT long-short return difference against the difference implied by its own estimated demand multipliers and realized demand shocks. The implied spread is constructed as in Equation (11), combining the leg-specific multipliers from Table 7 with the average equal-weighted PFIT in the short and long legs of each positive-PFIT tercile from the double sorts in Table 5, and is expressed in annualized percentage points. Plotted points are equal-count bins of anomalies formed after ranking on the implied spread; the fitted line, slope, t-statistic, and R^2 are estimated on all 146 anomaly-level observations with heteroskedasticity-robust standard errors. The sample is from 2002Q2 to 2023Q4.

Which of the two channels in Equation (3) accounts for this explanatory power? Within a PFIT tercile, the implied spread can be written symmetrically as

$$M^L D^L - M^S D^S = \left(M^L - M^S\right) \frac{D^L + D^S}{2} + \left(D^L - D^S\right) \frac{M^L + M^S}{2}, \tag{12}$$

where the first term is the multiplier-difference contribution and the second is the demand-difference contribution. Applying this identity in P1 and P3 and differencing, a symmetric allocation attributes 14.1% of the jointly explained cross-anomaly variation to multiplier differences and 5.4% to demand differences, leaving 80.5% of the cross-anomaly variation in realized compression unexplained. Two conclusions follow. First, essentially all of the explained variation comes from the multiplier channel rather than the exposure channel, which reinforces the portfolio evidence in Section 4. Second, most of the cross-anomaly variation in compression is not explained by our PFIT accounting at all. That is not surprising. Anomaly returns are noisy, our multipliers are estimated with error, and passive flow-induced trading is only one of many forces acting on realized anomaly spreads during a given quarter. We read Figure 5 as evidence that the multiplier channel is present and economically relevant in the cross section, not as a claim that passive flows are the dominant driver of anomaly returns.

5.2 Short Covering and Amplification

One reason short-leg multipliers may be larger is short covering. Passive inflows can push up the prices of heavily shorted stocks for reasons unrelated to stock-level fundamentals, acting like noise-trader demand shocks in the spirit of De Long et al. [1990]. Because this price move goes against short sellers, some may reduce positions by buying to cover. That buying pressure comes on top of the original passive-flow-induced trade, generating a larger total price response and, mechanically, a larger estimated multiplier.

We measure short covering using Markit utilization rather than the Compustat short-

interest ratio, following Muravyev et al. [2025]. Short interest reflects only the demand side of the lending market, the quantity of shares borrowed, so a decline could equally reflect fewer available shares to borrow. Utilization is the ratio of shares on loan to lendable supply and therefore nets out changes in supply, isolating demand-driven covering.

We test this channel using changes in securities-lending conditions after passive-flow shocks. For each anomaly, we keep the stocks in the short and long legs with positive PFIT and estimate

$$\Delta Y_{i,t+1} = \alpha_t + \beta_1 PFIT_{i,t+1} + \beta_2 ShortLeg_{i,t} + \beta_3 \left(PFIT_{i,t+1} \times ShortLeg_{i,t}\right) + \beta_4 Y_{i,t} + \epsilon_{i,t+1}, \tag{13}$$

where $Y_{i,t}$ denotes a short-sale outcome, either Markit utilization or the Markit indicative lending fee, and $\Delta Y_{i,t+1}$ is its change during the holding quarter. $ShortLeg_{i,t}$ is an indicator equal to one for stocks in D1 and zero for stocks in D10. The quarter fixed effects absorb common time-series variation in securities-lending conditions, the regression controls for the lagged level of the corresponding outcome, and standard errors are clustered by firm. Because Markit coverage improves substantially beginning in 2010, these tests use holding quarters from 2010Q2 through 2023Q4. The coefficient of interest is β_3, the difference between the short-leg and long-leg responses; the total short-leg slope is $\beta_1 + \beta_3$. A negative total slope indicates that passive-flow shocks are followed by declines in utilization or borrowing costs among heavily shorted stocks, consistent with short covering.

Short covering requires two ingredients, and our anomaly universe lets us test for both. It requires a strong price-impact channel, so that passive demand actually moves prices against short sellers, and it requires meaningful pre-existing short positions to cover. We therefore split the 146 anomalies two ways. The first split uses the sign of the estimated multiplier gap $M_S - M_L$ from Table 7: 86 anomalies have a positive gap and 60 have a negative gap. The second split uses the sign of the shorting-intensity gap, measured as the lagged difference

in Markit utilization between the short and long legs: 97 anomalies have a short leg that is more heavily shorted than their long leg, and 49 do not. Within each group we average the anomaly-specific coefficients and test the group mean across anomalies.

Table 8 reports the results, and the pattern is exactly what the short-covering story predicts. Passive inflows are followed by declines in utilization and indicative fees in every group, which is consistent with passive buying leads to short covering generally. What differs across groups is the *additional* short-leg response. Among the 86 anomalies with a positive multiplier gap, the interaction term is negative and significant for both outcomes (-9.30 with $t = -2.47$ for utilization and -0.125 with $t = -3.03$ for fees), and the total short-leg response is substantially larger in magnitude than the long-leg response. Among the 60 anomalies with a negative multiplier gap, the interaction is positive or insignificant, so the short leg responds no more than the long leg. The split on shorting intensity is even sharper: among the 97 anomalies whose short legs are more heavily shorted, the interaction is -14.71 ($t = -4.65$) for utilization and -0.164 ($t = -5.10$) for fees, while among the remaining 49 anomalies the interaction flips sign. In other words, the short-leg amplification appears precisely where price impact is largest and where there are short positions to cover, and it disappears where either ingredient is missing.

Table 8. Passive Flows and Short Covering

		Utilization change			Indicative-fee change		
Grouping rule	N	PFIT	PFIT × short	Total short	PFIT	PFIT × short	Total short
Positive multiplier gap	86	−19.334	−9.301	−28.635	−0.366	−0.125	−0.491
		(−6.09)	(−2.47)	(−8.54)	(−13.94)	(−3.03)	(−11.07)
Negative multiplier gap	60	−23.141	2.689	−20.452	−0.396	0.076	−0.320
		(−4.76)	(0.50)	(−8.29)	(−13.18)	(2.21)	(−14.05)
Positive shorting-intensity gap	97	−15.547	−14.708	−30.254	−0.311	−0.164	−0.474
		(−5.29)	(−4.65)	(−11.67)	(−14.96)	(−5.10)	(−12.11)
Negative shorting-intensity gap	49	−31.493	16.083	−15.409	−0.511	0.196	−0.315
		(−5.81)	(2.62)	(−3.92)	(−14.50)	(4.64)	(−10.45)

This table reports regressions of changes in shorting conditions on passive flow-induced trading, an indicator for the anomaly short leg, and their interaction, as in Equation (13). The outcomes are the change in Markit utilization and the change in the Markit indicative lending fee during the holding quarter. For each anomaly, the sample is restricted to the short and long legs, D1 and D10, and to observations with positive passive flow-induced trading; the regression absorbs quarter fixed effects, controls for the lagged level of the outcome, and clusters standard errors by firm. "PFIT" is β_1, the long-leg slope; "PFIT × short" is β_3, the difference between the short- and long-leg slopes; and "Total short" is $\beta_1 + \beta_3$. Anomalies are grouped by the sign of the estimated multiplier gap $M_S - M_L$ from Table 7 and by the sign of the lagged short-minus-long-leg difference in Markit utilization. The displayed coefficients are means of the anomaly-specific estimates within each group, and the t-statistics in parentheses test the cross-anomaly group mean. The sample is from 2010Q2 to 2023Q4.

The two splits are not independent, and their relation is suggestive about where multiplier heterogeneity comes from in the first place. We find that the cross-anomaly correlation between the estimated multiplier gap, $M_S - M_L$, and the shorting-intensity gap, measured as the lagged short-minus-long-leg difference in utilization, is above 0.3 and statistically significant. Anomalies whose short legs are more heavily shorted going into the holding quarter are the anomalies whose short legs have the steeper demand curves. This is what one would expect if pre-existing short positions are part of what makes short-leg demand curves steep: a heavily shorted stock has a larger stock of positions that can be forced to unwind when prices rise, and short sellers who cover are buying at the same time as the index funds. This suggests that heavier shorting may help explain why demand multipliers are larger in anomaly short legs.

Short covering is also not the only reason short-leg stocks may have larger passive-flow multipliers. More generally, anomaly short legs often contain stocks where arbitrage capital is scarce, disagreement is high, and current holders have especially strong beliefs. As one example of this, consider "meme" stocks during the COVID-19 pandemic: they featured optimistic valuations, highly engaged investors, and owners who were reluctant to sell [Greenwood, Laarits, and Wurgler, 2023]. In such stocks, arbitrageurs may be less willing or able to lean against price pressure, while current holders may require a larger price increase to supply shares. This interpretation is consistent with models in which disagreement and short-sale constraints interact to generate return continuation, producing dynamics that resemble price momentum [Daniel et al., 2023].

Thus, the short-covering tests should be read as evidence on one measurable amplification channel, not as a claim that all multiplier heterogeneity comes from literal short covering. Other limits-to-arbitrage mechanisms can generate similar price-impact asymmetry, and these mechanisms are likely to be strongest in anomaly short legs. The broader point is that passive demand shocks should have larger effects precisely in the stocks where arbitrage is most limited. This is the sense in which passive flow-induced trading creates a

limit to arbitrage for investors implementing convergence trades.

6 Conclusion

Sharpe's arithmetic is a useful tool for thinking about the effect of rising passive investing on the active management industry [Sharpe, 1991]. If passive investors hold the market portfolio and earn zero alpha relative to the market, active investors must also earn zero alpha in aggregate. But this aggregate accounting identity is not the experiment faced by each individual active manager. Active managers hold specific positions, in specific parts of the market, over specific holding periods. The realized payoff to those strategies depends not only on whether the underlying signal is informative about returns *on average*, but also on the demand shocks that arrive while the trade is in place.

This paper studies this issue across a broad universe of long-short anomaly strategies. We use passive flow-induced trading to measure the mechanical stock-level demand shocks created when index funds scale their holdings in response to investor flows. These shocks are well suited to the question: passive flows are hard to forecast out of sample and do not predict subsequent fund returns, so realized PFIT is close to an unanticipated and uninformed demand shock. Across 146 anomalies, long-short returns are substantially lower among stocks experiencing large passive inflows during the holding period, falling from 5.01% in the low-inflow tercile to 0.60% in the high-inflow tercile. The pattern holds for about 70% of individual anomalies and survives value weighting, risk adjustment, and stock-level controls.

The mechanism is not that short-leg stocks receive more passive demand than long-leg stocks. Average passive flow-induced trading and passive ownership are nearly identical in the two legs. Instead, the same passive demand shock has a larger price impact in the short leg, where arbitrage is more constrained. Estimated demand multipliers are larger in the short leg on average, and, more tellingly, the anomalies whose multipliers and demand

shocks imply more compression are the anomalies that display more compression in the data. Consistent with one amplification channel, securities-lending utilization and borrow fees fall by more in the short leg after passive-flow shocks, but only for the anomalies with large multiplier gaps and heavily shorted short legs, and those two features are themselves positively correlated in the cross section of anomalies.

The main takeaway is that passive flow-induced trading can act as a limit to arbitrage. Even if anomaly signals continue to contain information about future returns, the realized payoff to trading on those signals depends on mechanical demand pressure during the holding period. In this sense, passive investing need not change aggregate active alpha mechanically to affect active management. It can instead change the risks, timing, and realized returns of the trades active managers use to exploit mispricing.

References

Ian R Appel, Todd A Gormley, and Donald B Keim. Passive investors, not passive owners. *Journal of Financial Economics*, 121(1):111–141, 2016.

Pouya Behmaram and Carter Davis. Indexing and the elasticity of stock demand. Working paper, Indiana University, 2026.

Itzhak Ben-David, Francesco Franzoni, and Rabih Moussawi. Do ETFs Increase Volatility? *Journal of Finance*, 73(6), 2018. ISSN 15406261. doi: 10.1111/jofi.12727.

Adrian Buss and Savitar Sundaresan. More Risk, More Information: How Passive Ownership Can Improve Informational Efficiency. *SSRN Electronic Journal*, 2020. ISSN 1556-5068. doi: 10.2139/ssrn.3615173.

Yen-Cheng Chang, Harrison Hong, and Inessa Liskovich. Regression discontinuity and the price effects of stock market indexing. *The Review of Financial Studies*, 28(1):212–246, 2015.

Andrew Y. Chen and Tom Zimmermann. Open source cross-sectional asset pricing. *Critical Finance Review*, 27(2):207–264, 2022.

Judith Chevalier and Glenn Ellison. Risk taking by mutual funds as a response to incentives. *Journal of political economy*, 105(6):1167–1200, 1997.

Alex Chinco and Marco Sammon. The passive ownership share is double what you think it is. *Journal of Financial Economics*, 157:103860, 2024.

Jeffrey L Coles, Davidson Heath, and Matthew C Ringgenberg. On index investing. *Journal of Financial Economics*, 145(3):665–683, 2022.

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

Zhi Da and Sophie Shive. Exchange traded funds and asset return correlations. *European Financial Management*, 24(1), 2018. ISSN 1468036X. doi: 10.1111/eufm.12137.

Kent Daniel, Alexander Klos, and Simon Rottke. The dynamics of disagreement. *The Review of Financial Studies*, 36(6):2431–2467, 2023.

Carter Davis, Xiao Han, Stanislav Sokolinski, and Andrea Tamoni. Erasing alpha. Working paper, Indiana University, University of Bristol, and Rutgers University, 2026.

J Bradford De Long, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738, 1990.

Itamar Drechsler and Qingyi Freda Drechsler. The shorting premium and asset pricing anomalies. NBER Working Paper 20282, National Bureau of Economic Research, 2014.

Eugene F Fama and Kenneth R French. Average returns, b/m, and share issues. *The Journal of Finance*, 63(6):2971–2995, 2008a.

Eugene F Fama and Kenneth R French. Dissecting anomalies. *The journal of finance*, 63 (4):1653–1678, 2008b.

Ilias Filippou, Songrun He, Sophia Zhengzi Li, and Guofu Zhou. Etfs, anomalies and market efficiency. *Available at SSRN 4056260*, 2023.

Andrea Frazzini and Owen A Lamont. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of financial economics*, 88(2):299–322, 2008.

Xavier Gabaix and Ralph S. J. Koijen. In search of the origins of financial fluctuations: The inelastic markets hypothesis. NBER Working Paper 28967, National Bureau of Economic Research, 2021.

Xavier Gabaix and Ralph SJ Koijen. Granular instrumental variables. *Journal of Political Economy*, 132(7):2274–2303, 2024.

Lawrence Glosten, Suresh Nallareddy, and Yuan Zou. ETF activity and informational efficiency of underlying securities. *Management Science*, 67(1), 2021. ISSN 15265501. doi: 10.1287/mnsc.2019.3427.

Robin Greenwood. Short-and long-term demand curves for stocks: theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics*, 75(3):607–649, 2005.

Robin Greenwood, Toomas Laarits, and Jeffrey Wurgler. Stock market stimulus. *The Review of Financial Studies*, 36(10):4082–4112, 2023.

Valentin Haddad, Paul Huebner, and Erik Loualiche. How competitive is the stock mar-

ket? theory, evidence from portfolios, and implications for the rise of passive investing. *American Economic Review*, 2025.

Samuel G Hanson and Adi Sunderam. The growth and limits of arbitrage: Evidence from short interest. *The Review of Financial Studies*, 27(4):1238–1286, 2014.

Lawrence Harris and Eitan Gurel. 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(4):815–829, 1986.

Doron Israeli, Charles M.C. Lee, and Suhas A. Sridharan. Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies*, 22(3), 2017. ISSN 13806653. doi: 10.1007/s11142-017-9400-8.

Aditya Kaul, Vikas Mehrotra, and Randall Morck. Demand curves for stocks do slope down: New evidence from an index weights adjustment. *The Journal of Finance*, 55(2):893–912, 2000.

Ralph SJ Koijen and Motohiro Yogo. A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515, 2019.

Jonathan Lewellen. Institutional investors and the limits of arbitrage. *Journal of Financial Economics*, 102(1):62–80, 2011.

Dong Lou. A flow-based explanation for return predictability. *The Review of Financial Studies*, 25(12):3457–3489, 2012.

Dmitriy Muravyev, Neil D Pearson, and Joshua M Pollet. Anomalies and their short-sale costs. *The Journal of Finance*, 80(6):3639–3694, 2025.

Stefan Nagel. Short sales, institutional investors and the cross-section of stock returns. *Journal of financial economics*, 78(2):277–309, 2005.

Darius Palia and Stanislav Sokolinski. Passive Asset Management, Securities Lending and Stock Prices. *SSRN Electronic Journal*, 2019. ISSN 1556-5068. doi: 10.2139/ssrn.3335283.

Anna Pavlova and Taisiya Sikorskaya. Benchmarking intensity. *The Review of Financial Studies*, 36(3):859–903, 2023.

Stefano Pegoraro, Marco Sammon, and John J. Shim. Optimal index-linked rebalancing with anticipatory trading. Working paper, University of Notre Dame and Harvard Business School, May 2026.

Jeffrey Pontiff and Artemiza Woodgate. Share issuance and cross-sectional returns. *The Journal of Finance*, 63(2):921–945, 2008.

Nan Qin and Vijay Singal. Indexing and stock price efficiency. *Financial management*, 44 (4):875–904, 2015.

Marco Sammon. Passive ownership and price informativeness. *Management Science*, 71(6): 4582–4598, 2025.

Marco Sammon and John J Shim. Do active funds do better in what they trade? *Available at SSRN*, 2023.

Marco Sammon and John J. Shim. Who clears the market when passive investors trade? *The Review of Financial Studies*, 2026a.

Marco Sammon and John J Shim. Index rebalancing and stock market composition: Do indexes time the market? *Journal of Financial Economics*, 177:104229, 2026b.

Myron S Scholes. The market for securities: Substitution versus price pressure and the effects of information on share prices. *The Journal of Business*, 45(2):179–211, 1972.

William F Sharpe. The arithmetic of active management. *Financial Analysts Journal*, 47 (1):7–9, 1991.

Andrei Shleifer. Do demand curves for stocks slope down? *The Journal of Finance*, 41(3): 579–590, 1986.

Andrew Ross Sorkin. Seth klarman warns about a risky market under trump. The New York Times (DealBook), February 2017. URL `https://www.nytimes.com/2017/02/06/business/dealbook/sorkin-seth-klarman-trump-investors.html`. Accessed 2026-02-19.

Robert F Stambaugh, Jianfeng Yu, and Yu Yuan. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2):288–302, 2012.

Robert F Stambaugh, Jianfeng Yu, and Yu Yuan. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance*, 70(5):1903–1948, 2015.

Philippe Van der Beck. Flow-driven esg returns. Swiss finance institute research paper, Swiss Finance Institute, 2021.

Philippe van der Beck. Identifying elasticities in demand-based asset pricing. *Available at SSRN*, 2022.

Bastian Von Beschwitz, Pekka Honkanen, and Daniel Schmidt. Passive ownership and short selling. *Review of Finance*, page rfaf023, 2025.

Online Appendix

Passive Flows and the Limits to Arbitrage

This appendix reports the supporting analyses referenced in the main text. Section OA.1 tests whether fund flows forecast subsequent fund returns, which is the informativeness test discussed in Section 3.3 of the paper. Section OA.2 reports the out-of-sample fund-flow calibration when fund-quarters are weighted by total net assets. Section OA.3 reports the value-weighted counterpart of the main double sorts, and Section OA.4 adds the negative- and zero-PFIT groups. Sections OA.5 and OA.6 report the value-weighted counterparts of the longer-horizon and implied-spread results. The anomaly universe is the same 146 signals used in the main text, and, unless stated otherwise, portfolio returns and alphas are annualized percentage points and cross-anomaly inference is based on a robust constant-only regression across the anomaly-level estimates. All tables and figures are collected at the end of this appendix.

OA.1 Fund Flows and Subsequent Fund Returns

Section 3.3 argues that PFIT is a useful proxy for an inelastic demand shock in part because passive flows do not appear to be informed. Table OA1 reports the underlying test. The dependent variable is fund f's return in quarter $t+1$, and, separately for passive and active funds, we estimate

$$R^{fund}_{f,t+1} = \alpha + \sum_{j=0}^{4} \rho_j Flow_{f,t-j} + \sum_{j=0}^{4} \gamma_j R^{fund}_{f,t-j} + \varepsilon_{f,t+1}, \tag{OA.1}$$

If investors direct money to funds because they have information about the securities those funds hold, current and recent flows should forecast subsequent fund performance. Equal-weighted specifications give each fund-quarter equal weight, while value-weighted specifications use analytic weights proportional to end-of-quarter total net assets within the fund-type

by quarter cell. The dependent variable and all regressors are winsorized at the 1st and 99th percentiles within quarter and passive/active status, and standard errors are two-way clustered by fund and quarter.

The results give no support for a timing interpretation. Current flow does not significantly predict next-quarter returns in any of the four specifications, for either fund type and under either weighting scheme. A handful of individual lag coefficients are statistically significant, but their signs are not stable across specifications: for example, the coefficient on the second flow lag is positive and significant in the equal-weighted specifications and insignificant in the value-weighted ones, while the coefficient on the fourth lag switches sign across weighting schemes. This is what one would expect from sampling variation across twenty estimated coefficients rather than from a systematic timing ability. The R^2 values reinforce the point: they range from 0.014 to 0.047, so current and recent flows account for almost none of the variation in subsequent fund returns.

OA.2 TNA-Weighted Fund-Flow Calibration

Figure OA5 reproduces the out-of-sample fund-flow calibration in Figure 1 of the main text, weighting fund-quarters by total net assets rather than equally. This addresses the concern that the low predictability of passive flows is a small-fund phenomenon driven by the many tiny index products in the CRSP mutual fund database. It is not. Predicted passive flows explain 1.7% of the variation in subsequently realized passive flows, with a calibration slope of 0.312, compared with 19.0% and a slope of 0.691 for active funds. If anything, TNA weighting widens the gap between passive and active flow predictability relative to the equal-weighted results in Table 3.

OA.3 Value-Weighted Double Sorts

Our baseline portfolios are equal-weighted, following the convention in the anomaly literature and in Chen and Zimmermann [2022]. Table OA2 reports the value-weighted counterpart of main-text Table 5. Portfolios are formed identically: within each quarter, stocks with positive realized PFIT are sorted into terciles and then into anomaly deciles within tercile. Returns use formation-quarter market capitalization as weights, and a portfolio-quarter is retained only when its weights sum to more than 0.99.

The main pattern survives value weighting. The average long-short return falls from 3.593% in the low-PFIT tercile, to 1.099% in the middle tercile, to −0.708% in the high-PFIT tercile. The low-minus-high difference is 4.301% with a cross-anomaly t-statistic of 6.34, which is close to the equal-weighted difference of 4.412%. The CAPM-alpha difference is 3.720% ($t = 6.03$). As in the equal-weighted results, realized passive demand raises returns in both legs, but by more in the short leg: moving from the low- to the high-PFIT tercile raises value-weighted short-leg returns by 20.178% and long-leg returns by 15.877%. Both leg-level effects are larger than their equal-weighted counterparts of 14.767% and 10.355%, which is consistent with passive flow-induced trading mattering more for the large stocks that dominate value-weighted portfolios, and the difference between the two legs is exactly the 4.301% compression of the spread.

Figure OA2 plots the value-weighted leg returns, and Figure OA1 plots the distribution of the 146 anomaly-level value-weighted low-minus-high differences. The value-weighted distribution closely resembles its equal-weighted counterpart in Figure 3: the cross-anomaly mean is 4.30% with a t-statistic of 6.34, and 69.2% of anomalies have a positive difference, compared with 69.9% with equal weighting. The dispersion is somewhat wider in the tails, ranging from about −20% to +32%.

OA.4 Negative- and Zero-PFIT Observations

The main analysis restricts attention to stock-quarters with positive PFIT, in order to isolate passive inflow-induced buying pressure and because inflows account for about 70% of nonzero observations. Table OA3 adds the two omitted groups, stock-quarters with negative PFIT and those with exactly zero measured PFIT, and reports both equal- and value-weighted results. Zero PFIT means zero measured net flow-induced trading in that stock-quarter; it does not mean zero passive ownership, because index-fund flows may be zero or may offset one another across the funds holding the stock.

For the negative PFIT group, because $PFIT^L$ and $PFIT^S$ enter Equation (3) with opposite signs, a larger short-leg multiplier implies an asymmetric prediction for outflows: it should widen the long-short spread when PFIT is negative, just as it compresses the spread when PFIT is positive. The findings are largely consistent with it.

Adding these groups extends rather than qualifies the main pattern. With equal weighting, the long-short return is 4.359% in the negative-PFIT group and 10.444% in the zero-PFIT group, compared with 0.600% in the high-positive-PFIT group. The negative-minus-high difference is 3.759% ($t = 6.15$) and the zero-minus-high difference is 9.843% ($t = 5.62$). The value-weighted differences are 4.037% ($t = 5.86$) and 6.684% ($t = 3.38$). CAPM-alpha differences have the same sign and significance under both weighting schemes. Reading across the five groups, long-short returns are highest where measured passive demand is zero, high where it is negative, and lowest where inflow-induced buying is largest, which is the ordering the price-pressure interpretation implies. The unusually large spread in the zero-PFIT group should be interpreted with care: it is a zero-net-demand group rather than a zero-passive-ownership group, and it contains a different, generally smaller set of stocks.

OA.5 Value-Weighted Longer-Horizon Returns

Figure OA3 reports the value-weighted counterpart of main-text Figure 4. As in the equal-weighted case, the anomaly-decile and PFIT-tercile assignments are formed in the original holding quarter and then held fixed for eight additional quarters without re-sorting.

The message is the same and, if anything, sharper. At horizon zero, high-PFIT portfolios earn much higher returns than low-PFIT portfolios in both legs, with gaps of roughly 20% and 16% in the short and long legs respectively. From horizon one onward the gaps collapse and the two paths become quite similar afterwards. The robust conclusion, common to both weighting schemes, is that the large return gap between high- and low-PFIT stocks is concentrated in the quarter in which the passive demand arrives and does not persist afterward.

OA.6 Value-Weighted Implied Versus Realized Spreads

Figure OA4 reports the value-weighted counterpart of main-text Figure 5. For each anomaly, we combine the leg-specific demand multipliers with the average realized demand shocks in the short and long legs of each positive-PFIT tercile to form the PFIT-implied low-minus-high spread in Equation (11), and regress the realized value-weighted spread difference on it.

The relation remains positive and significant. The slope is 0.830 with a heteroskedasticity-robust t-statistic of 2.86, and the R^2 is 8.2%, compared with a slope of 1.313 and an R^2 of 19.3% in the equal-weighted specification.

Online Appendix Tables and Figures

Table OA1. Fund Flows and Next-Quarter Fund Returns

	Passive funds		Active funds	
Regressor	Equal weighted	Value weighted	Equal weighted	Value weighted
Current flow	−0.005	−0.082	−0.005	0.041
	(−0.37)	(−1.74)	(−0.30)	(1.10)
Flow, $t-1$	−0.002	0.014	−0.004	−0.034
	(−0.19)	(0.26)	(−0.30)	(−1.35)
Flow, $t-2$	0.028	0.023	0.023	0.045
	(2.15)	(0.66)	(2.09)	(1.25)
Flow, $t-3$	−0.002	0.025	−0.022	−0.047
	(−0.26)	(0.94)	(−1.71)	(−1.84)
Flow, $t-4$	−0.027	−0.003	−0.013	−0.011
	(−2.22)	(−0.11)	(−1.22)	(−0.48)
Current fund return	−0.055	−0.011	0.004	0.017
	(−0.41)	(−0.12)	(0.03)	(0.21)
Fund return, $t-1$	0.044	−0.004	0.029	0.011
	(0.43)	(−0.05)	(0.32)	(0.16)
Fund return, $t-2$	−0.113	−0.065	−0.100	−0.083
	(−1.10)	(−0.74)	(−1.11)	(−1.06)
Fund return, $t-3$	−0.136	−0.085	−0.106	−0.065
	(−1.43)	(−0.99)	(−1.40)	(−0.86)
Fund return, $t-4$	0.105	0.063	0.078	0.023
	(0.84)	(0.81)	(0.80)	(0.28)
Observations	31,568	31,568	85,195	85,195
R^2	0.047	0.021	0.029	0.014

This table reports fund-quarter regressions of next-quarter fund returns on current and four lags of fund flows and fund returns, as in Equation (OA.1). The four regressions are estimated separately by passive/active status and weighting scheme. Value-weighted specifications use analytic weights proportional to end-of-quarter total net assets within the fund-type by quarter cell. Parentheses contain two-way clustered t-statistics by fund and quarter. Constants are estimated but not reported. The sample is from 2002Q2 to 2023Q4.

Table OA2. PFIT and Anomaly Double-Sorted Portfolio Returns (Value Weighted)

Outcome	Portfolio	Low	Middle	High	Low−High
Return	Short leg	4.185	13.536	24.438	−20.253
		(11.85)	(38.63)	(80.30)	(−42.46)
Return	Long leg	7.833	14.602	23.710	−15.877
		(20.18)	(50.60)	(86.99)	(−32.30)
Return	Long-short	3.593	1.099	−0.708	4.301
		(7.26)	(2.45)	(−1.63)	(6.34)
CAPM alpha	Short leg	−8.340	0.514	11.038	−19.378
		(−19.57)	(1.32)	(38.45)	(−44.14)
CAPM alpha	Long leg	−4.287	2.142	11.355	−15.643
		(−12.18)	(7.02)	(42.39)	(−34.75)
CAPM alpha	Long-short	4.006	1.634	0.286	3.720
		(6.90)	(2.98)	(0.68)	(6.03)

This table is the value-weighted counterpart of Table 5. Each quarter, stocks with positive realized PFIT during the holding quarter are sorted into terciles based on PFIT and then, within each tercile, into deciles based on the anomaly signal available at the end of the previous quarter. Portfolio returns use formation-quarter market capitalization as weights. The short leg is D1, the long leg is D10, and the long-short portfolio is D10 minus D1. Anomaly-level means and CAPM alphas are estimated with heteroskedasticity-robust standard errors. The table reports robust cross-anomaly means with cross-anomaly t-statistics in parentheses. Returns and alphas are annualized percentage points. The sample is from 2002Q2 to 2023Q4.

Table OA3. Negative and Zero PFIT Sample

Outcome	Portfolio	Negative	Zero	Low	Middle	High	Neg.−High	Zero−High	Low−High
Panel A: Equal-weighted portfolios									
Return	Short leg	0.785	−9.576	9.655	13.970	24.422	−23.638	−33.999	−14.767
		(2.17)	(−8.87)	(25.08)	(49.19)	(68.25)	(−53.97)	(−28.67)	(−33.80)
Return	Long leg	5.144	0.867	14.668	16.638	25.023	−19.879	−24.155	−10.355
		(14.03)	(0.72)	(37.26)	(60.38)	(82.86)	(−49.74)	(−18.97)	(−27.41)
Return	Long-short	4.359	10.444	5.013	2.668	0.600	3.759	9.843	4.412
		(7.77)	(6.56)	(8.60)	(6.67)	(1.20)	(6.15)	(5.62)	(7.15)
CAPM alpha	Short leg	−12.262	−23.633	−5.345	−0.316	9.141	−21.403	−32.774	−14.487
		(−33.66)	(−20.41)	(−14.25)	(−1.24)	(36.42)	(−56.13)	(−28.90)	(−38.51)
CAPM alpha	Long leg	−7.449	−12.900	0.284	3.040	10.211	−17.660	−23.111	−9.927
		(−23.49)	(−10.59)	(0.79)	(12.60)	(55.98)	(−50.04)	(−18.91)	(−26.32)
CAPM alpha	Long-short	4.813	10.733	5.629	3.356	1.070	3.743	9.663	4.559
		(8.71)	(6.55)	(9.65)	(9.32)	(3.30)	(7.08)	(6.04)	(8.36)
Panel B: Value-weighted portfolios									
Return	Short leg	−4.273	−4.978	4.240	13.503	24.418	−28.692	−29.397	−20.178
		(−11.95)	(−3.67)	(11.75)	(38.77)	(79.77)	(−62.88)	(−20.49)	(−41.34)
Return	Long leg	−0.945	0.997	7.833	14.602	23.710	−24.655	−22.713	−15.877
		(−2.37)	(0.89)	(20.18)	(50.60)	(86.99)	(−50.97)	(−19.51)	(−32.30)
Return	Long-short	3.328	5.975	3.593	1.099	−0.708	4.037	6.684	4.301
		(6.07)	(3.21)	(7.26)	(2.45)	(−1.63)	(5.86)	(3.38)	(6.34)
CAPM alpha	Short leg	−15.332	−17.756	−8.293	0.508	11.069	−26.402	−28.826	−19.363
		(−38.11)	(−12.10)	(−19.56)	(1.31)	(38.99)	(−63.48)	(−20.73)	(−43.80)
CAPM alpha	Long leg	−11.802	−11.168	−4.287	2.142	11.355	−23.158	−22.523	−15.643
		(−30.69)	(−9.11)	(−12.18)	(7.02)	(42.39)	(−50.85)	(−18.58)	(−34.75)
CAPM alpha	Long-short	3.530	6.589	4.006	1.634	0.286	3.244	6.302	3.720
		(6.00)	(3.24)	(6.90)	(2.98)	(0.68)	(5.17)	(3.25)	(6.03)

This table extends the main double sorts to stock-quarters with negative and with exactly zero realized PFIT. Negative and Zero denote PFIT below and equal to zero; Low, Middle, and High are terciles among observations with positive PFIT. Within each of the five PFIT groups, stocks are sorted into deciles based on the anomaly signal available at the end of the previous quarter, and portfolio returns are measured during the holding quarter. Panel A reports equal-weighted portfolios and Panel B reports value-weighted portfolios. The table reports robust cross-anomaly means, with cross-anomaly t-statistics in parentheses. Returns and alphas are annualized percentage points. The sample is from 2002Q2 to 2023Q4.

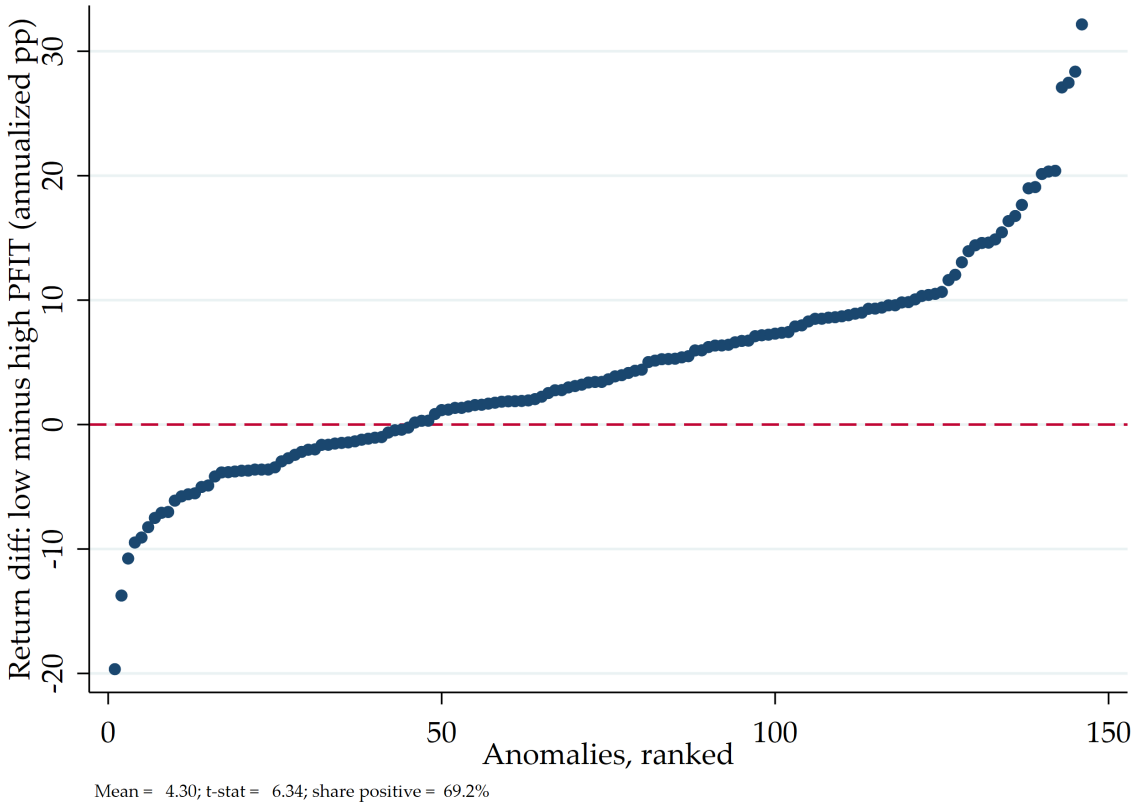


Figure OA1. Low minus high PFIT spread return distribution (value weighted). This figure is the value-weighted counterpart of Figure 3. It ranks the 146 anomaly-level estimates of the difference between the value-weighted long-short return in the low-positive-PFIT tercile and the value-weighted long-short return in the high-positive-PFIT tercile. Portfolios are formed as in Table OA2, and each point is one anomaly. Differences are expressed in annualized percentage points. The reported mean, cross-anomaly t-statistic, and share of positive differences are computed across the 146 anomalies. The sample is from 2002Q2 to 2023Q4.

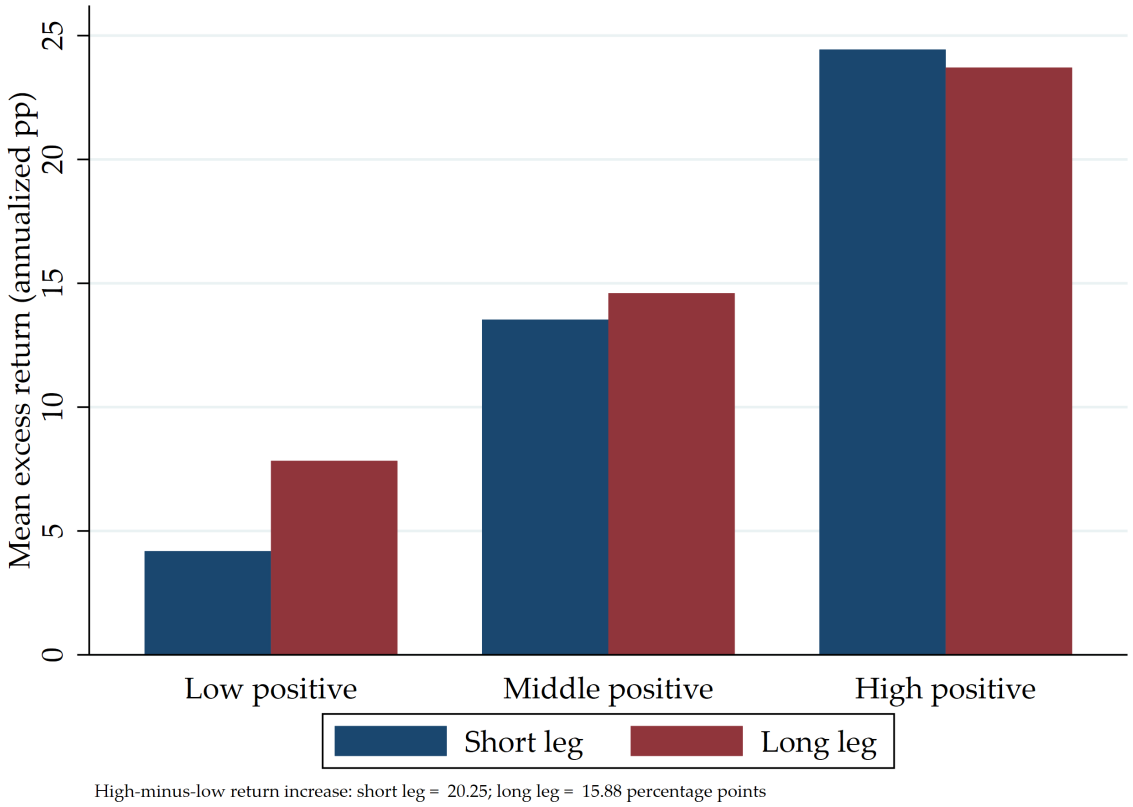


Figure OA2. PFIT and anomaly double-sorted portfolio returns (value weighted). This figure is the value-weighted counterpart of Figure 2. Bars report cross-anomaly averages of anomaly-level mean value-weighted excess returns during the holding quarter for the short leg (D1) and the long leg (D10) in each realized positive-PFIT tercile, expressed in annualized percentage points and averaged across the 146 anomalies. The note below the figure reports the increase in each leg's average return in moving from the low- to the high-PFIT tercile. The sample is from 2002Q2 to 2023Q4.

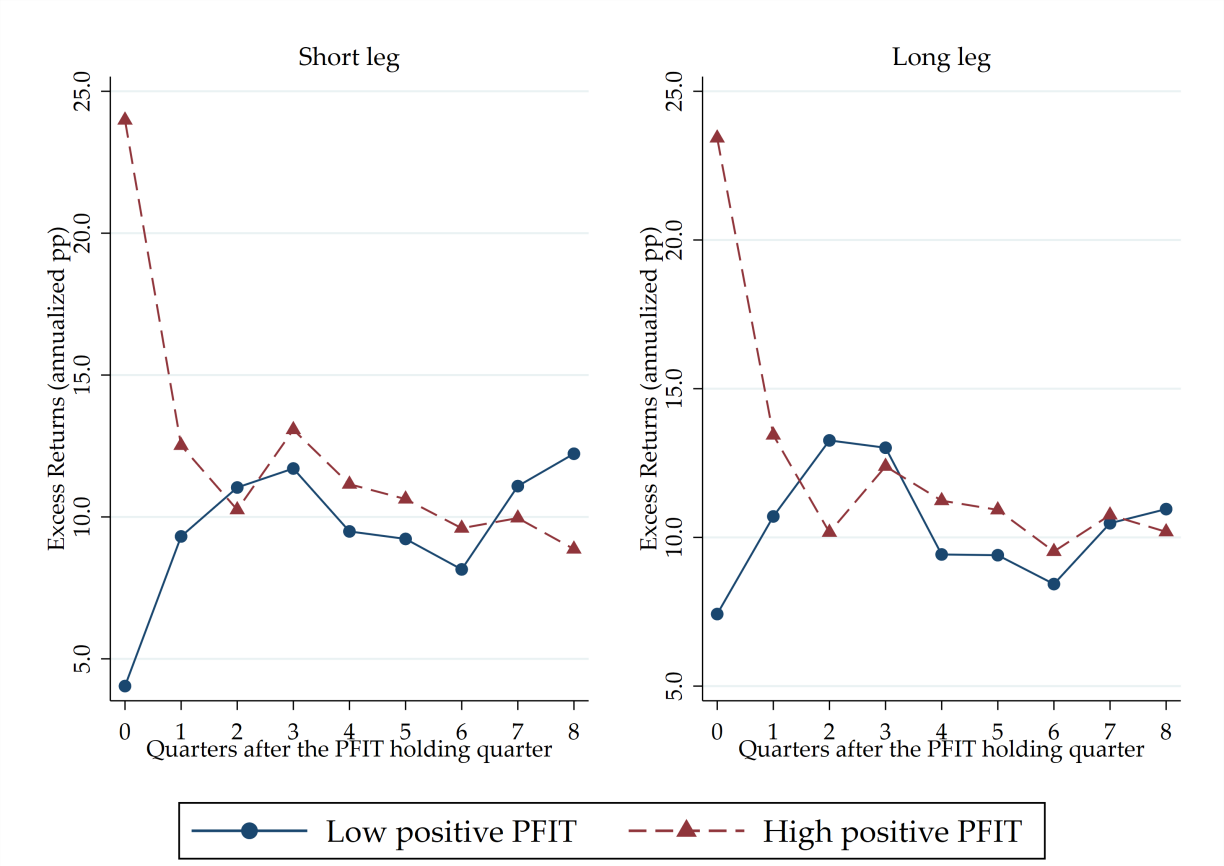


Figure OA3. Long- and short-leg returns across holding horizons (value weighted). This figure is the value-weighted counterpart of Figure 4. Stocks are sorted each quarter into positive-PFIT terciles and anomaly deciles as in Table OA2. The anomaly-decile and PFIT-tercile assignments are then held fixed and portfolio returns are measured for eight additional quarters without re-sorting. Horizon zero is the quarter in which the realized PFIT shock occurs. Returns are value-weighted annualized excess returns at each horizon, averaged across the 146 anomalies. The sample is from 2002Q2 to 2023Q4.

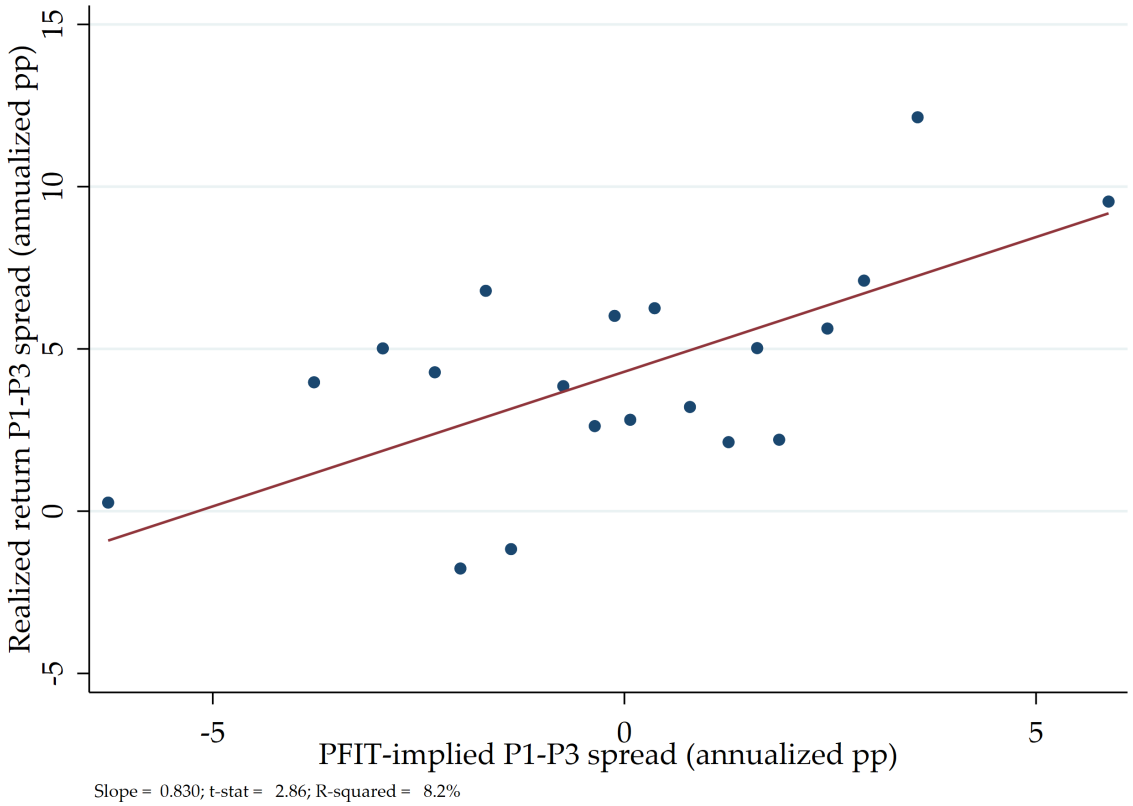


Figure OA4. Implied versus realized low minus high PFIT spread return (value weighted). This figure is the value-weighted counterpart of Figure 5. It plots each anomaly's realized value-weighted low-minus-high PFIT long-short return difference against the difference implied by its own estimated demand multipliers and realized demand shocks, constructed as in Equation (11) and expressed in annualized percentage points. Plotted points are equal-count bins of anomalies formed after ranking on the implied spread; the fitted line, slope, t-statistic, and R^2 are estimated on all anomaly-level observations with heteroskedasticity-robust standard errors. The sample is from 2002Q2 to 2023Q4.

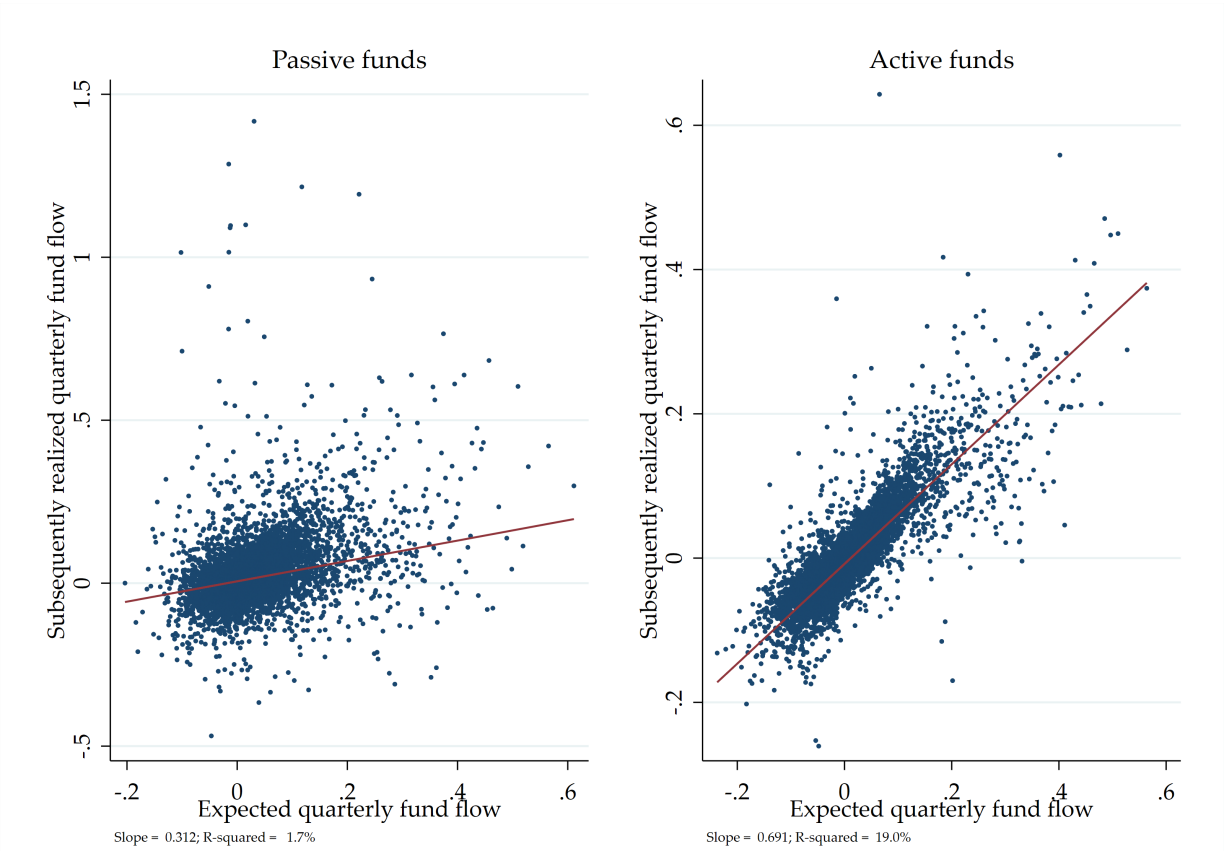


Figure OA5. Predicted and realized fund flows: out-of-sample fit with TNA weights. This figure is the TNA-weighted counterpart of Figure 1. It plots subsequently realized quarterly fund flows against out-of-sample predicted flows, separately for passive funds (left panel) and active funds (right panel), weighting fund-quarters by total net assets. Predicted flows for quarter t are constructed by estimating the cross-sectional forecasting regression in Equation (8) separately in each of the eight preceding quarters, averaging the coefficients, and applying them to each fund's three lagged flows and three lagged returns. Plotted points are equal-count bins used for rendering only; the reported slope and R^2 are estimated on the full unbinned sample. Fund flows are expressed as a fraction of lagged total net assets. The sample is from 2002Q2 to 2023Q4.