# Who Clears the Market When Passive Investors Trade?*

Marco Sammon† John Shim‡

First version: March 29, 2024
This version: July 9, 2024

**Abstract**

Each time a stock is bought or sold by a passive index fund, who takes the other side? We use quarterly holdings, transactions and shares outstanding data from 2002 to 2021 to form 10 mutually exclusive groups, including index funds, active mutual funds, large financial institutions, insiders, short sellers, and firms. We combine a simple regression framework with a market clearing condition to assess who tends to take the other side of trades by passive vehicles. Over the past 20 years across all stocks, *firms* are the largest providers of shares to passive investors on average and on the margin. At the stock-year-quarter level, for every 1 percentage point (pp) change in ownership by index funds, firms take the other side at a rate of 0.64 pp. When restricting to the subsample where index funds are net buyers, firms issue at a rate of 0.95 pp. To isolate inelastic index fund demand, we construct an instrument for each stock-year-quarter using returns of unrelated stocks held in the same funds and leverage the tendency for index fund flows to chase returns. This analysis confirms our estimated magnitudes, and sheds light on the role played by prices in coordinating market clearing between firms and index funds. The instrument also helps rule out other potential drivers like common fundamental shocks and reverse causality. Our results suggest that adjustments in the supply of shares, through, e.g., seasoned equity offerings and employee compensation, are the single-most responsive group to inelastic demand.

**Keywords:** Market Clearing, Index Funds, Passive Ownership, Mutual Funds, ETFs, Active Management, Institutional Investors

**JEL Classification:** G11, G23

*We thank Malcolm Baker, Rob Battalio, Alon Brav, Eric Budish, John Campbell, Alex Chinco, Zhi Da, Robin Greenwood, Sam Hanson, Victoria Ivashina, Ralph Koijen, Tim Loughran, David Mclean (Discussant), Simon Oh, Stefano Pegoraro, Chris Reilly, Andrei Shleifer, Jeremy Stein, Adi Sunderam, Erik Stafford, Philippe van der Beck, Catherine Yoshimoto, and conference and seminar audiences at Four Corners, Harvard Business School, Notre Dame, and University of Cologne for helpful comments.
†Harvard Business School, Harvard University; mcsammon@gmail.com
‡Mendoza College of Business, University of Notre Dame; jshim2@nd.edu

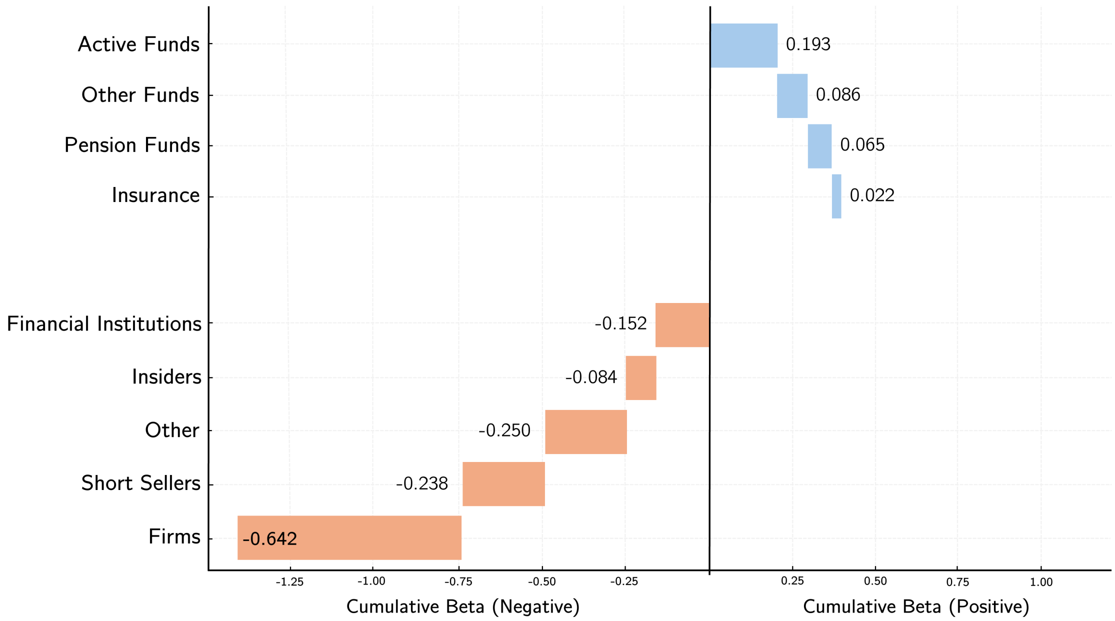
# 1 Introduction

Passive investors are a unique group, operating differently than most other types of investors: they provide semi-regular flows into index mutual funds and exchange-traded funds (ETFs), who then follow mechanical rules on how to allocate those contributions. This combination of regular flows and mechanical rules suggests that a significant proportion of what index funds buy represents plausibly inelastic demand. Further, as a group, index funds have grown significantly, coming to own nearly 20% of the U.S. stock market. A foundation of asset markets, however, is market clearing — with every buyer, there must be a seller. In this paper, we ask: When index funds trade, who takes the other side to clear the market? And, more broadly, who has accommodated the rise of passive ownership from 2% to nearly 20% of the market over the last 20 years by selling a significant quantity of shares?

To answer this question, we combine a simple regression framework with a market clearing condition, which allows us to quantify how other types of investors trade relative to index funds. We apply this methodology to quarterly change-in-holdings and transaction data for each stock included in at least one major index from 2002 to 2021. From these data, we construct quarterly position changes for several mutually-exclusive groups, designed to account for every share that must have changed hands: Index Funds, Active Funds, Other Funds, Pension Funds, Insurance, Financial Institutions, Insiders, Other (which includes retail investors and small institutions), Short Sellers, and Firms. Throughout the paper, we capitalize our group names to distinguish our specific groups from more common usage of the words — i.e., Insurance vs. insurance. It is important to note that our Firms group captures all the ways in which a stock's shares outstanding can change, including not only direct Firm activity through buybacks and seasoned equity offerings (SEOs), but also less direct channels like employee stock compensation and the exercise of warrants and convertible debt. These less direct mechanisms are quantitatively large, explaining just over 50% of the variation in total Firm activity.

Our main finding is that, for the average passive demand shock – positive or negative – Firms are the most important group in clearing the market. In terms of magnitudes, when passive investors demand 1 percentage point (pp) of a stock's shares outstanding – either buying or selling – Firms on average accommodate by adjusting the supply of shares by 0.64 pp. Figure 1 provides a graphical illustration of who clears the market for the marginal unit of Index Fund demand. Most large demand-side institutions tend to buy and sell in the same direction as Index Funds. The lone exception is Financial Institutions, which tend to trade in the opposite direction of Index Funds, though at a rate less than Firms and Short Sellers. Collectively, the supply side – i.e., Firms and Short Sellers – provides 0.88 pp for every 1 pp of Index Fund demand.

**Figure 1:** Who Clears the Market for 1 pp of Index Fund Demand?

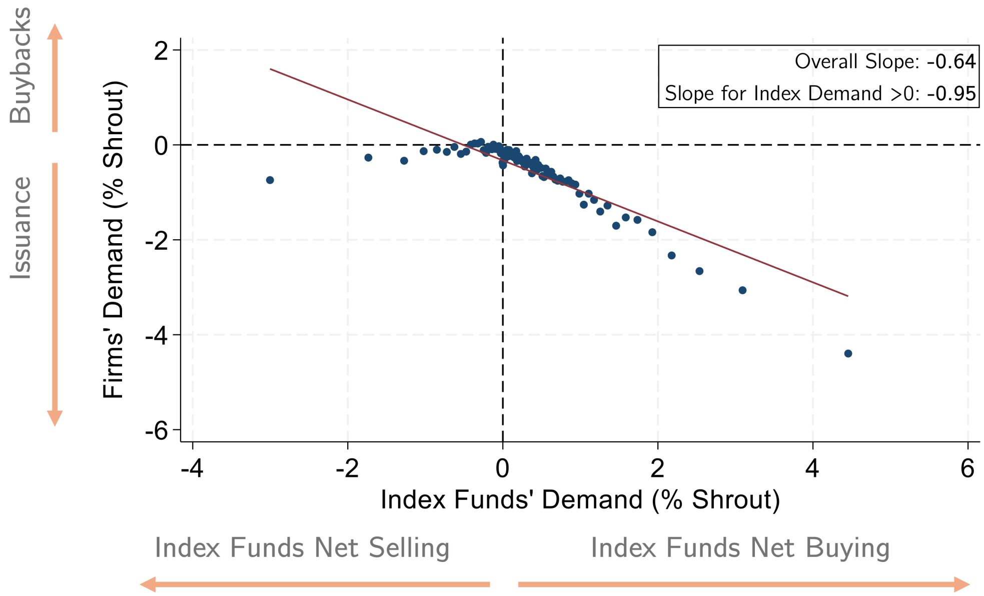



*Notes.* We plot how much each of our 9 groups contributes to clear the market for 1 pp of Index Fund demand (i.e., each group's "beta") at the stock-year-quarter level. We provide the beta for each group and collect all positive beta groups on the top half and negative beta groups on the bottom. A group with a positive beta indicates that when Index Funds are, for example, net buyers, the group also tends to buy. A group with a negative beta tends to sell when Index Funds buy and buy when Index Funds sell. See Section 2 for details on the data and methodology. See Section 3 for details on the results.

The propensity for Firms to take the other side of Index Fund demand is not symmetric with respect to Index Fund buying and selling. Specifically, there is a strong tendency for Firms to provide shares when passive investors are net buyers, but Firms' responses are muted when Index Funds are net sellers. Figure 2 plots the net change in shares outstanding (i.e., additional shares supplied or removed from the market by Firms) against the change in holdings by Index Funds, all in units of percent of total lagged shares outstanding. The overall slope corresponds to our estimate of -0.64, representing how Firms on average take the other side of Index Fund demand. The figure also highlights the asymmetry – in stock-quarters where Index Funds are net purchasers, the estimated slope is -0.95. That is, when Index Funds are net buyers, the market ultimately clears from Firms providing shares at nearly a one-for-one rate.

Traditionally with inelastic demand, we think of prices as the mechanism to facilitate market clearing – buyers will push up price until they can find a willing seller (in our case, Firms). There are alternative stories, however, including an omitted variable (e.g., common fundamental shocks which coordinate Firm and Index Fund demand) and reverse causality (index funds need to mechanically adjust their holdings in the face of primary market activity by firms). We introduce a novel instrument to show that these alternatives are unlikely to be the main drivers of our market clearing facts.

**Figure 2:** Firm Demand vs. Index Fund Demand




*Notes.* We show a binscatter plot of Firm demand against Index Fund demand using stock-year-quarter observations of changes in shares outstanding (Firm demand) and changes in Index Fund positions in shares (Index Fund demand) between 2002 and 2021. We divide by lagged shares outstanding so the units are in percent ownership of a stock. The binscatter plots divides the observations into 100 equally sized bins and plots the average within each bin. Positive and negative Firm demand represent net buybacks and net issuance, respectively. Positive and negative Index Fund demand represent net buying and selling across all Index Funds. We report the slope of a regression of Firm demand on Index Fund demand with all observations (red line) and conditional on Index Fund buying. See Section 2 for details on the data and methodology. See Section 3 for details on the results.

The instrument leverages the tendency of index fund investors to chase returns, isolating the returns from the subset of stocks that are fundamentally unrelated to a focal stock, but held in the same funds. The intuition is that the focal stock will have Index Fund buying or selling because unrelated stocks had either positive or negative returns. These unrelated returns trigger flows into those funds, which leads to buying or selling of all holdings – as index funds scale up and down their holdings proportionally in response to flows – including the stock in question. Therefore, our instrument is designed to isolate passive flows unrelated to the focal stock's own performance.

We test two instruments built on this logic and find an estimated Firm sensitivity of -0.86 and -0.88, similar to our baseline estimate of -0.64. These tests suggest that an omitted variable and reverse causality are unlikely to be the main drivers our results, and that Firms are responding to price pressure induce by passive demand. In addition, the tests further support that a significant fraction of passive demand is inelastic and/or market clearing from inelastic passive demand is similar to market clearing for all passive demand.

We also provide more direct evidence that the price mechanism facilitates market clearing, and that Firms are the most responsive to prices. Specifically, we augment our baseline empirical framework with an indicator for positive excess earnings yield – designed to capture relatively high or low prices (Ben-David and Chinco, 2024). We find that Firms are half as responsive in accommodating passive demand when their excess earnings yield is positive. That is, if prices are not high enough, Firms are less likely to respond and clear the market. Market clearing, however, must still hold, and we find that in these cases, Financial Institutions tend play a larger role.

The responsiveness of Firms to inelastic passive demand then begs the question – what is the source of shares provided by Firms? We decompose Firm demand into contributions from buybacks, SEOs, and compensation/other sources. We find that SEOs and compensation/other each account for about half of the responsiveness of Firms, with buybacks playing a smaller role. It is reassuring that changes in shares outstanding from sources other than buybacks and SEOs play an important role, given that issuance linked to compensation represents option-like claims to a company's cash flows that have not yet materialized in the form of shares. That is, employees can exercise options in response to high prices that result from inelastic Index Fund buying, but do not have a built-in mechanism to induce Firms to buy back shares when prices fall as a result of Index Fund selling. This option-like claim maps nicely to the asymmetric response of Firms to Index Fund demand in Figure 2.

Our paper's first major contribution is to establish a new fact: Firms are the single largest group in accommodating Index Fund demand, both on average and on the margin.[1] Second, we establish that Firms are a much more active participant in responding to other investors' trading activity than previously understood. Given our support for prices as the market clearing mechanism, our findings imply that Firms are likely important for clearing the market not only for passive inelastic demand but also inelastic demand in general. Third, our results suggest that a significant proportion of Firm's issuance comes not from companies making decisions to adjust the supply of shares via buybacks and SEOs, but from stakeholders that have existing but yet-to-materialize claims on company shares (e.g., employees and convertible debt holders). Lastly, the facts we establish have implications for demand-system models in asset pricing (Koijen and Yogo, 2019; Gabaix and Koijen, 2021), which often assume the supply of shares is fixed. One way to interpret our findings is that, without the responsiveness of Firms, prices would have risen even further in the face of consistent passive demand for shares over the past 20 years.

[1] This is a statement about the equal-weighted average stock-year-quarter. As a group, Firms have bought back shares on net in dollar terms over the past 20 years. Over this same period, Index Funds were also net dollar buyers of shares. Therefore, Firms cannot have cleared the market for *average* passive demand on a value-weighted or dollar basis. Still, for the *marginal* unit of passive demand, even on a value-weighted basis, we find that Firms remain the most responsive group. See Section 3.2.1 and Appendix B.3 for more details.

## 1.1 Related Literature

Our work has implications for several areas of research in asset pricing and corporate finance. First, an old literature has argued that demand shocks unrelated to fundamentals should have no effect on prices (Scholes, 1972). More recent evidence, however, suggests that even non-fundamental demand shocks are crucial for explaining asset price fluctuations (Koijen and Yogo, 2019; Gabaix and Koijen, 2021). We uncover an important part of the story, identifying which groups are on the other side of every passive demand shock over the past 20 years. Our findings add to a growing literature showing that the elasticity of investors who provide liquidity to passive demand and investor heterogeneity matters for asset prices (Van der Beck, 2021; Haddad et al., 2022; Balasubramaniam et al., 2023). In these papers, however, the supply of shares is not endogenized, i.e., the firm and short sellers are not modeled as being responsive to prices. We show that neglecting the role of the firm (and the whole supply side) in market clearing – especially in the case of buying by passive index funds and ETFs – omits a potentially central player in these demand systems.

In addition, our finding that firms are the ones that clear the market in the face of passive demand has implications for corporate finance. Fama and French (2005) find that issuance is large and very common. Past literature has also shown that firms tend to issue equity when they think their equity is overvalued (Baker and Wurgler, 2002), substitute between debt and equity depending on their relative valuations (Ma, 2019), and issue debt at specific maturities in response to maturity gaps in government debt (Greenwood et al., 2010). We add to this evidence, showing that firms respond not just to information about future fundamentals (e.g., revenue and earnings), but also to inelastic demand and price pressure. These findings have broader implications for the real effects of passive ownership, including its influence on capital structure and payout policy.

Our work is also related to the literature that connects the growth in passive ownership with company buybacks and issuance. Brav et al. (2024) study the growth of institutional ownership by the "Big Three" institutions that primarily offer passive investment vehicles, and they touch on topics including flows, fees, and how firm activity itself can affect passive growth.

More broadly, our paper contributes to a large literature on the effects of inelastic demand by passive funds. Many papers have focused on changes in index membership (see e.g., Madhavan (2003), Petajisto (2011), Chang et al. (2015), Coles et al. (2022), Van der Beck (2021)), which while important, account for a relatively small share of total trading by passive investors. In this paper, we develop a methodology to study every stock-level quarter-over-quarter change in passive ownership. Importantly, we show that the process for market clearing around index changes is not representative of the average way the market accommodates demand from passive investors. This suggests that the results from studies focused on index changes may not generalize to buying and selling by passive funds in response to flows, which are the predominant source of dollar buying and selling by passive funds.

There are two papers that directly study the connection between passive investment vehicles and equity issuance. Tamburelli (2024) studies the effect of index fund demand on firms' behavior in the context of S&P 500 additions and – in line with our findings – shows that firms are the most important provider of shares. Evans et al. (2023) find that greater ETF ownership increases the probability of an SEO. Both papers compliment our broader approach, where Firms emerge as the single-most important group in clearing the market for passive demand.

Finally, perhaps the most closely related paper to ours from a methodological perspective is McLean et al. (2020), who also conduct a market clearing exercise, examining the changes in holdings by 9 groups of investors. Their paper, however, is focused on the implications for return predictability. Specifically, they aim to understand whether any particular group's buying/selling is related to future expected returns and anomalies. Our focus is instead on the market clearing itself, in terms of which investors are likely to take the other side of trades with passive ownership – and how this may vary depending on the direction of passive trading, the reason for passive trading and across time. To this end, we develop a novel regression framework to quantify the average tendency by our 9 non-passive groups to take the same or opposite side as net passive demand.

# 2 Data & Empirical Methodology

In this section, we outline the data sources we use. We then describe how we construct our mutually-exclusive "investor" groups, for which we will estimate buying or selling aggregated at the group level for each stock in each quarter. Lastly, we provide our empirical regression methodology to document who clears the market when passive investors buy or sell.

## 2.1 Data

We outline each of our data sources below and highlight some data cleaning decisions. Appendix A.1 contains a detailed description of the data and our data cleaning steps.

**Mutual Fund Holdings Data** We use the Thomson Reuters S12 data for quarterly holdings in individual stocks of mutual funds, exchange-traded funds (ETFs), closed-end funds, and unit investment trusts. We separate all funds into three categories: index (passive), active, and other. We classify a fund as an index fund based on the index fund flag and the fund name in the CRSP mutual fund database using the method in Appel et al. (2016). We classify a fund as active if it is in the universe of funds that can be linked between the CRSP mutual fund dataset and the Thomson S12 dataset using the WRDS MF links database but it is not otherwise classified as passive. Any remaining funds that cannot be matched between Thompson and the CRSP mutual fund database are what constitute our "other funds" group. As discussed in Sammon and Shim (2023), the prevalence of stale filings can create problems when working with changes in holdings. To address this issue, we linearly interpolate holdings of each stock at the fund level across stale quarters, for up to 3 quarters.

**Institutional Holdings Data** We obtain data on institutional investors' holdings from 13F filings recorded in the Thomson Reuters S34 dataset. Institutions are required to file a 13F if they hold more than $100M in qualified securities. To classify institutional investors into groups, we use the 13F classification procedure in Bushee (2001), and data for the classification from Brian Bushee's website. The classification assigns each institution to one of the following categories: banks, investment companies, independent investment advisors, insurance companies, corporate pension funds, public pension funds, university and foundation endowments, and miscellaneous.

**Stock Data** We use the CRSP monthly stock dataset for data on shares outstanding. While shares outstanding is not traditionally an "investor" category, to completely account for all sources of share changes, we also include the firm itself, which can issue or buy back shares. We identify net share issuance/buybacks based on changes in split-adjusted shares outstanding. Importantly, CRSP shares outstanding does not include treasury or authorized shares.[2] This means that if stock awards have been approved by the board of directors and shares have been authorized – but are held in the firm's treasury stock – those shares will not be counted toward shares outstanding. These shares will only affect net issuance when they are actually awarded to employees and, thus, available in the market if those employees choose to sell shares.

A concern in our setting is that the shares outstanding data in CRSP can be stale – and therefore we may be mismeasuring the timing of Firms' net issuance. Specifically, shares outstanding data for stocks primarily listed on Nasdaq (CRSP exchange code 3) is updated daily, based on data directly from Nasdaq itself. But,

[2]See the CRSP US Stock Data Description Guide (page 125) for more details.

for stocks primarily listed on other exchanges (e.g., those listed on the NYSE), shares outstanding data may only be updated monthly or quarterly. We have three robustness checks to ensure that potentially stale shares outstanding data is not driving our findings. First, we utilized shares outstanding data from FactSet – which uses data from SEC filings to retroactively update shares outstanding data based on when the change actually occurred – and found nearly identical results. Second, we restricted our analysis to Nasdaq-listed stocks, where we are confident the data is not stale, and again observed similar results. Finally, in Section 3.2.2, we examine net issuance at longer horizons, which should be less susceptible to stale data concerns, and again find our results are qualitatively unchanged.

**Short Interest Data** We use short interest data from Compustat following the method in Hanson and Sunderam (2014). We find that the short interest ratio computed using Compustat data is highly correlated with the short interest ratio reported by S&P Global's Markit database.

**Insider Transactions Data** We use the Thomson Reuters Insiders dataset for changes in holdings for company insider transactions, which we aggregate to the stock and quarter level.

**Index Constituents Data** We obtain S&P 500 and S&P 1500 membership data directly from S&P, while we get data on S&P MidCap 400 and S&P SmallCap 600 membership from Siblis Research. The data from S&P data starts in 2002 and ends in 2021. We also use data from Siblis research to determine Nasdaq 100 membership, which is available from 2015 to 2021. Russell index membership data is obtained directly from FTSE Russell, and runs from 2009 to 2021. Finally, CRSP index membership is provided directly by CRSP, and runs from 2015 to 2021.

**Data Filters** Our main analyses study the period 2002 to 2021. To be included in our sample, stocks must pass several filters. First, we only include ordinary common shares (CRSP share codes 10-11) traded on major exchanges (CRSP exchange codes 1-3). Second, we exclude stocks that are an acquiring permno or have an acquiring permno in either quarter $t$ or quarter $t-1$. In such quarters, there can be large changes in split-adjusted shares outstanding, which can create extreme outliers in company issuance activity. We require that each stock is included in one of the major index families (S&P 1500, Russell 3000 or CRSP Total Market) in quarter $t$, $t-1$ or $t+1$, because our primary objective is to study market clearing when index funds trade.

## 2.2 Investor Groups

We describe how we use the data described in Section 2.1 to measure position changes for each investor group in each stock and quarter. We first discuss how we address potentially overlapping data (e.g., the same shares being reported both in S12 and 13F filings), then we describe how we construct our investor groups.

Note that throughout the paper we capitalize the name of each of the groups we define below. We do this because many group names are also common finance-related words, and capitalization for group names distinguishes our specific groups to more common usage of the words (e.g., Insurance vs. insurance).

### 2.2.1 Eliminating Overlapping Groups

Nearly all mutual funds and ETFs from the S12 mutual fund holdings data are under the umbrella of some financial institution that also files a 13F. This means that mutual fund holdings and 13F holdings have significant overlap. We address this overlap by first identifying all of the 13F institution categories that could include mutual fund holdings in their filings: banks, investment companies, independent investment advisors, and miscellaneous. We then combine the holdings of these four 13F categories together into a single category and subtract all S12 holdings. The remaining holdings form a single category that represents financial institutions excluding fund holdings. This is done for each stock in each quarter.

We see this as a conservative way to avoid double counting the holdings of mutual funds and other funds. We think of the holdings of these financial institutions excluding mutual funds as representing the holdings of hedge funds, large family offices, the proprietary arms of banks, and other large institutional investors.

### 2.2.2 Constructing Groups

We form 10 mutually exclusive groups. The first set of investor groups comes directly from holdings data, where we can observe quarter-over-quarter changes in holdings per stock, aggregated within group. We use mutual fund and ETF holdings to form the first three groups: Index Funds, Active Funds, and Other Funds (which are the funds in the Thomson Reuters S12 dataset that cannot be matched with the CRSP mutual fund dataset). Throughout the paper, we will often use the term "passive investors" when describing the aggregate Index Fund group.

We use 13F holdings to form three other groups: Insurance (which combines insurance companies and university & foundation endowments), Pension Funds (which combines corporate and public pension plans), and Financial Institutions.[3]

For each of these groups, we measure the net buying or selling for a group in each stock and in each quarter by aggregating (adjusted) shares held at the stock and quarter level across all investors within a certain group. We then examine changes in holdings at the stock and quarter level for each group and normalize by shares outstanding in the previous quarter. This gives the aggregate buying or selling activity for a group in a stock-quarter, measured as percent ownership of a company. We define $q_{i,j,t}$ to capture the aggregate buying or selling activity, and is given by

$$q_{i,j,t} = 100 \cdot \frac{shares_{i,j,t} - shares_{i,j,t-1}}{shrout_{i,t-1}}, \tag{1}$$

where $shares_{i,j,t}$ is the adjusted shares held in stock $i$ by group $j$ in quarter $t$ and $shrout_{i,j,t-1}$ is the shares outstanding in stock $i$ in quarter $t-1$.

As an example using our Index Fund group, if some index funds and ETFs buy shares in a particular stock-quarter and many other index funds and ETFs sell shares, the Index Fund group buying/selling $q_{i,j,t}$ captures net buying or selling across all index funds and ETFs. If there is some imbalance within the Index Fund group, one or more investor groups collectively must have an imbalance of the opposite sign to clear the market.

The next investor group we measure is Insiders. We do not have data on insider holdings. Instead, we use insider buys and sells to record aggregated insider transactions at the stock-quarter level. For insiders, we have

$$q_{i,Insiders,t} = 100 \cdot \frac{\sum_{i,t} insiderbuys_{i,t} - \sum_{i,t} insidersales_{i,t}}{shrout_{i,t-1}}, \tag{2}$$

where $insiderbuys_{i,t}$ and $insidersales_{i,t}$ are individual insider transactions in units of split-adjusted shares.

The next two groups account for possible changes in the supply of available shares in the market. The most direct way for the supply of shares to change is for the firm itself to either issue shares or buy back shares. For example, if Index Funds buy 1,000,000 shares of a stock and no other groups trade, it is possible that the firm issues 1,000,000 additional shares, which allows the market to clear. We label this group "Firm,"

[3]We combine insurance companies and university & foundation endowments because they are both very long-horizon investors and because the endowments group is relatively small. We also combine corporate and public pension plans because they have common objectives. Financial Institutions excludes all fund holdings, as described in Section 2.1, to avoid double counting.

and use changes in shares outstanding as one measure of changes in the number of available shares, or

$$q_{i,Firm,t} = 100 \cdot \frac{shrout_{i,t-1} - shrout_{i,t}}{shrout_{i,t-1}}. \tag{3}$$

For consistency with other groups, we represent Firm issuance as "selling" (i.e., negative changes), which requires another group to buy in order to clear the market. We similarly label buybacks as Firm "buying".

The supply of shares can also change through increases in short interest. If an index fund buys 1,000,000 shares and no other groups adjust their positions (including the Firm), a hedge fund or another institution may sell shares short by borrowing from existing holders. For the purposes of market clearing, the accounting of shares is similar for short selling and Firm issuance. For example, suppose that Index Funds buy 1,000,000 shares, which a hedge fund accommodates by borrowing 1,000,000 shares and selling them. In this scenario, the hedge fund has a net position which is more negative by 1,000,000 shares. We label this group "Short Sellers," and use changes in short interest as a measure of this group's buying and selling, or

$$q_{i,ShortSellers,t} = 100 \cdot \frac{shortinterest_{i,t-1} - shortinterest_{i,t}}{shrout_{i,t-1}}. \tag{4}$$

To be clear, Short Sellers' position changes are signed just like all other groups – Short Sellers' buying corresponds to decreases in short interest, and selling corresponds to increases in short interest. It is also important to note that many 13F institutions, especially our Financial Institutions group, are likely to be short sellers. However, 13F holdings only capture ownership. So one way to think of our Short Sellers group is that it is separating our many other groups' short positions from their long positions (which are captured in the groups defined above).

We list our 10 groups in Figure 3. The first 9 groups provide a careful accounting of all of the ways in which shares can change hands in the data. However, we know that our data are incomplete and do not account for every share of every stock. A nontrivial fraction of holdings, and thus of share changes, will be missed because of the omission of retail investors and institutions that are too small to file 13Fs, as well as other groups we cannot measure directly. Because we know that the market must clear, we attribute the remainder of share changes to a group that we call "Other" to ensure that the market clears. That is, if the first 9 groups above collectively are net purchasers, then it must be that our Other group is a net seller for exactly the same number of shares. In this sense, we can think of the Other group as a residual group that enforces market clearing. This residual Other group is the 10th and final group. See Appendix A.2 for an illustration of how our Other group is constructed.

**Figure 3:** Investor Groups

- Mutual Funds
  1. **<u>Index Funds</u>**: Passively-managed index funds
  2. **<u>Active Funds</u>**: Actively-managed funds
  3. **<u>Other Funds</u>**: Funds which cannot be linked between Thompson S12 and the CRSP Mutual Fund data
- 13F Institutions
  4. **<u>Financial Institutions</u>**: All financial institutions excluding mutual fund holdings
  5. **<u>Insurance</u>**: Insurance and endowments
  6. **<u>Pension Funds</u>**: Corporate and public pension funds
- Share Suppliers
  7. **<u>Firms</u>**: Share buybacks and issuance
  8. **<u>Short Sellers</u>**: Changes in the effective supply of shares from either increases in short interest or short covering
- Miscellaneous
  9. **<u>Insiders</u>**: Company insiders required to disclose share transactions
  10. **<u>Other</u>**: Retail, foreign institutions, small institutions and other groups we cannot measure directly. A residual category to enforce market clearing

The Other group can still be economically relevant in that it can accurately capture the economic activity of small investors. For example, one type of investor that must reside in this category is retail traders. Although we do not claim that our Other category is a direct quantification of total retail trading activity, we do find that our Other category's position changes are correlated with proxies of net buying and selling by retail investors. We document these patterns in Appendix A.2.3.

Since the Other group guarantees market clearing, it may also be affected by data errors. For example, if there is an erroneous position change in the Pension Funds group, this error will force the residual Other group to take the opposite position to clear the market. We will keep this in mind when assessing the empirical results and conduct analyses that try to assess the degree to which this group captures economic activity or data errors.

### 2.2.3 Outliers

Lastly, we trim outliers in each group below the 0.5 percentile and above the 99.5 percentile in our data. That is, we simply delete stock-quarter observations if any one of our groups has a percentage share change that is in the extreme tails of that group's percentage position change distribution. Because some the extreme observations are overlapping across groups, this ends up removing about 7.5% of security-quarter observations. Of course, it is possible that some groups drastically change their positions and that this filter erroneously deletes "true" observations from our dataset. However, we suspect that many of these outliers come from data errors in the underlying holdings data.[4] In addition, the magnitudes for a subset of the data for some groups are improbably large, even after correcting issues with stale data as described in Section 2.1.

If the outliers are the result of data errors, replacing them with the 0.5 or 99.5 percentile values – i.e., Winsorizing the data – may not do much to correct the errors. Moreover, an error for one group will have knock-on effects in our Other category. Specifically, in order to clear the market, it will force the residual Other group to also make an unrealistically large offsetting position change. This will affect the overall estimation of betas and lead to less informative inference. For these reasons, we delete these observations for our analysis.

Nevertheless, as a robustness test, we repeat our analysis in Section 3 using (1) the raw data and (2) Winsorizing the outliers within each group at the 0.5 and 99.5 percentiles. These alternative specifications yield results that are qualitatively similar to our main results. See Appendix B.5 for these additional results.

## 2.3 Empirical Methodology

Because passive funds have plausibly inelastic demand, we treat their position changes as a starting point to understand who clears the market. That is, our framework is built on the idea that passive funds demand liquidity by initiating the buying and selling. We develop a simple methodology to understand who takes the other side of passive trades.

It is important to note that our methodology is designed to identify long-term buyers and sellers of shares. While it is common for intermediaries, such as algorithmic trading firms and financial institutions, to be the most common counterparty on a day-by-day or trade-by-trade basis, their goal is often not to hold any portfolio in particular but to bridge the gap between final buyers and sellers. As a result, we will often not capture these types of investors. However, to the extent that these intermediaries take on strategic long-term bets and thus hold shares (or adjust their holdings of shares) over a quarter, they too will be counted as playing a role in market clearing.

Our methodology is built on estimating a series of univariate regressions, one for each of the nine, non-Index

[4]See Appendix B in Sammon and Shim (2023) for a detailed analysis of the mutual fund holdings data.

Fund groups, of the form

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}. \tag{5}$$

$\beta_j$ represents the degree to which group $j$'s demand responds to Index Fund demand. Throughout the paper, we use the language "Index Fund demand," "passive demand," "Index Fund changes," and "passive changes" interchangeably when referring to $q_{i,\text{IDX},t}$.

Because the market must clear in each stock each quarter, some other group or groups must be on the other side of passive changes. This can be mathematically represented as

$$\sum_j \beta_j = -1. \tag{6}$$

In words, Equation 6 says that all other groups in aggregate must collectively take the other side of passive position changes on a one-for-one basis due to market clearing. See Appendix A.3 for a derivation of this market clearing expression. In addition, the market must also clear for average passive demand. That is, the average Index Fund change across stocks and quarters as a percentage of lagged adjusted shares outstanding, denoted $\overline{q}_{\text{IDX}}$, must equal the sum of the average group change across stocks and quarters, or $\overline{q}_j$.

The estimated alphas and betas from Equation 5 tell us two things about market clearing. First, they tell us how the market clears *on average*. For example, if Active Funds have an average change of 0.011, this means that Active Funds collectively purchase 0.011% of the total ownership of the average stock each quarter.[5]

The second thing we can learn from estimating Equation 5 is each group's *sensitivity* to Index Fund changes. The beta tells us how much each group's demand varies as a function of Index Fund demand. Specifically, for each additional percentage point of a stock's shares outstanding bought by Index Funds, group $j$ tends to adjust their demand by $\beta_j$ percentage points of shares outstanding. A group with positive beta means that, as Index Funds increase demand for a given stock-quarter, this group also increases its demand in that stock-quarter. This is also why the sum of the betas enters into the market clearing condition of Equation 6 – if one group demands more shares when Index Funds demand more shares, some other group must demand fewer shares.

[5]The alphas by themselves tell us the fitted-value (i.e., predicted) for a stock-quarter with zero Index demand *given* the estimated sensitivity of each group's demand to Index Funds. For example, Active Funds have an alpha of around -0.05, which means in a quarter with zero passive demand, they are predicted to sell around 0.05% of each stock.

# 3 Who Clears the Market?

In this section, we present our baseline linear estimates from the empirical specification described in Section 2.3.

## 3.1 Baseline Linear Estimates

We estimate Equation 5 for each of the non-passive investor groups to account for how the market clears when Index Funds, in aggregate, add to or decrease their positions. Table 1 reports both the alpha and the beta for each group, as well as t-statistics, the number of observations, and the $R^2$.

In addition, we report the average position change for each group $j$, which we denote as $\overline{q}_j$. In the context of our regression framework, this can be interpreted as $\overline{q}_j = \alpha_j + \beta_j \cdot \overline{q}_{\text{IDX}}$, or the fitted value for $q_j$ at the average level of passive demand $\overline{q}_{\text{IDX}}$. The average equal-weighted quarterly position change for Index Funds is 0.34pp (i.e., a position increase of 0.34% of a stock's total shares outstanding). Therefore, the sum of $\overline{q}_j$ for all $j$ must be -0.34% of shares outstanding, because the shares that passive buys must come from a combination of the other groups.[6]

First, we highlight the average position change for each group $\overline{q}_j$. There are three groups that, for the typical firm, provide shares to Index Funds: Firms (-0.673 pp), Insiders (-0.074 pp), and Short Sellers (-0.053 pp). Note that these three groups collectively sell about 0.8 pp, much more than the 0.34 pp demanded by Index Funds on average. The reason for this is that all other groups, at least on an equal-weighted basis, also add to their positions on average, just like Index Funds.[7] This first result shows that the typical share demanded by Index Funds comes from the supply side. Additionally, Firms and Short Sellers, on average, provide *all* of the shares typically demanded by not only Index Funds, but all other institutional investor groups.

Second, we highlight the sensitivities of each group to Index Fund demand. These estimates give a sense of the equilibrium responsiveness of each group: When Index Funds demand relatively more shares, who provides those *additional* shares? Here, the answer is again the supply side – when Index Funds demand an additional 1pp of a company's shares outstanding, Firms and Short Sellers collectively provide about 0.96

[6]The sum of $\overline{q}_j$s may vary across tables because of subsample analysis or weighting differences. E.g., Table 8 reports value-weighted regressions and uses the corresponding value-weighted average change in passive ownership to compute $\overline{q}_j$ which differs from the equally-weighted average.

[7]It has been well documented that Active Funds in aggregate have seen redemptions over our sample period. This can be observed in our value-weighted analysis in Section 3.2.1, which captures this pattern in $\overline{q}_j$. In fact, given that the equal-weighted $\overline{q}_j$ for Active Funds is positive and the value-weighted $\overline{q}_j$ is negative, this implies that, as a group, Active Funds have tilted their portfolios toward small- and mid-cap stocks, while selling relatively large stocks to satisfy redemptions.

**Table 1:** Beta Estimates

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\overline{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.193 | 7.015 | -0.052 | -2.209 | 172,807 | 0.007 | 0.011 |
| Other Funds | 0.086 | 3.148 | 0.060 | 3.361 | 172,807 | 0.012 | 0.088 |
| Pension Funds | 0.022 | 7.423 | -0.007 | -1.243 | 172,807 | 0.004 | 0.000 |
| Insurance | 0.065 | 7.894 | -0.029 | -2.281 | 172,807 | 0.008 | -0.008 |
| Financial Institutions | -0.152 | -2.265 | 0.123 | 1.445 | 172,807 | 0.001 | 0.073 |
| Insiders | -0.084 | -10.920 | -0.033 | -4.684 | 172,807 | 0.005 | -0.060 |
| Other | -0.250 | -4.674 | 0.228 | 2.918 | 172,807 | 0.003 | 0.146 |
| Short Sellers | -0.238 | -6.433 | 0.037 | 0.802 | 172,807 | 0.015 | -0.041 |
| Firms | -0.642 | -16.196 | -0.327 | -8.757 | 172,807 | 0.032 | -0.537 |
| Total | -1.000 | | 0.000 | | | | -0.327 |

Notes. The table provides estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t},$$

for each investor group $j$. $q_{i,j,t}$ is the quarterly holdings change in stock $i$ for group $j$ in year-quarter $t$ in units of percent ownership of the company. $q_{i,\mathrm{IDX},t}$ is the ownership change for Index Funds. T-statistics are computed from standard errors double clustered at the stock and time (year-quarter) level. The last column reports the average quantity change for each group across all stocks and quarters ($\overline{q}_j$). See Section 2.2 for more details on the investor groups, Section 2.3 for the empirical methodology, and Section 3.1 for more details on the table.

pp. The combination of the market clearing averages and sensitivities suggests that Firms and Short Sellers on average provide all of the shares demanded by Index Funds.

Table 1 also shows that Active Funds, Pension Funds, and Insurance all have negative $\alpha_j$ and positive $\beta_j$. That is, the estimates suggest these groups are predicted to sell on average when Index Funds do not change their positions. These groups' positive betas imply that they tend to buy more of the stocks that Index Funds are buying relatively more intensely. Finally, in the average stock-quarter, these groups all also increase their positions ($\overline{q}_j > 0$).

Financial Institutions have a positive alpha but a negative beta, suggesting that they take the other side of marginal Index Fund demand on average. The residual Other group has a positive intercept but a much larger (in-magnitude) negative beta, suggesting that the residual group is more sensitive to large Index Fund demand shocks than Financial Institutions. We will show that the responsiveness of both groups can be better understood by examining non-linearities in how their position changes vary as a function of Index Fund position changes in Section 3.1.2.

Before proceeding, we would like to highlight some patterns on statistical significance in Table 1 (we double-cluster standard errors at the stock and year-quarter level). The most statistically significant estimates are Firms' alpha and beta, with a t-statistic of -10.1 and -15.8, respectively. So not only are Firms' estimated coefficients the most negative, they are also the most reliably different from zero.

### 3.1.1 Fixed Effects

One concern with the results in Table 1 is that common time-series or cross-sectional shocks are driving our results, instead of Firms *responding* to Index Fund demand. To allay such concerns, we re-estimate Equation 5 including quarter-year, industry-by-quarter-year, and stock and quarter-year fixed effects. These specifications are designed to account for commonalities in Index Fund demand on a variety of dimensions. For example, the year-quarter fixed effects account for heterogeneity in Index Fund demand over time, while the industry-by-quarter-year fixed effects account for common demand shocks to particular sectors of the economy period-by-period.

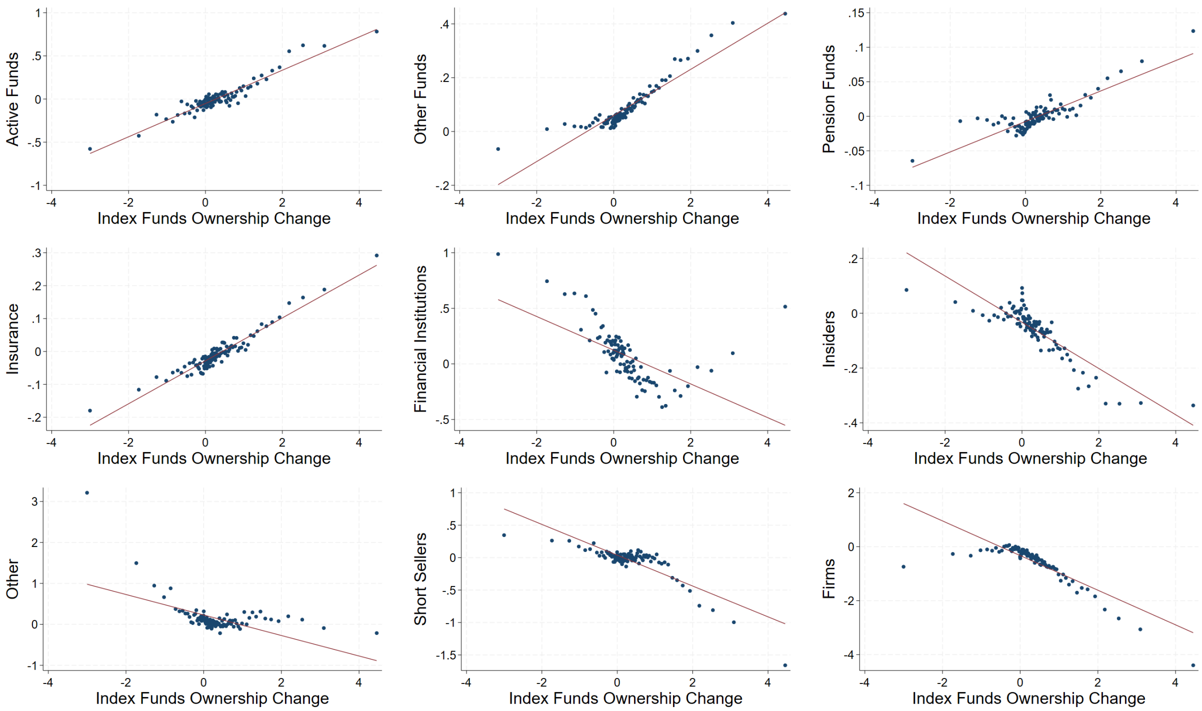
Appendix B.1 provides the regression estimates for each fixed effects specification. Most of the beta estimates are largely unchanged – the supply side still accounts for more than 75% of the marginal shares needed to clear the market for passive position changes, with Firms providing more than half. This suggests that the overarching message of the baseline results also holds within a stock over time. This is in spite of the fact that the fixed effects soak up a significant amount of the variation in position changes. The $R^2$ of nearly every group's regression jumps significantly. In particular, the $R^2$ for Firms jumps from 0.034 in the baseline specification to 0.226 with fixed effects. Although not reported in Table 6, we find that most of the increase in $R^2$ comes from stock fixed effects, not year-quarter fixed effects.

### 3.1.2 Binscatter Plots

To identify possible nonlinearties in the relationship between each group's demand and Index Fund demand, we produce binscatter plots for each group's position change, measured in percent of shares outstanding purchased or sold, as a function of Index Fund position changes. We form 100 bins of equal size for each group, which means that each bin has more than 1,300 observations. These plots illustrate the expected position change of each group conditional on Index Fund changes.

Our first takeaway is that the binscatter plots reinforce some patterns documented in Table 1. The conditional mean for Active Funds, Other Funds, Pension Funds, and Insurance is roughly linear in Index Fund demand. The figures also confirm the significant negative relationship between passive changes and the changes of Firms, Insiders, and Short Sellers. All three exhibit a similar pattern that is not clear from the beta estimates of Table 1: the negative relation is much stronger for positive passive changes than for negative. In fact,

**Figure 4:** Ownership Changes: Binscatter by Group




*Notes.* Each panel presents a bin scatter of net demand by each investor group – $q_{i,j,t}$ – against net demand by Index Funds – $q_{i,\text{IDX},t}$. The unit of observation is security-year-quarter.

for Firms, it appears that, if anything, there is a weak positive relation between Firms' and Index Funds' position changes when passive changes are negative. In Appendix B.2, we show that the estimate for Firm beta is around -1 for positive Index Fund demand and is 0.165 when Index Funds are net sellers.

Figure 4 also uncovers some patterns for Financial Institutions that are not obvious from Table 1 alone. Financial Institutions' position changes seem to have a clear negative relation with passive changes for all but the largest positive changes. For the largest passive changes, it appears that Financial Institutions mimic Index Funds. This suggests that Financial Institutions, which includes the trades of hedge funds, may indeed be on the other side of Index Funds for all but relatively large positive passive position changes.

Lastly, these figures provide a reasonable sanity check on the data: Can we reliably clear the market for Index Funds' position changes with our data-driven groups? Or, do we have to regularly rely on the residual Other group to clear the market? Figure 4 suggests that our results are not heavily dependent on our Other group clearing the market. For all but the extreme negative passive changes, the Other group has nearly a flat relation with passive share changes. That is, we can, on average, roughly capture market clearing amongst the groups in our data with the exception of the extreme negative share changes from Index Funds.

For these extreme observations, the groups in our data do not seem to take the other side of Index Funds and nearly all of our observed passive changes must be cleared by the residual group. In other words, the negative slope for the Other group in Table 1 is driven by stock-quarters where Index Funds are significant net sellers. Given the general trend toward passive buying over the past 20 years, in the cases of extreme passive selling there may be a concurrent event which explains why the groups in our sample do not clear the market. For example, suppose that due to a corporate event (e.g., redomiciling the firm for tax reasons) the stock became ineligible for many types of passive index funds and ETFs. In such cases, many institutions may also have mandates that prevent them from buying such stocks (Beber et al., 2021). This may be exactly the type of case where foreign institutional investors may take the other side, which would show up in our residual Other group.

## 3.2 Robustness Tests

We perform several robustness tests for the main equal-weighted regression results. We briefly describe each test below – and provide the full details and tables in the appendix.

### 3.2.1 Value-Weighted Regressions

To capture market clearing for the marginal *dollar* of passive demand, we re-estimate the regressions from Section 3.1 on a value-weighted basis. With value weights, we find that Firms are still the group that contributes the most to clearing the market for the marginal unit bought or sold by Index Funds. On the other hand, relative to our baseline equal-weighted specification, Active Funds and Financial Institutions play a larger role in clearing the market for average Index Fund demand. We describe the value-weighted results in detail in Appendix B.3.

These findings help establish that our baseline results on marginal market clearing are not dominated by small- and mid-cap companies and are robust to the increased prevalence of buybacks, especially for very large companies. It is important to remember that, while buybacks have drawn attention from the media, 80% of company-quarters have net issuance, and the issuance as a percentage of the overall size of the company is larger on average than buybacks. We provide these supporting facts and greater detail in Appendix D.

#### 3.2.2 Year-over-Year Changes

While we attempt to capture the long-term buyers and sellers of stock, a quarter may still be too short to capture how some groups clear the market for others in the long run. For example, while there are many ways in which new shares can come to the market via Firms (see Section 5), traditional issuance via a seasoned equity offering can take more than a quarter. Appendix B.4 provides the details for estimating Equation 5 with year-over-year changes. The takeaway is that Firms play an even larger role in clearing the market for marginal passive demand with a beta of -1.441. The beta is larger-in-magnitude than -1 because other demand-side funds and institutions have even larger positive betas over longer horizons than they do over a quarter, necessitating some other group (in this case Firms) to clear the market for passive on greater than a one-for-one basis.

#### 3.2.3 Sample Selection and Treatment of Outliers

In Appendix B.5, we provide additional estimates for different sample periods (2009-2021 and 2015-2021) and different treatment of outliers (raw data and Winsorization). Most of the patterns in our baseline sample are also found in these alternative samples. Notably, Firms' beta estimates are even more negative in recent years than over the whole sample. Moreover, Active Funds consistently respond in the same direction as Index Funds, with beta estimates ranging from 0.180 to 0.506.

The most notable differences in these alternative samples are for the Other group when adding back extreme observations. The beta estimates range from -0.642 (winsorized) to -1.325 (raw), a large difference from our baseline estimate of -0.250. This suggests that our residual Other group does more work in clearing the market when including more extreme observations (i.e., when the market does not clear among the investor groups we can directly observe in the data). This provides some suggestive evidence that these extreme observations may be data errors, supporting our decision to remove them from our main sample.

#### 3.2.4 Data Errors

As described in Section 2 and documented in detail in Sammon and Shim (2023), the S12 data are littered with many types of errors, some of which involve staleness in reported holdings. In the Appendix, we discuss two sets of tests designed to address such data errors.

The first type of issue is due to general data errors where a group appears to increase or decrease its ownership of a stock but does not in reality. Such an error will force the Other category to absorb this seemingly unmet demand. In Appendix B.6, we rerun a version of our baseline regressions, limiting the sample to only stock-quarters where the Other group has an ownership change of less than 0.5 pp in magnitude, i.e., stock-quarter observations where we are confident these types of data errors are less likely to alter the beta estimates. Table 14 shows that the results for this subsample are consistent with the baseline results – Firms and Short Sellers collectively account for a significant fraction of the marginal shares demanded by passive.

The second and more specific issue that could contaminate our results is stale data. Suppose, for example, that Index Funds buy a stock in period $t$ and Active Funds sell to them, but Active Funds' sales are erroneously not recorded in the data and stale holdings from the previous quarter are reported instead. Then, the passive buying in quarter $t$ will appear to be cleared by the Other group selling in period $t$. Furthermore, if Active Funds' sales, that actually occurred in $t$, are recorded in quarter $t+1$, the Other group will need to buy to clear the market.

To quantify the effects of possibly stale data on our baseline regression results, in Appendix B.6 we test the degree to which passive changes in quarter $t$ are related to other groups' position changes in the same stock but in quarter $t+1$. Given the example above, in the presence of stale data we expect to observe a negative relation between the residual Other group's position change in $t$ and Index Fund demand at $t$, but a positive relation between the residual group's position change at $t+1$ and Index fund demand at $t$. Table 15 shows that this is indeed the case, but suggests that the degree to which stale data affect our findings is low since the magnitude is significantly smaller in $t+1$ than in $t$ (0.124 vs -0.271).

#### 3.2.5 Over Time

We estimate our set of regressions each quarter and report the beta estimates over time. That is, these betas are only identified using *cross-sectional* variation in Index Fund demand period-by-period. We plot an 8-quarter moving average of our beta estimates in Figure 10. We find that Firms' betas have been getting more negative and closer to -1 over time. See Appendix B.8 for more details.

#### 3.2.6 Thematic and Broad-Based Funds

So far, we have treated Index Funds as one group. However, as discussed in the introduction, one may be concerned that our results are driven by common fundamental shocks which coordinate Index Fund and Firm demand. In this case, one might believe that there are significant differences in our findings for broad-based index funds (e.g., S&P 500 or Russell 1000 funds) and style funds (e.g., industry or factor funds). The logic is that demand for style funds is more likely to be coordinated by common fundamental shocks – because their positions are more concentrated – and therefore these types of funds may be entirely driving our main results.

To see the contribution of different types of index funds, we separate demand in our Index Fund group into demand from broad-based funds and style funds. We draw two main conclusions from this analysis. First, Firms are still the most responsive group to Index Fund demand, regardless of whether we focus on broad-based or style Index Funds. Second, and perhaps more surprisingly, Firms are much more responsive to broad-based than style Index Funds' demand. We believe this is in part because, for broad-based Index Fund buying, there is more overall inelastic demand that needs to be accommodated, as style funds tend to partially mimic the demand of broad-based funds. We provide details in Appendix B.9.

#### 3.2.7 Index Switching Stocks

We analyze market clearing with respect to shocks to index inclusion. We provide two sets of results, which we describe in detail in Appendix B.10: (1) Estimates for stocks that did or did not switch from one major index to another (e.g., switches within the Russell or S&P family of indices), and (2) more granular estimates based on switching from a specific major index to another (e.g., Russell 1000 to 2000) or direct additions or deletions (e.g., migrations into or out of the S&P 1500 universe).

We find that the results for index stayers largely tell a similar story to our baseline findings: Firms account for most of the other side of passive demand, with Short Sellers and Insiders also consistently contributing to clearing the market for Index Fund buying/selling. Market clearing, however, is quite different for index switchers. Most notably, Firms play a much larger *average* role: $\overline{q}_j$ for Firms is -1.082pp, nearly all of the 1.102pp average Index Fund demand for index switchers. However, Firms are much less *responsive*, with a beta estimate of -0.201 (as opposed to -0.763 for stayers). This analysis is roughly consistent with the results in Tamburelli (2024), who shows that firms accommodate inelastic demand from index trackers around S&P 500 index inclusion events.

# 4 Mechanism and Identification

So far, our results establish a strong negative correlation between Index Fund demand and Firm demand. In this section, we aim to understand the mechanism behind this relationship, specifically *why* Firms appear to respond to Index Fund demand by issuing shares. We propose three likely channels that could generate the observed relationship in the data.

Our preferred mechanism, and the one that is most supported by the evidence, is that prices coordinate market clearing. Specifically, when Index Fund demand shocks occur, due to their inelastic nature (Haddad et al., 2022), prices increase. In response, Firms issue shares to capitalize on these high prices (Baker and Wurgler, 2002; Dong et al., 2012; Ma, 2019).

Another possible mechanism is that common fundamental shocks coordinate Firm and Index Fund demand. For example, the commercialization of AI (Eisfeldt et al., 2023) might drive technology firms to invest in GPUs and hire new employees, which they fund by issuing equity and attract human capital by granting stock to employees. Simultaneously, investors excited about the potential of commercialized AI may increase their demand for ETFs that hold stocks poised to benefit from AI technology. In this case, an omitted variable – namely, the common technology shock – coordinates Index Fund and Firm demand.

A final possible mechanism is reverse causality, due to the mechanical way Index Funds trade in response to Firm issuance and buyback activity. For instance, when a Firm issues shares, the number of "index-eligible shares" increases, leading to mechanical buying by Index Funds.

To isolate prices as the coordinating device for Index Fund and Firm demand, we propose an instrumental variables (IV) specification. Our IV results suggest that the effect of Index Fund demand on prices, rather than omitted common fundamental shocks or reverse causality, is the primary driver of our OLS regression results in Section 3.1. We also provide direct evidence for the price channel, demonstrating that the relative valuation of a stock – as measured by excess earnings yield – is a strong predictor of how Firms respond to Index Fund demand. We conclude this section by discussing a series of additional tests designed to rule out the mechanical/reverse causality channel.

Before presenting the IV findings, we would like to emphasize that these results will not alter our main conclusions as market clearing is an identity. In other words, our IV results *cannot* change the cross-sectional and time-series regularity of Firms clearing the market for Index Fund demand (and passive buying in particular). Instead, they can help inform *why* Firms are the most responsive group to Index Fund demand.

### 4.1 Instrumental Variables Approach

As outlined above, our preferred explanation for the relationship between Index Fund and Firm demand is that prices coordinate market clearing. One alternative explanation is the omitted variables problem, i.e., that firm-level or common fundamental shocks coordinate Index Fund and Firm demand. Therefore, we aim to identify shocks to Index Fund demand that are uncorrelated with stock-level, industry-level, and factor-level fundamentals.

To achieve this, we leverage the flow-performance relationship in passive funds, i.e., the fact that passive funds experience inflows after good fund returns and outflows after poor returns.[8] More specifically, we aim to identify stocks that experienced buying or selling because their funds experienced flows based on the performance of "co-holdings." We need, however, to ensure that these flows are not driven by common fundamental shocks affecting both the focal stock and the co-holdings. So, we construct a *leave-out* return which excludes stocks that are likely similar to the focal stock.

To construct these leave-out co-holdings returns, for each focal stock $i$ at the end of each quarter $t$, we identify all funds $k$ in the set of all funds $K$ that hold the stock. Then for each fund $k \in K$, at $t-1$ we drop the stock itself and stocks in the same Daniel et al. (1997) $5 \times 5 \times 5$ portfolio formed on size, book-to-market, and past returns (hereafter DGTW portfolio). We also exclude any stock in the same Fama-French 10 industry. We believe that excluding all stocks in the broader FF 10 industry group -— as opposed to, e.g., the more granular Fama-French 49 industry group – is a conservative choice to exclude as many common fundamental shocks as possible. The goal of this filtering strategy is to identify co-holdings that are not similar on the dimension of size, book-to-market, past return, and industry. Then, within each fund $k$, we reweight all the stocks that survive these filters in proportion to how much the fund held of each stock at $t-1$. Finally, we compute the buy-and-hold returns to this reweighted portfolio from the end of quarter $t-1$ to the end of quarter $t$, which we denote as $r_{i,k,t}^{coholdings}$, and refer to this as the stock-fund-quarter co-holdings return in quarter $t$. We also calculate the co-holdings return for fund $k$ the previous quarter – $r_{i,k,t-1}^{coholdings}$ – following the same methodology, i.e., excluding stocks that were in the same industry and DGTW portfolio as stock $i$ in quarter $t-2$, and computing the hypothetical buy-and-hold returns from the end of quarter $t-2$ to the end of quarter $t-1$. We compute $r_{i,k,t}^{coholdings}$ and $r_{i,k,t-1}^{coholdings}$ even if stock $i$ was not held by fund $k$ at the end of quarter $t-2$ or $t-1$, as these past co-holdings returns could still drive flows into stock $i$ when it was

[8]See Appendix C.3 for our quantitative estimates of the strength of the flow-performance relationship for passive funds.

held in quarter $t$.[9]

After computing $r_{i,k,t}^{coholdings}$ and $r_{i,k,t-1}^{coholdings}$ for each fund $k$ that holds stock $i$, we aggregate these returns to the stock-quarter level. We do this by taking a weighted average of the co-holdings returns for all funds $k \in K$ holding the stock at the end of quarter $t$, where the weights are proportional to stock $i$'s weight in each fund's portfolio.[10] Specifically, we compute:

$$\overline{r}_{i,t}^{coholdings} = \sum_{k \in K} w_{i,k,t} \cdot r_{i,k,t}^{coholdings},$$

where $w_{i,k,t}$ is the weight of stock $i$ in fund $k$ at time $t$, which have been renormalized to sum to one within each stock-quarter. We use the same weighting scheme to compute $\overline{r}_{i,t-1}^{coholdings}$. To reduce the influence of outliers, we Winsorize $\overline{r}_{i,t}^{coholdings}$ and $\overline{r}_{i,t-1}^{coholdings}$ at the 0.5% and 99.5% level.[11] We then use these quantities in our first-stage regression to predict Index Fund demand in stock $i$ in quarter $t$:

$$q_{i,IDX,t} = \gamma_t \cdot \overline{r}_{i,t}^{coholdings} + \gamma_{t-1} \cdot \overline{r}_{i,t-1}^{coholdings} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{Fixed Effects} + e_{i,t}. \tag{7}$$

Given the goal of our IV strategy is to identify passive demand unrelated to own-firm fundamentals, we include the firm's own standardized unexpected earnings (SUE) in quarters $t$ and $t-1$ as controls. To account for common technology shocks, we include Fama-French 10 industry-by-year-quarter fixed effects. We include two lags of co-holdings returns and SUE because the flow-performance relationship is persistent. As an additional check, in Appendix C.2, we show our IV results are robust to excluding the contemporaneous co-holdings returns $\overline{r}_{i,t}^{coholdings}$.

Intuitively, our instrument consisting of co-holdings returns is meant to capture the performance of relatively

[9] Our IV specification has fewer observations than our OLS regression, as some stocks do not have any co-holdings that survive all filters proposed above, and thus have missing values for $\overline{r}_{i,t}^{coholdings}$ and $\overline{r}_{i,t-1}^{coholdings}$.

[10] One natural alternative is to weight by the percent of stock $i$'s market capitalization held by each fund, which should predict the size of the demand shock – as a fraction of stock $i$'s market capitalization – given percentage fund flows (Sammon and Shim, 2023). We prefer weighting by portfolio weights for several reasons. From a practical perspective, we find the strongest first-stage regression when weighting by the stock's weight in each fund (i.e., it makes our weighted average co-holdings returns a relatively better predictor of Index Fund demand). This is likely due to the stronger flow-performance relationship for funds with more concentrated holdings, as shown in Appendix C.3. More broadly, our goal is to identify demand shocks uncorrelated with own-firm fundamentals and similar firms' fundamentals. Funds with more concentrated holdings – and thus larger individual stock weights – likely contain more idiosyncratic shocks from a few stocks unrelated to the focal stock $i$'s returns. A large diversified fund like IWM (Russell 2000) or IWB (Russell 1000), with numerous holdings, is unlikely to have any small individual group of co-holdings driving the fund's returns without a significant systematic component. As additional robustness, we explicitly estimate expected flows given the flow-performance relationship at the fund-level to match the units between expected and realized Index Fund demand in Appendix C.3.

[11] This differs from our treatment of outliers in Section 2.2.3, were we *trimmed* our $q_{i,j,t}$s at the 0.5% and 99.5% levels – removing 7.5% of all observations – instead of Winsorizing. We Winsorize $\overline{r}_{i,t}^{coholdings}$ and $\overline{r}_{i,t-1}^{coholdings}$ because we do not want to drop additional data when constructing the IV. This is in contrast to our data in the market clearing exercise since extreme outliers create the need to clear the market by some other group, i.e., the errors may propagate to the Other group to ensure that the market clears. If instead we use the raw values for $\overline{r}_{i,t}^{coholdings}$ and $\overline{r}_{i,t-1}^{coholdings}$ in our IV specification, the results are even stronger, both in terms of the estimated magnitudes and statistical significance.

unrelated stocks held in the same funds. A high value of $\overline{r}_{i,t}^{coholdings}$ and $\overline{r}_{i,t-1}^{coholdings}$ indicates that a stock's co-holdings performed well; similarly, a low value indicates the co-holdings performed poorly. This variation in co-holdings returns will predict Index Fund buying and selling in the first stage if passive investors chase returns, even if they came from fundamentally unrelated stocks. In addition, the inclusion of $SUE$ as a control and industry-by-year-quarter fixed effects helps remove firm-specific information and common components of returns that may remain in the instrument.

Our second-stage regression is:

$$q_{i,j,t} = \beta \cdot \hat{q}_{i,IDX,t} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{Fixed Effects} + \varepsilon_{i,j,t}, \tag{8}$$

Equation 8 contains the same controls and fixed effects as the first-stage regression in Equation 7. With these controls and fixed effects, the key source of variation we are exploiting is heterogeneity within a Fama-French 10 industry at a given time between stocks whose unrelated co-holdings performed relatively better versus those that performed relatively worse.

A natural question is: If we are interested in measuring the effect of passive demand on prices, why don't we include $r_{i,t}$ on the right-hand side of Equation 8? We believe $r_{i,t}$ is a "bad control" in the sense that our IV approach focuses on inelastic demand shocks, which affect prices and coordinate market clearing. Controlling for prices would be controlling for the mechanism we think facilitates market clearing in the first place.

Table 2 presents the results. In this subsection, we focus on the case where Firm demand, $q_{i,Firm,t}$, is on the left-hand side of Equation 8. The first column shows a strong first stage with an F-statistic over 20. Further, each instrument predicts passive demand with the expected sign, i.e., higher co-holdings returns predict more passive inflows. The second column presents the IV results, where the magnitude is slightly larger and consistent with the OLS results, reported in column 4. The reduced form regression in column 3 shows the regression of $q_{i,Firm,t}$ directly on the instruments themselves. Again, each instrument predicts $q_{i,Firm,t}$ with the expected sign and is (marginally) significant, addressing concerns about weak instruments potentially driving our results (Chernozhukov and Hansen, 2008).[12]

We interpret the results in Table 2 as supporting the conclusion that prices, rather than omitted common shocks or reverse causality, coordinate Firms clearing the market for inelastic Index Fund demand. While previous research has established that Firms respond to prices Baker and Wurgler (2002); Dong et al. (2012);

[12]In Appendix C.1, we present the IV results for all of our individual investor groups. Table 20 confirms that our main finding – i.e., that Firms are the most responsive to Index Fund demand – holds true in our IV setting.

**Table 2:** Instrumental Variables Specification

| | First Stage | IV | RF | OLS |
|---|---|---|---|---|
| $\overline{r}_{i,t-1}^{coholdings}$ | 1.045*** | | -1.433* | |
| | (0.253) | | (0.842) | |
| $\overline{r}_{i,t}^{coholdings}$ | 1.567*** | | -1.004 | |
| | (0.330) | | (0.758) | |
| $SUE_{i,t}$ | 0.00463*** | 0.0105** | 0.006 | 0.00946** |
| | (0.001) | (0.005) | (0.005) | (0.005) |
| $SUE_{i,t-1}$ | 0.00373*** | 0.0145** | 0.0112** | 0.0136** |
| | (0.001) | (0.006) | (0.005) | (0.005) |
| $q_{i,IDX,t}$ | | -0.864** | | -0.627*** |
| | | (0.389) | | (0.037) |
| Observations | 130,794 | 130,794 | 130,794 | 130,794 |
| R-squared | 0.1 | 0.026 | 0.033 | 0.062 |
| F-statistic | 23.26 | | | |
| Fixed Effects | FF 10 Industries × Year-Quarter | | | |

*Notes.* The table provides estimates from the first and second stages of our IV regression, as well as the associated reduced form and OLS regressions:

$$q_{i,IDX,t} = \gamma_t \cdot \overline{r}_{i,t}^{coholdings} + \gamma_{t-1} \cdot \overline{r}_{i,t-1}^{coholdings} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + e_{i,t},$$
$$q_{i,Firm,t} = \beta \cdot \hat{q}_{i,IDX,t} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + \varepsilon_{i,Firm,t},$$

where $\overline{r}_{i,t}^{coholdings}$ and $\overline{r}_{i,t-1}^{coholdings}$ are the weighted average leave-out co-holdings returns across all funds $k \in K$ that held stock $i$ at the end of quarter $t$. $SUE_{i,t}$ is the standardized unexpected earnings (SUE) of stock $i$ in quarter $t$. FE are a set of Fama-French 10 industry-by-year-quarter fixed effects. Standard errors are double clustered at the stock and time (year-quarter) level. The first column reports the first stage regression, and the last row reports the associated F-statistic. The second column reports the IV specification, while the third column reports the reduced-form regression. The final column reports the OLS results, restricting to the sample with non-missing values for the instruments and a matched set of controls and fixed effects to the first three columns.

Ma (2019), none of these results on their own imply that Firms would be the *most* responsive group.

The natural next question is that if prices are the coordinating driver of Firm and Index Fund demand, how does that change the interpretation of our OLS results? Recall that the IV is designed to isolate non-fundamental Index Fund demand shocks, i.e, inelastic demand. Given the similarity in the magnitudes of the estimated effect in the IV and OLS regressions, this suggests that Firms respond just as much to non-fundamental Index Fund demand shocks as to the endogenous demand shocks in the OLS regression. This suggests that the Index Fund demand *itself* is what's important for how Firms respond. This is outside of a traditional Q-theoretic framework where firms only raise capital when they have attractive investment opportunities.

**Alternative Instrument Specification** Implicitly, our first stage regression in Equation 7 assumes that returns of co-holdings predict flows and does so equally across funds. This is despite the fact that, empirically, there is significant heterogeneity in the flow-performance relationship for index funds. Even with this simplification, the strength of our first-stage regression suggests that this is true on average. In Appendix C.3, we refine this result, estimating the flow-performance sensitivity at the fund level, then use each fund's sensitivity along with the previously calculated stock-quarter-level leave-out co-holdings returns to compute expected flows into each fund each quarter.[13] We then use these expected flows to calculate expected demand shocks at the stock-fund-quarter level (because index funds proportionally scale all holdings up and down in response to flows), and then directly aggregate across all funds that hold each stock. We call this alternative measure of expected Index Fund buying $\tilde{q}_{i,IDX,t}$, and is akin to the expected Index Fund demand that emerges from a first-stage regression and is used in a second-stage regression. Although this approach potentially introduces noise by using the *estimated* flow-performance relationships, it can help reconcile differences in magnitudes between the IV and OLS regressions. This is because – unlike the returns in the reduced form regression in Table 2 – the estimated demand shock will be in the same units as $q_{i,IDX,t}$.

We report the results in Appendix Table 22. The regression of $q_{i,Firm,t}$ on $\tilde{q}_{i,IDX,t}$, i.e., our alternative measure of expected index fund buying yields a coefficient of -0.88. This is reassuring on two levels. First, it again confirms the *causal* nature of the relationship between Index Fund demand and Firm issuance. Second, because $\tilde{q}_{i,IDX,t}$ is in the same units as $q_{i,IDX,t}$, it suggests that the estimated magnitudes in Table 2 are not distorted by the instrument being in units of returns. In addition, it directly aggregates quantities to "build" a first stage instead of using returns as an instrument in the first stage. This helps address the

[13] While we compute expected flows at the fund-quarter level, we would like to highlight that this quantity varies across stocks in the same fund at a given point in time. This is because we use the stock-fund-quarter-level leave-out co-holdings returns to compute expected flows, and these co-holdings returns could be different for every stock.

concern that aggregating co-holdings returns, even from unrelated stocks, can recover common factors that could capture the focal stock itself.

As an additional robustness check against a possible look-ahead bias and the endogeneity of time $t$ returns, we construct an alternative version of $\tilde{q}_{i,IDX,t}$ that uses only the lagged flow-performance relationship and lagged co-holdings returns. These findings, also reported in Appendix Table 22, confirm that our main results still hold.

## 4.2 Prices as a Coordinating Device

In the last subsection, we constructed an IV designed to rule out the effect of common fundamental shocks and therefore "rule-in" the role of prices. In this subsection, our objective is to provide direct evidence for the channel of prices coordinating market clearing. To this end, we are interested in quantifying how the responsiveness of each group is affected by how "expensive" the stock is. As a simple measure to capture this, we use excess earnings yield ($EXEY$), defined as:

$$EXEY_{i,t} = \frac{TTM\ EPS_{i,t}}{P_{i,t}} - r_{rf,t},$$

where $TTM\ EPS_{i,t}$ is the trailing 12 months earnings per share (EPS), $P_{i,t}$ is the price at the end of the quarter, and $r_{rf,t}$ is the 10-year Treasury rate. To strip out the effect of extraordinary items, we use "street" earnings from IBES as our measure of EPS (Hillenbrand and McCarthy, 2024). The logic behind using $EXEY$ is that if a stock has an earnings yield higher than the risk-free rate, the stock is relatively cheap, i.e., has a relatively low P/E ratio. Ben-David and Chinco (2024) provide evidence that high values of $EXEY$ are a strong predictor of buybacks, and low values of $EXEY$ predict issuance, validating its use as a measure of a stock's relative valuation relevant for firm decision making.

To simplify the interpretation, we construct an indicator variable for whether the excess earnings yield is positive $\mathbb{1}_{EXEY_{i,t}>0}$. We then augment our baseline OLS regression with $\mathbb{1}_{EXEY_{i,t}>0}$ and an interaction term between $\mathbb{1}_{EXEY_{i,t}>0}$ and Index Fund demand:

$$q_{i,j,t} = \beta_j \cdot q_{i,IDX,t} + \gamma_j \cdot \mathbb{1}_{EXEY_{i,t}>0} + \psi_j \cdot q_{i,IDX,t} \times \mathbb{1}_{EXEY_{i,t}>0} + \rho_t + e_{i,j,t}, \tag{9}$$

where $\rho_t$ are a set of year-quarter fixed effects to account for the common time-series component in $EXEY$. The key term of interest here is $\psi_j$, which indicates how our groups $j$ adjust their behavior in response to the stock being relatively cheap.

Table 3 contains the results. Our primary focus is on column 9, which contains the results for Firms. There, we see that, as in our baseline results, $\beta_j$ is negative. The coefficient on $\gamma_j$ is positive, consistent with the notion that when the stock is cheap, the Firm will unconditionally want to do more buybacks/less issuance (Ben-David and Chinco, 2024). Finally, the interaction term $\psi_j$ is negative, suggesting that when the stock is cheap, the firm trades less against passive demand. In terms of magnitudes, suppose we call stocks with positive excess earnings yield "value stocks," as their earnings are high relative to the market value of the firm. Value stocks typically do not issue equity because their equity valuation is relatively lower, and therefore it is usually cheaper – from a cost of capital perspective – to raise financing in debt markets. The level effects are on the same scale as the interaction term, so combining the two effects implies that additional Index Fund demand is half of the effect of switching from being a value stock to a growth stock. Our interpretation is that Firms do not want to accommodate Index Fund demand by issuing equity when prices are relatively low.

Next, we focus on Short Sellers in column 8. As in the baseline OLS results, $\beta_j$ is negative. However, the coefficient on $\gamma_j$ is positive. This suggests that Short Sellers are less likely to accommodate Index Fund buying when stocks are cheap. Similar to the Firms results in column 9, this is in line with our preferred interpretation: Short Sellers don't want to short when prices are already low.

Finally, if when prices are low, Short Sellers and Firms play a smaller role, a natural question is: which of our groups takes on a larger role to clear the market for Index Fund demand? As column 5 shows, Financial Institutions (e.g., Hedge Funds, Broker Dealers) end up playing a large part in clearing the market for marginal passive demand when prices are low. While the $\beta_j$ (i.e., unconditional responsiveness) is close to zero, the $\gamma_j$ is -0.512, suggesting that when prices are low, Financial Institutions are more likely to take the other side of Passive Demand. It's important to note that in this regression, we are not conditioning on the observed change in prices. So, it's possible that when prices are low, prices need to move more than usual to get Financial Institutions to step up and clear the market for Index Fund demand.

### 4.3 Ruling out Mechanical Stories and Reverse Causality

As discussed above, our IV results suggest that an omitted variables problem is not driving our baseline OLS results. However, one might still be worried that the mechanical response of Index Funds to Firm activity creates a reverse-causality problem and this reverse causality is the main driver of our findings.

**Table 3:** Earnings Yield and Market Clearing

| | Act. Fnds.<br>(1) | Oth. Fnds.<br>(2) | Pens.<br>(3) | Ins.<br>(4) | Fin Inst.<br>(5) | Insiders<br>(6) | Others<br>(7) | Shorts<br>(8) | Firms<br>(9) |
|---|---|---|---|---|---|---|---|---|---|
| $q_{i,IDX,t}$ | 0.238*** | 0.0938*** | 0.0243*** | 0.0628*** | 0.0349 | -0.0759*** | -0.249*** | -0.269*** | -0.859*** |
| | (0.029) | (0.017) | (0.003) | (0.008) | (0.068) | (0.009) | (0.070) | (0.039) | (0.066) |
| $1_{EXEY_{i,t}>0}$ | -0.112*** | -0.0167* | -0.0227*** | -0.0247** | -0.290*** | -0.00582 | -0.653*** | 0.0374 | 1.088*** |
| | (0.023) | (0.009) | (0.005) | (0.010) | (0.062) | (0.009) | (0.081) | (0.035) | (0.082) |
| $q_{i,IDX,t} \times 1_{EXEY_{i,t}>0}$ | -0.106*** | -0.0081 | -0.00755*** | 0.00298 | -0.512*** | 0.0137 | 0.0566 | 0.120*** | 0.441*** |
| | (0.035) | (0.032) | (0.003) | (0.012) | (0.056) | (0.009) | (0.066) | (0.028) | (0.062) |
| Observations | 146,874 | 146,874 | 146,874 | 146,874 | 146,874 | 146,874 | 146,874 | 146,874 | 146,874 |
| R-squared | 0.007 | 0.012 | 0.005 | 0.007 | 0.01 | 0.004 | 0.008 | 0.011 | 0.073 |
| Fixed-Effects | | | | | YQ | | | | |

*Notes.* The table provides estimates from our baseline OLS regression augmented to account for the effect of relative valuation on each groups' response to Index Fund demand:

$$q_{i,j,t} = \beta_j q_{i,IDX,t} + \gamma_j 1_{EXEY_{i,t}>0} + \psi_j q_{i,IDX,t} \times 1_{EXEY_{i,t}>0} + \rho_t + e_{i,j,t},$$

where $EXEY_{i,t}$ is firm $i$'s excess earnings yield in quarter $t$, defined as the difference between the firm's trailing 12-month earnings yield and the risk-free rate, as measured by the yield on 10-year treasuries. $\rho_t$ are a set of year-quarter fixed effects. Standard errors are double clustered at the stock and time (year-quarter) level.

To clarify this point, consider the following numerical example. Suppose that there is only one passive fund, which is a value-weighted index fund that holds the entire market. For simplicity, assume that this fund applies no float adjustments and has enough AUM such that it holds 10% of each firm's shares outstanding. Now, suppose one firm conducts an issuance equal to 5% of its shares outstanding, while all the other firms do neither net issuance nor net buybacks. This will trigger trading by the value-weighted index fund, as it must maintain a constant ownership percentage of the (float-adjusted) shares outstanding of each constituent (Sammon and Shim, 2023). To do this, the fund must buy shares of the firm that issued, which it will fund by selling shares of everything else.

In this case, the fund will need to buy roughly 10% of the issued shares, i.e., 10% $\times$ 5% = 50 bps of shares outstanding.[14] Then, the "beta" in this example from our baseline regression of Firm demand on Index Fund demand would be -5% = $\beta\times$ 0.5%, i.e., $\beta = -10$.

The exact problem that might cause concern is as follows: Suppose most of the time, Firms don't respond to Index Fund demand. But, sometimes Firms act in the primary market and Index Funds mechanically respond. Then we would be mixing "betas" of -10 with "betas" of 0 and we end up recovering a number like -0.6, i.e., our baseline OLS estimate.

Against this mechanism, we have already shown our IV results, which partially rule out the mechanical story.

[14]To see this, suppose the firm initially had 100 shares outstanding, so after the issuance, if the value-weighted index fund didn't trade, it would own 10/105 = 9.5% of shares outstanding. By buying 10% of the issuance, the fund would own 10.5/105 =10% of the firm's shares outstanding. We say "roughly" here because this is actually a fixed-point problem for the fund. Specifically, the fund will need to determine the right amount of all the other stocks to sell such that it will hold a constant percentage of each firm's share outstanding while respecting the constraint of fully investing its AUM – and as a result will buy less than 10% of the issuance, as after selling the other stocks to "make room" for the issuance, it will own less than 10% of their shares outstanding.

This is because our IV is designed to isolate the part of Index Fund demand coming from *flows*, as opposed to mechanical rebalancing in the face of Firm issuance or buybacks. Our augmented version in Appendix C.3 goes even further, estimating the flow performance relationship for each individual fund. Our IV, however, does not fully rule out the mechanical story, as it is still possible that, for reasons *unrelated* to the associated Index Fund demand shock, Firms happen to issue equity when co-holdings returns are high.

In this subsection, we discuss several additional pieces of evidence for why we believe the mechanical/reverse causality story isn't driving our results. First, as highlighted in the example above, the mechanical effect is proportional to -1 times the inverse of lagged passive ownership. As passive ownership has increased, therefore, the part of the "beta" coming from the mechanical response should have shrunk in absolute value. So, if the mechanical trading of Index Funds was entirely driving our findings, our $\beta_{Firm}$ should be approaching 0 over time. Figure 10 shows the opposite – $\beta_{Firm}$ has become more negative over time, inconsistent with the mechanical channel as the primary driver of our results.

Another important aspect of the mechanical story is that it should be symmetric with respect to issuance and buybacks. Going back to our numerical example, if the Firm bought back 5% of shares, the value-weighted index fund would be forced to sell roughly 50 basis points of the firm's shares outstanding. However, Figure 4 shows that our effect is asymmetric, with Firms only responding to Index Fund demand when Index Funds are net buyers. Again, this is inconsistent with the mechanical effect entirely driving our results.

Finally, we perform a decomposition exercise aimed at explicitly separating out the part of Index Fund demand coming from flows and the part of Index Fund demand coming from the mechanical response of Index Funds to Firm activity. In Appendix C.4, we show that – although the baseline OLS effect is attenuated when using our flows-only-based measure of expected Index Fund demand – our main finding of Firms responding to Index Fund demand is qualitatively unchanged. In summary, while the mechanical effect is certainly present in the data, between our IV approach, the time-series trends in the beta coefficients, the asymmetry with respect to Index Fund buying and selling, and our decomposition exercise, we have significant evidence against this channel being the primary driver of our findings.

# 5 How Do Firms Provide Shares?

We further explore how exactly Firms issue shares and/or reduce the size of buybacks to accommodate greater passive demand.

We decompose Firm changes into three sources: Buybacks, SEOs, and other sources. We identify the quarterly total dollar value of buybacks using PRSTKCY in Compustat – which we convert into share terms by assuming that Firms buy back at the volume-weighted average (split-adjusted) price over the quarter. We hand collect data on SEOs from Bloomberg. All other sources of shares outstanding changes are accounted for by the remainder. This gives us three measures of Firm activity to replace our single measure that we have used throughout the paper. Two measures capture direct activity from Firms (buybacks and SEOs). We think of the remainder as mainly capturing compensation (i.e., restricted or performance stock units or exercised stock options), but it may also include other ways in which shares can be created (e.g., the exercise of warrants or convertible debt).[15] As in Section 2.2.2, we construct the $q_{i,j,t}$ variables so that a positive value represents Firms buying shares and a negative value corresponds to selling shares (i.e., issuance). Also as before, the $q_{i,j,t}$ variables are represented as a percentage of all shares outstanding.

Table 4 presents the regression estimates. It shows that both SEOs and compensation/other sources account for the vast majority of Firm responsiveness to Index Fund demand. The largest-in-magnitude coefficient comes through our "remainder" group – which includes employee compensation – with an estimate of -0.312; SEOs are the second largest at -0.254. Although Firms are less responsive in terms of adjusting buybacks in the face of Index Fund demand, the coefficient is still negative at -0.077.[16] In addition, given the role that compensation is likely to play, a natural question is whether our findings are dominated by a few industries (e.g., the tech sector) that tend to use stock as a source of compensation. In Appendix B.7.1, we show that, while there is variation in Firms' betas by industry, all but one industry has a negative beta.

Another pattern we have documented is that Firms are particularly responsive to buying by Index Funds. We show in Section 3.1.2 and in Appendix B.2 that the slope between Firm and Index Fund demand is much steeper in firm-quarters with net Index Fund buying and estimate that it is nearly -1. This is unlikely to be a coincidence. Shares issued for the purpose of employee compensation are likely to have an option-like structure – in the sense that shares awarded to employees are likely to increase when prices are high. And, according to the evidence in Section 4, Firms clear the market for Index Fund demand using prices as a coordinating device.

This suggests that Firm issuance through compensation and other sources should be much more sensitive to passive buying than selling. Similar to the tests in Appendix B.2, we split the sample based on whether

[15] The CRSP data on shares outstanding does not include treasury shares. This means that if shares are held in the treasury stock but are slated to be awarded to employees, they will only be counted into shares outstanding once they are actually awarded.

[16] The smaller coefficient could also be because buybacks tend to be conducted by a minority of Firms and the average size (as a fraction of total shares outstanding) tends to be smaller. See Appendix D for stylized facts on buybacks.

**Table 4:** Beta Estimates: Decomposing Firm Changes

| | No FEs | | Ind. x YQ FE | |
|---|---|---|---|---|
| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\beta_j$ | $t(\beta_j)$ |
| Active Funds | 0.193 | 7.015 | 0.179 | 11.453 |
| Other Funds | 0.086 | 3.148 | 0.062 | 8.716 |
| Pension Funds | 0.022 | 7.423 | 0.025 | 10.158 |
| Insurance | 0.065 | 7.894 | 0.063 | 13.299 |
| Financial Institutions | -0.152 | -2.265 | -0.085 | -1.794 |
| Insiders | -0.084 | -10.920 | -0.086 | -12.141 |
| Other | -0.250 | -4.674 | -0.282 | -6.166 |
| Short Sellers | -0.238 | -6.433 | -0.283 | -10.337 |
| Firms (Buybacks) | -0.077 | -8.822 | -0.060 | -8.609 |
| Firms (SEOs) | -0.254 | -9.584 | -0.242 | -9.476 |
| Firms (Compensation, Other) | -0.312 | -13.039 | -0.292 | -13.930 |
| Total | -1.001 | | -1.001 | |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. We decompose Firms' changes into two new groups: Firms' primary market activity (seasoned equity offerings, or SEOs, and buybacks) and Firms' other issuance activity. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

Index Funds were net buyers or sellers in a stock-quarter. Table 5 provides the regression estimates. We find that half of the responsiveness of Firms to positive passive demand comes through compensation and other sources. Less than half comes from the combination of SEOs and buybacks, though buybacks is a relatively small component. Both compensation/other sources and SEO betas are larger in magnitude than any other group for the passive buying sample. For negative demand, we find that all three Firm betas are positive, and the largest of the three is compensation. First, this suggests that even when Index Funds are selling, Firms are issuing some shares for compensation, though the magnitude is much smaller than when Index Funds are net buyers.

# 6 Conclusion

We aim to answer a basic question: Who sells when passive investors buy, and who buys when passive investors sell? That is, when passive investors trade, who ultimately clears the market? To this end, we start by combining several datasets on investors' holdings with data on short-interest, insider transactions, and shares outstanding. We aggregate all holdings and/or changes by group to study which groups take the other side of passive demand, both positive and negative, at the stock and quarter level to clear the market.

**Table 5:** Regression Estimates: Decomposing Firms by Index Fund Demand Direction

| | Negative Index Fund Demand | | | | | Positive Index Fund Demand | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | $\hat{q}_j$ | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | $\hat{q}_j$ |
| Active Funds | 0.179 | 5.635 | -0.044 | -1.792 | -0.087 | 0.207 | 5.623 | -0.066 | -3.036 | -0.004 |
| Other Funds | 0.026 | 2.457 | 0.041 | 2.787 | 0.035 | 0.109 | 2.982 | 0.041 | 3.162 | 0.073 |
| Pension Funds | 0.008 | 1.181 | -0.013 | -1.866 | -0.015 | 0.027 | 6.742 | -0.011 | -1.741 | -0.003 |
| Insurance | 0.047 | 5.309 | -0.034 | -2.476 | -0.045 | 0.073 | 6.564 | -0.034 | -2.675 | -0.012 |
| Financial Institutions | -0.319 | -3.801 | 0.143 | 1.677 | 0.220 | -0.033 | -0.376 | 0.011 | 0.124 | 0.001 |
| Insiders | -0.024 | -2.766 | -0.008 | -0.752 | -0.002 | -0.104 | -8.925 | -0.019 | -2.661 | -0.050 |
| Other | -0.960 | -9.621 | -0.049 | -0.608 | 0.182 | -0.007 | -0.097 | 0.051 | 0.627 | 0.049 |
| Short Sellers | -0.137 | -2.958 | 0.008 | 0.184 | 0.041 | -0.321 | -7.045 | 0.118 | 2.594 | 0.022 |
| Firms (Buybacks) | 0.058 | 3.684 | 0.611 | 18.416 | 0.597 | -0.065 | -6.799 | 0.444 | 22.826 | 0.425 |
| Firms (SEOs) | 0.011 | 0.981 | -0.043 | -4.017 | -0.046 | -0.401 | -9.900 | 0.065 | 4.944 | -0.054 |
| Firms (Compensation, Other) | 0.110 | 5.106 | -0.611 | -24.413 | -0.638 | -0.484 | -13.193 | -0.601 | -25.500 | -0.745 |
| Total | -1.001 | | 0.001 | | 0.242 | -0.999 | | -0.001 | | -0.299 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. We decompose Firms' changes into two new groups: Firms' primary market activity (seasoned equity offerings, or SEOs, and buybacks) and Firms' other issuance activity. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

We then develop a regression framework to assess (1) which groups are on average taking the opposite side of passive investors, and (2) which groups are the most responsive to marginal passive demand. Treating passive investors as the focal group is built on the logic that passive demand is inelastic, and therefore one can view passive index funds and ETFs as initiating trades when adjusting their holdings. That being said, our methodology could be applied to any group of investors, e.g., it could be used to study which groups are on average trading against active mutual funds.

Our main finding is that Firms have been the single most significant provider of the shares purchased by Index Funds. We estimate that when passive investors demand 1pp more of Firms' shares outstanding, Firms respond at rate of 0.64 percentage points more shares issued/fewer shares repurchased. Short sellers are another important group for market clearing, with a response coefficient of -0.24. These two groups alone account for 88% of the marginal shares needed to clear the market in the face of additional passive demand. Firms do not play a role in clearing the market when Index Funds are net sellers of a stock, which is where Financial Institutions play a bigger role in clearing the market for passive demand. Long-term investors (e.g., Active Funds, Insurance, and Pension Funds) typically mimic the direction of passive demand: they buy more/sell less in the stocks that passive investors buy more of.

Our proposed mechanism is that prices coordinate market clearing. To test this, we construct an instrument of inelastic Index Fund demand designed to be orthogonal to own-firm fundamentals. Specifically, we identify stocks which are expected to receive passive flows because the unrelated stocks they were co-held with performed relatively well. We obtain strikingly similar results in this better identified setting, evidence that Index Fund demand *causes* Firms to issue equity. Further, we provide direct evidence that Firms are less likely to issue in the face of Index Fund demand when prices are relatively low – further advancing the narrative of prices being the crucial factor that determines how Firms respond to inelastic demand.

We believe that our results speak to several fields in finance. First, and perhaps most clearly, our findings are related to work that studies the rise in passive ownership in financial markets. Several other papers have studied the effects of passive demand on asset prices, and how the effects depend on who is on the other side of the trade. However, existing analysis neglects a crucially important player: the firm itself.

Our findings also have important implications for corporate finance in that it suggests that passive ownership may impact firms, including capital structure and payout policy, as well as real effects like investment. As discussed in Morck et al. (1990), there are many reasons why the stock market may matter for the real economy, and is not just a sideshow. We provide evidence on a particular channel, specifically that passive ownership may affect firms' financing decisions. Our findings suggest that firms respond to increased inelastic demand by issuing more equity. And, given our evidence that prices coordinate market clearing, this suggests that Firms are issuing at relatively high prices, taking advantage of a perceived lower cost of capital. We document that this issuance comes through two channels. The first is primary-market offerings (SEOs), where the firm itself receives cash. The second is employee compensation and other sources of issuance, where the firm may be using shares as a seemingly less expensive way to compensate employees. What firms do with this additional cash is a promising area for future research.

## References


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

Baker, M. and Wurgler, J. (2002). Market timing and capital structure. *The journal of finance*, 57(1):1–32.

Balasubramaniam, V., Campbell, J. Y., Ramadorai, T., and Ranish, B. (2023). Who owns what? a factor model for direct stockholding. *The Journal of Finance*, 78(3):1545–1591.

Barber, B. M., Huang, X., Jorion, P., Odean, T., and Schwarz, C. (2022). A (sub) penny for your thoughts: Tracking retail investor activity in taq. *Available at SSRN 4202874*.

Barber, B. M., Huang, X., and Odean, T. (2016). Which factors matter to investors? evidence from mutual fund flows. *The Review of Financial Studies*, 29(10):2600–2642.

Barber, B. M. and Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, 55(2):773–806.

Battalio, R. H., Jennings, R. H., Saglam, M., and Wu, J. (2023). Identifying market maker trades as' retail'from taq: No shortage of false negatives and false positives. *Available at SSRN 4579159*.

Beber, A., Brandt, M. W., Cen, J., and Kavajecz, K. A. (2021). Mutual fund performance: Using bespoke benchmarks to disentangle mandates, constraints and skill. *Journal of Empirical Finance*, 60:74–93.

Ben-David, I. and Chinco, A. M. (2024). Expected eps × trailing p/e. Technical report.

Boehmer, E., Jones, C. M., Zhang, X., and Zhang, X. (2021). Tracking retail investor activity. *The Journal of Finance*, 76(5):2249–2305.

Brav, A., Lund, D. S., and Zhao, L. (2024). Flows, financing decisions, and institutional ownership of the us equity market. *Available at SSRN 4693837*.

Bushee, B. J. (2001). Do institutional investors prefer near-term earnings over long-run value? *Contemporary accounting research*, 18(2):207–246.

Chang, Y.-C., Hong, H., and Liskovich, I. (2015). Regression discontinuity and the price effects of stock market indexing. *The Review of Financial Studies*, 28(1):212–246.

Chernozhukov, V. and Hansen, C. (2008). The reduced form: A simple approach to inference with weak instruments. *Economics Letters*, 100(1):68–71.

Chinco, A. and Fos, V. (2021). The sound of many funds rebalancing. *The Review of Asset Pricing Studies*, 11(3):502–551.

Chinco, A. and Sammon, M. (2023). The passive-ownership share is double what you think it is. *Available at SSRN 4188052*.

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

Crane, A. D. and Crotty, K. (2018). Passive versus active fund performance: do index funds have skill? *Journal of Financial and Quantitative Analysis*, 53(1):33–64.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance*, 52(3):1035–1058.

Dong, M., Hirshleifer, D., and Teoh, S. H. (2012). Overvalued equity and financing decisions. *The Review of Financial Studies*, 25(12):3645–3683.

Eisfeldt, A. L., Schubert, G., and Zhang, M. B. (2023). Generative ai and firm values. Technical report, National Bureau of Economic Research.

Evans, K. P., Leung, W. S., Li, J., and Mazouz, K. (2023). Etf ownership and seasoned equity offerings. *Journal of Financial and Quantitative Analysis*, pages 1–28.

Fama, E. F. and French, K. R. (2005). Financing decisions: who issues stock? *Journal of financial economics*, 76(3):549–582.

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

Greenwood, R., Hanson, S., and Stein, J. C. (2010). A gap-filling theory of corporate debt maturity choice. *The Journal of Finance*, 65(3):993–1028.

Greenwood, R. and Sammon, M. C. (2022). The disappearing index effect. Technical report, National Bureau of Economic Research.

Haddad, V., Huebner, P., and Loualiche, E. (2022). How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. In *Working Paper*.

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

Harris, L. and Gurel, E. (1986). Price and volume effects associated with changes in the s&p 500 list: New evidence for the existence of price pressures. *the Journal of Finance*, 41(4):815–829.

Hillenbrand, S. and McCarthy, O. (2024). Valuing stockswith earnings.

Investment Company Institute (2023). Fact book. Technical report. [link].

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

Laarits, T. and Sammon, M. (2023). The retail habitat. *Available at SSRN 4262861*.

Ma, Y. (2019). Nonfinancial firms as cross-market arbitrageurs. *The Journal of Finance*, 74(6):3041–3087.

Madhavan, A. (2003). The russell reconstitution effect. *Financial Analysts Journal*, 59(4):51–64.

McLean, R. D., Pontiff, J., and Reilly, C. (2020). Taking sides on return predictability. *Available at SSRN 3637649*.

Morck, R., Shleifer, A., Vishny, R. W., Shapiro, M., and Poterba, J. M. (1990). The stock market and investment: is the market a sideshow? *Brookings papers on economic Activity*, 1990(2):157–215.

Petajisto, A. (2011). The index premium and its hidden cost for index funds. *Journal of empirical Finance*, 18(2):271–288.

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

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

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

Tamburelli, T. (2024). Firms issue shares to satisfy inelastic demand. *Available at SSRN*.

Van der Beck, P. (2021). Flow-driven esg returns. *Swiss Finance Institute Research Paper*, (21-71).

# Who Clears the Market when Passive Investors Trade?

Marco Sammon and John J. Shim

# INTERNET APPENDIX

## A Data and Methodology Appendix

### A.1 Data Sources and Cleaning Details

In this appendix, we provide greater detail on the the data sources and cleaning decisions, which were outlined in Section 2.1.

#### A.1.1 Thomson Reuters S12 13F Holdings Data

We use the Thomson Reuters S12 data for mutual holdings data for all funds registered under the Investment Company Act of 1940 (commonly referred to as 40 Act Funds). These are mostly mutual funds, exchange-traded funds (ETFs), closed-end funds, and unit-investment trusts.

We separate all funds into three categories: Index (passive), Active, and Other. We classify a fund as an Index Fund based on the index fund flag and the fund name in the CRSP mutual fund database using the method in Appel et al. (2016).[17] We classify a fund as Active if it is in the universe of funds that can be linked between the CRSP mutual fund dataset and the Thomson S12 dataset using the WRDS MF links database but it is not otherwise classified as passive. Any remaining funds that cannot be matched between Thompson and the CRSP mutual fund database are included in a separate "Other Funds" group. It is worth highlighting that an update to the S12 data, which took place in February 2022 and retroactively

[17] Specifically, we classify a fund as passive if it meets either of the following criteria: (1) It has a non-missing value for the index fund flag in the CRSP mutual fund database. This includes funds with code "D" (pure index funds), code "B" (index-based funds) and code "E" (enhanced index funds) (2) It has a name that makes it look like an index fund. To identify these funds, we use the same list of strings as Appel et al. (2016), which includes permutations of index names like "S&P" and "500". Although this is a less conservative definition of passive funds than used in other papers (e.g., Crane and Crotty (2018), which only includes funds with an index fund flag of "D"), including these additional funds has little effect on the level of passive ownership. For example, in 12/2022, the level of passive ownership under the Crane and Crotty (2018) definition is 16.7%, while under our definition it is 17.1%.

updated previous S12 data, dramatically increased the size of the Other Funds group from 2017-present, mainly coming from increased coverage of foreign funds which hold US equities.

As discussed in Sammon and Shim (2023), the prevalence of stale filings can create problems when working with changes in holdings. To address this issue, we linearly interpolate holdings of each stock at the fund level across stale quarters, for up to three consecutive stale quarters.

#### A.1.2 Thomson Reuters S34 Mutual Fund Holdings Data

We obtain data on institutional investors' holdings from 13F filings recorded in the Thomson S34 dataset. Institutions are required to file a 13F if they hold more than $100M in qualified securities. To classify institutional investors into groups, we use the 13F classification procedure in Bushee (2001), and data for the classification from Brian Bushee's website. The classification assigns each institution to one of the following categories: banks, investment companies, independent investment advisors, insurance companies, corporate pension funds, public pension funds, university and foundation endowments, and miscellaneous. When forming our groups, we combine insurance companies and university & foundation endowments into our "Insurance" group because they are both very long-horizon investors (and the endowments group is relatively small). We also combine corporate and public pension plans into our "Pension Funds" group because they have common objectives.

#### A.1.3 CRSP

We use the CRSP monthly stock database for data on shares outstanding. While shares outstanding is not traditionally an investor category, to do a complete accounting for how shares change hands, we also include the Firm itself (i.e., it can issue or buy back shares). We identify share issuance/buybacks based on changes in split-adjusted shares outstanding. Importantly, CRSP shares outstanding does not include treasury or authorized shares.[18] This means that stock awards which have been approved by the board of directors and even shares have been authorized/issued but held in the treasury will not be counted. Such shares will only be counted toward CRSP's definition of shares outstanding when they are actually awarded to employees (and, thus, theoretically available in the market if those employees choose to sell shares).

[18]See the CRSP US Stock Data Description Guide (page 125) for more details.

### A.1.4 Computstat/Markit

Another potentially important source of shares is Short Sellers. Specifically, each shorted share effectively creates an additional share that needs to be held by another investor. In addition, when Short Sellers close their positions, the effective supply of shares decreases. Therefore, to get a complete accounting of how shares can change hands, we also examine changes in short interest. Short interest data are obtained from Compustat following the method in Hanson and Sunderam (2014). The short interest ratio computed using Compustat data is highly correlated with the level of short interest reported by S&P Global's Markit database. We also use Markit to obtain data on the shares available for shorting, utilization rates, and shorting costs.

### A.1.5 Insiders

For insiders, we do not have the level of holdings but only changes in holdings via their publicly reported buying and selling activity. We get data on insider transactions from the Thomson Reuters Insiders dataset, which we aggregate at the firm and quarter level.

### A.1.6 Index Constituents Data

The next collection of datasets we leverage include information on index membership, index weights and float adjustments. We obtain S&P 500 and S&P 1500 membership data directly from S&P. Starting in 2002, this includes float adjustments for all stocks in the 1500 universe. We get data on the S&P MidCap 400 and S&P SmallCap 600 membership from Siblis Research, and match this to the S&P 1500 data to obtain float adjustments, as the same float adjustment is applied to all sub-indices within the S&P 1500 index family. We get Russell index membership data from FTSE Russell. Starting in 2008, this includes daily index membership, as well as daily float adjustments and index weights. We get Nasdaq 100 index membership from Siblis research, which starts in 2014.

We get CRSP index membership directly from CRSP. This includes daily index membership, weights and float adjustments starting in 2014 for all CRSP sub-indices. In our analysis, we pool together all CRSP index-based funds which do not track the CRSP total market index, as the AUM tracking these is small relative to the AUM in the three funds tracking the CRSP total market index (VTSAX, VTI and VITNX[19]).

[19]Note that VITNX does not exactly track the CRSP total market index. According to Vanguard's website, "The fund replicates more than 95% of the market capitalization of the index and invests in a representative sample of the balance using

We identify migrations within families of indices by identifying stocks which were simultaneously added to one index in the family and dropped from another (e.g., a stock which is dropped from the Russell 1000 and added to the Russell 2000 at the same time is classified as a Russell migration).

In addition to classifying funds based on whether they are active or passive, we also aim to identify the index each passive fund is tracking. To identify funds tracking CRSP indices, we obtain a list of fund tickers from Vanguard's website. To identify funds tracking Russell, S&P and Nasdaq indices, we use the funds' names in the CRSP mutual fund database. For example, to identify S&P 500-tracking funds, we look for combinations of "S&P", "S & P", "SandP", "S and P", "SP" (all non-case sensitive) and "500". For the S&P 500, we validate our name-based classification by comparing it a classification based on the Lipper Objective Code "SP" (i.e., the Lipper Objective Code for S&P 500 funds). We find these two methods yield similar results, allaying concerns about misclassification using our names-based methodology. We also hand check the largest funds tracking each index to ensure they are classified correctly.

We then compute the ratio of the AUM of all funds tracking each index to the total index float (i.e., the sum of the float adjusted market capitalization of all index members). Then, at the stock level, we compute the expected number of shares held by each family of index funds as shares outstanding $\times$ AUM Tracking/Index Capitalization $\times$ IWF (where IWF is the investable weight factor, expressed as a decimal). The logic is that an index tracker – by construction – holds a constant percentage of each constituent's float (Sammon and Shim, 2023). As a specific example, suppose that S&P 500 tracking funds own 10% of the index's float. And then, consider an individual stock with a float adjustment of 0.8. S&P 500 index funds are expected to own 10% of the stock's float i.e., 10% $\times$ 0.8 = 8% of the firm's shares outstanding.

One concern with this method for computing expected ownership by each family of index funds is that it likely understates the true size of index trackers, as there are many investors tracking these benchmarks – e.g., direct replication by institutional investors and shadow indexing by active investors – which will not be captured by index fund holdings alone (Chinco and Sammon, 2023). Given that our market clearing exercise is based on investor type, not investor mandate, we believe this is not an issue in our setting.

Another concern with this method is that it will identify funds tracking subsets of the indices we're actually interested in e.g., S&P 500 value funds like IVE. And, empirically, there is significant variation in the passive ownership share across stocks in these sub indices. That being said, we perform several validation exercises to ensure that our measure of expected buying around index change events is not biased by ignoring these

a portfolio-optimization technique to avoid the expense and impracticality of full replication." This is in contrast to VTSAX and VTI, which are designed to fully replicate the CRSP total market index.

sub-index classifications. The logic is that if about half the stocks added to a particular index are growth stocks, and half the stocks added to the index are value stocks, the fraction of the index's float held by index funds will still capture the average buying across these sub indices and thus the average expected buying by such funds.

#### A.1.7 Data Filters

To be included in our sample, stocks must pass several additional filters. First, we only include ordinary common shares (CRSP share codes 10-11) traded on major exchanges (CRSP exchange codes 1-3). Second, we exclude stocks that are an acquiring permno or have an acquiring permno in either quarter $t$ or quarter $t-1$. This is because in such quarters, there can be large changes in split-adjusted shares outstanding because, e.g., a firm issues shares to acquire another company, which can create extreme outliers in $q_{i,j,t}$.

Furthermore, we require that each stock is included in one of the major index families (S&P 1500, Russell 3000 or CRSP Total Market), because our primary objective is to study market clearing when index funds trade. Our index data for the S&P 1500 universe starts in 2002, our Russell index data starts in 2009, and our CRSP index data (now used by many Vanguard index funds) starts in 2015.[20] We use 2002 to 2021 as our baseline sample, as it covers most stocks in most indices and still captures the significant growth in passive ownership. As a robustness exercise, we re-estimate our baseline empirical tests in Section 3 on various time subsamples. The findings are qualitatively similar to our baseline results. We provide these additional results in Appendix B.5.

### A.2 Forming Investor Groups

#### A.2.1 Overlapping Groups

Figure 5 provides a visual illustration of the data, and the adjustment we use to ensure the categories are mutually exclusive. In the top panel of Figure 5, the circle represents 100% which is the total number of shares outstanding in a hypothetical stock. The orange region represents the ownership of all 40-Act Funds (roughly 35% in this example) and the blue region represents all 13F institutions (roughly 75%). These numbers are roughly in line with a typical stock-quarter in our sample.

The figure also illustrates the overlap of the mutual fund and 13F datasets – nearly all of the 40-Act Fund

[20] Although Vanguard's transition to CRSP indices was initiated in 2012, it was not finished until 2014.

holdings are recorded in some combination of 13F filings for banks, investment companies, independent investment advisors, and miscellaneous institutions. We combine these four 13F categories and subtract all fund holdings to create a separate category. The bottom panel of Figure 5 shows this new financial institutions category, which does not overlap with other groups. This leaves us with 6 mutually-exclusive investor groups that together sum up to total 13F institutional holdings but now with a separate accounting for index funds, active funds and other mutual funds.

The bottom panel of Figure 5 also highlights a placeholder for a residual category that represents all ownership that is outside of 13F filings. We attribute this group largely to retail investors, small and foreign institutional investors, as well as other miscellaneous sources of holdings. This is meant to provide a sense of who might account for owning the remainder of the shares of each company. We will describe this residual category in more detail in Section 2.2, where we discuss the methodology.

[21]

#### A.2.2 Illustration

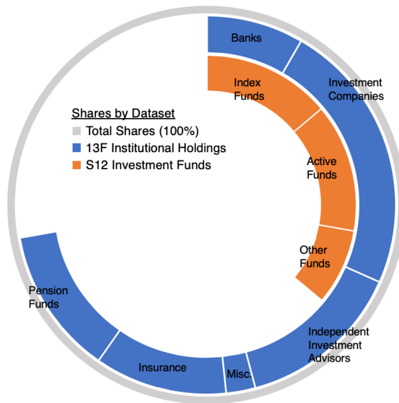
Figure 6 provides an illustrative example of what the data might look like. The figure shows each of the 10 groups and their respective share change. The quantity of the share change is represented by the block width and the direction of their share change is represented by which side of 0% they fall on. This example shows that Index Funds, Pension Funds and Insurance make up a majority of the buying, while Firms, Active Funds, and Financial Institutions make up a majority of the selling. The example also shows that shares changes for the first 9 groups do not clear the market, i.e., they do not sum to zero. Thus, the residual group appears on the selling side to make sure total buying equals total selling.

#### A.2.3 Residual Other Category and Retail Investors

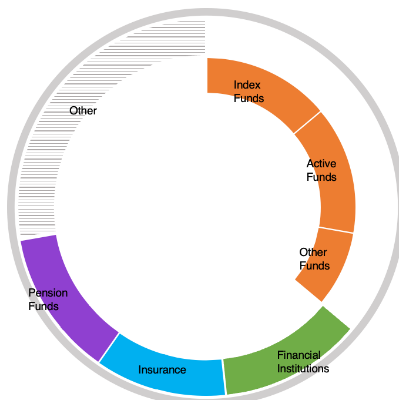
As discussed in the previous section, we have an Other category, whose demand is set to clear the market conditional on the demand of all the investor groups we can observe. This will capture several groups which

[21] It is possible that some of the Other Funds are foreign funds which are not part of institutions which file 13Fs. As we outline above, we subtract *all* the holdings in S12 filings from the 13F filings for groups known to manage mutual funds. We do this because, due to matching issues between S12 and 13F data, we cannot unambiguously determine if an S12 filing institution is part of a 13F filing institution or not. For example, VanEck's individual funds' S12s cannot be matched based on Thompson's identifiers to VanEck's overall 13F. In this case, if we assume that a lack of a match means that the S12 funds are not part of a 13F institution, we would be double counting VanEck's position changes. So, to avoid such double counting, we assume all S12 filings are also part of a 13F filing. This will prevent double counting VanEck's position changes, and thus prevent creating an erroneous offsetting change in our Other group. On the other hand, in the case of a mutual fund which is part of an institution which does not file a 13F, we will erroneously create an offsetting Other trade.

**Figure 5:** Dataset Decomposition and Investor Categories: Example



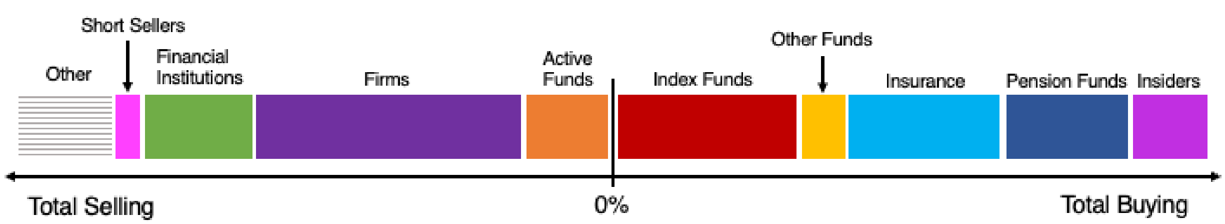

Panel A: Datasets and Categories




Panel B: Mutually-Exclusive Categories

*Notes.* Panel A presents the fraction of the average stock's shares outstanding owned by the investor categories we can observe in the S12 and 13F data. Panel B presents the same breakdown, except the categories have been refined to be mutually exclusive.

**Figure 6:** Share Changes by Group: Example




*Notes.* Example breakdown of net buying and selling by our 10 mutually exclusive investor categories. Bars above zero denote net buying, while bars below zero denote net selling. Because markets must clear, total net buying is by construction equal to total net selling.

we know are not included in our data including retail investors, small institutional investors who do not file 13Fs, and foreign institutional investors who also do not file 13Fs. In this section, we aim to understand whether or not this Other group's behavior is related to proxies of retail trading activity.

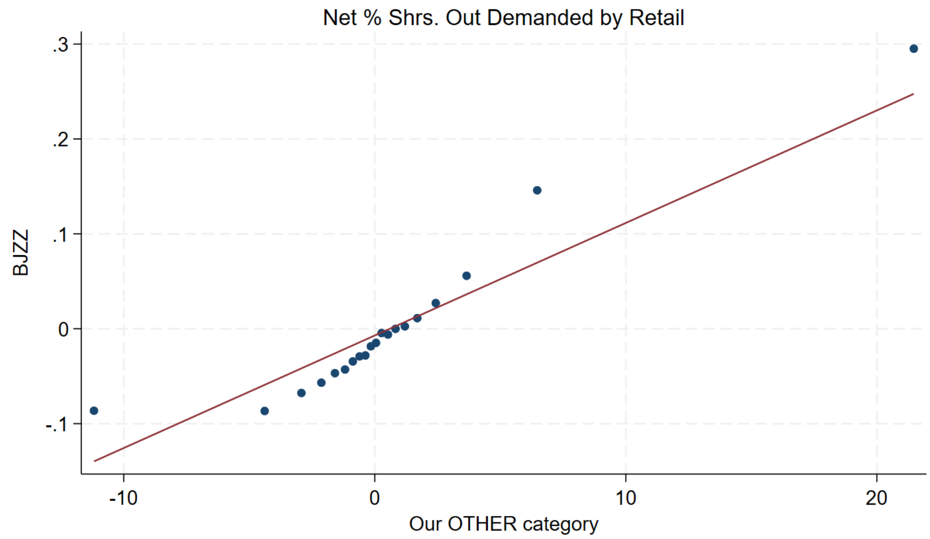
First, we compare our measure of Other buying and selling to the measure of retail buying and selling in Boehmer et al. (2021). Specifically, leveraging a regulatory requirement for wholesalers, we identify marketable retail buy and sell orders using sub-penny price improvements in the TAQ data. Then, for each stock, each day, we are able to construct a measure of net buying by retail investors. We aggregate this up to the stock-quarter level to match the frequency of our measure of Other demand. Of course, this procedure may produce false positives and false negatives at the individual trade level (i.e., the BJZZ algorithm can classify some retail trades as institutional and classify some institutional trades as retail) (Barber et al., 2022; Battalio et al., 2023), but, as discussed in Laarits and Sammon (2023), aggregated versions of this measure are useful for ranking stocks based on retail trading intensity. We use this procedure to identify retail trades between 2010-2021, as before 2010, the algorithm is relatively less effective at identifying retail trades.

Figure 7 presents a binned scatter plot of our measure of Other demand against the measure of net retail trading activity described above which we construced using the algorithm in Boehmer et al. (2021) (hereafter BJZZ). The figure shows that the two measures are strongly positively correlated – suggestive evidence that our measure is indeed capturing retail trading activity. That being said, the scale of the measure constructed using the method in BJZZ is roughly two orders of magnitude smaller than our measure of "Other" trading activity. This could be because BJZZ only captures a fraction of all retail orders (as discussed in Barber et al. (2022) and Battalio et al. (2023), the false negative rate is around 70%), and because our measure – by construction – will pick up net trading by non-retail groups like foreign and small institutions.

Our second validation exercise leverages the retail investor data in Barber and Odean (2000). Specifically, we start with the trade-level data they obtained from a retail brokerage, and aggregate it up to a stock-day measure of net buying by retail investors. This data runs from 1991-1996, and only represents trades at one individual retail brokerage. Therefore, even though the level will likely not match what we find, if this brokerage is representative of the population of retail investors as a whole, we would expected differences in net retail demand to match differences in our measure of "Other" activity. We aggregate this measure of net retail demand to the stock-quarter level to match the frequency of our measure of Other demand.

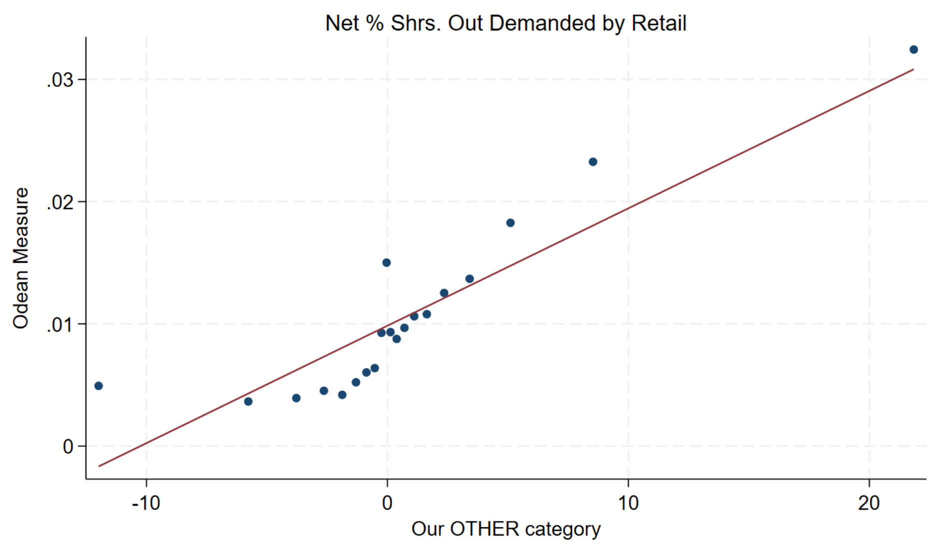
Figure 8 presents a bin scatter of our measure of Other demand against the measure of retail activity constructed using trades in Barber and Odean (2000). As with Figure 7, Figure 8 shows that the two

**Figure 7:** Validation 1: Comparison to BJZZ




*Notes.* The x-axis variable is our measure of net Other demand, expressed as a percentage of shares outstanding. The y-axis variable is the net demand by retail investors, expressed as a percentage of shares outstanding, where retail buy and sell orders are identified using the algorithm in Boehmer et al. (2021) (BJZZ). The unit of observation is stock-quarter.

**Figure 8:** Validation 2: Comparison to Odean Data




*Notes.* The x-axis variable is our measure of net Other demand, expressed as a percentage of shares outstanding. The y-axis variable is the net demand by retail investors, expressed as a percentage of shares outstanding, where retail buy and sell orders are identified using the transaction-level data in Barber and Odean (2000). The unit of observation is stock-quarter.

measures are strongly positively correlated – further evidence that our measure is indeed capturing retail trading activity. That being said, the scale of the measure constructed using the Barber and Odean (2000) data several orders of magnitude smaller than our measure of retail trading activity. This could be because, as discussed above, their data only includes a single retail brokerage.

## A.3 Market Clearing Derivation

From market clearing, we have for a given stock $i$ in a quarter $t$ that

$$\sum_j q_{i,j,t} = 0, \tag{10}$$

and rewriting, we have

$$\sum_j q_{i,j,t} = -q_{i,\mathrm{IDX},t} \tag{11}$$

where $j$ now indexes all groups except Index Funds. We substitute Equation 5 into the expression to yield

$$\sum_j \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t} = -q_{i,\mathrm{IDX},t}. \tag{12}$$

This holds for each stock $i$ in each quarter $t$. This means we can sum over stocks and quarters, or

$$\sum_i \sum_t \sum_j \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t} = \sum_i \sum_t -q_{i,\mathrm{IDX},t}. \tag{13}$$

We can simplify this to

$$\sum_i \sum_t \sum_j \alpha_j + \sum_j \beta_j \left( \sum_i \sum_t q_{i,\mathrm{IDX},t} \right) = \sum_i \sum_t -q_{i,\mathrm{IDX},t}, \tag{14}$$

since the sum of the error term over time and stocks is zero per group $j$. This further simplifies to

$$\frac{\sum_i \sum_t \sum_j \alpha_j}{\sum_i \sum_t q_{i,\mathrm{IDX},t}} + \sum_j \beta_j = -1, \tag{15}$$

or

$$\frac{\sum_j \alpha_j}{\overline{q}_{i,\mathrm{IDX},t}} + \sum_j \beta_j = -1, \tag{16}$$

where $\overline{q}_{\mathrm{IDX}}$ is the average Index Fund ownership change over all stocks and quarters. If $\alpha_j = 0$ for all $j$, this yields

$$\sum_j \beta_j = -1. \tag{17}$$

Given that the alpha represents the estimated quantity bought or sold when Index Fund demand is zero, the market clearing condition also holds in that the sum of the alphas must be zero. That is, the market must clear amongst the other groups when Index Funds do nothing.

# B Baseline Empirical Results Appendix

In this appendix, we present supporting analyses for Section 3.

## B.1 Fixed Effects

We re-estimate our baseline set of regressions and incorporate a set of fixed effects to address potential common drivers of index fund demand and the demand of other investor groups. Specifically, we implement three distinct specifications.

First, we include year-quarter fixed effects. This accounts for commonalities in flows into index funds within a quarter or aggregate market trends that may affect a group's demand for stocks in general in a particular quarter. This isolates within-quarter variation, and helps address the concern that what drives our results is differences in aggregate effects driving both flows and group asset allocation decisions.

Second, we use industry-by-year-quarter fixed effects to more finely control for commonalities in flows and demand within a quarter. For industry classifications, we use the Fama French 49 industries which are based on four-digit SIC codes (See Ken French's data library for more details). This specification isolates variation within an industry in a given quarter. This specification addresses the impact of industry-specific shocks that might simultaneously influence flows into industry-themed funds and patterns in firm issuance within that industry. For example, if there is a technological shock to computer hardware and semiconductor companies because of a revolution in artificial intelligence, this specification will isolate variation within this set of stocks. That is, we test if more Index Fund demand relative to other stocks in the same industry-quarter is associated with relatively more Firm issuance.

Third, we use year-quarter and firm fixed effects to control for aggregate effects across quarters and patterns at the stock level. One possible explanation of our results is that it is driven by cross-sectional variation between stocks. For example, some stocks tend to regularly issue shares and others tend to regular conduct buybacks, and these patterns are also related to what stocks Index Funds tend to buy or sell. This fixed effects specification asks whether each group's demand is related to variation in Index Fund demand within a company over time.

The betas from these specifications are reported in Table 6. The results are qualitatively unchanged and quantitatively similar to the main estimates reported in Section 3.1. The biggest change is that Firms' and

**Table 6:** Beta Estimates with Fixed Effects

| | YQ FE | | | FF49 Industry x YQ FE | | | Stock and YQ FE | | |
|---|---|---|---|---|---|---|---|---|---|
| Investor Group | $\beta_j$ | $t(\beta_j)$ | $R^2$ | $\beta_j$ | $t(\beta_j)$ | $R^2$ | $\beta_j$ | $t(\beta_j)$ | $R^2$ |
| Active Funds | 0.188 | 11.937 | 0.022 | 0.179 | 11.453 | 0.052 | 0.144 | 9.507 | 0.066 |
| Other Funds | 0.065 | 8.790 | 0.100 | 0.062 | 8.716 | 0.131 | 0.051 | 7.231 | 0.142 |
| Pension Funds | 0.026 | 10.362 | 0.044 | 0.025 | 10.158 | 0.075 | 0.022 | 9.036 | 0.077 |
| Insurance | 0.065 | 13.513 | 0.036 | 0.063 | 13.299 | 0.067 | 0.061 | 12.916 | 0.060 |
| Financial Institutions | -0.057 | -1.198 | 0.050 | -0.085 | -1.794 | 0.083 | -0.153 | -3.298 | 0.096 |
| Insiders | -0.089 | -12.228 | 0.008 | -0.086 | -12.141 | 0.033 | -0.062 | -11.633 | 0.132 |
| Other | -0.267 | -5.703 | 0.039 | -0.282 | -6.166 | 0.071 | -0.299 | -7.063 | 0.126 |
| Short Sellers | -0.287 | -10.534 | 0.079 | -0.283 | -10.337 | 0.112 | -0.273 | -9.967 | 0.108 |
| Firms | -0.644 | -16.358 | 0.044 | -0.594 | -17.069 | 0.092 | -0.490 | -16.256 | 0.212 |
| Total | -1.000 | | | -1.001 | | | -0.999 | | |

*Notes.* Estimates from a modified version of our baseline regression specification with stock and quarter fixed effects:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \gamma_i + \psi_t + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

Active Funds' beta are smaller in magnitude with stock and year-quarter fixed effects: Firms' beta is -0.490 (vs. the baseline estimate of -0.642) and Active Funds' beta is 0.144 (vs. 0.193).

## B.2 Non-Linearities

The binscatter plots in Section 3.1.2 show that Firms' and Short Sellers' demand has a negative relationship to Index Fund demand, but that it is driven entirely by firm-quarters where passive investors are net buyers of shares. To quantify this difference, we re-estimate our baseline regressions but split the sample based on whether Index Funds' were net buyers or sellers of a stock. Table 7 presents the estimates for each subsample. Unsurprisingly, the number of observations in the positive change sample is much larger than the negative change sample, given the consistent growth of passive funds over this time period. There is still a sizeable sample, about a quarter of all observations, that saw Index Funds sell shares on net.

The beta estimates reveal several stark differences between these two subsamples. First, the right panel shows that when Index Funds buy shares, Firms on average issue shares on a one-for-one basis, with an estimated $\beta_j$ of almost exactly -1. The left panel shows that when Index Funds sell shares, Firms do the opposite as when Index Funds buy shares by selling alongside Index Funds, i.e., by issuing shares, although the magnitude is significantly smaller. Second, Short Sellers exhibit a similar pattern to Firms, in that they sell when Index Funds are buying but do not respond much to passive selling. Third, Financial Institutions trade much more with Index Funds for passive sales, consistent with the scatter plots presented in Section

**Table 7:** Regression Estimates: Positive vs. Negative Passive Position Change

| | Negative Passive Ownership Change | | | | | | Positive Passive Ownership Change | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | Obs. | $R^2$ | $\hat{q}_j$ | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | Obs. | $R^2$ | $\hat{q}_j$ |
| Active Funds | 0.176 | 5.563 | -0.045 | 46,131 | 0.003 | -0.130 | 0.207 | 5.605 | -0.066 | 125,870 | 0.006 | 0.065 |
| Other Funds | 0.027 | 2.438 | 0.040 | 46,131 | 0.001 | 0.027 | 0.109 | 2.971 | 0.041 | 125,870 | 0.014 | 0.128 |
| Pension Funds | 0.007 | 1.184 | -0.014 | 46,131 | 0.000 | -0.017 | 0.027 | 6.725 | -0.011 | 125,870 | 0.005 | 0.007 |
| Insurance | 0.046 | 5.169 | -0.035 | 46,131 | 0.002 | -0.057 | 0.073 | 6.545 | -0.034 | 125,870 | 0.007 | 0.022 |
| Financial Institutions | -0.333 | -3.949 | 0.134 | 46,131 | 0.004 | 0.295 | -0.033 | -0.366 | 0.011 | 125,870 | 0.000 | 0.030 |
| Insiders | -0.024 | -2.668 | -0.007 | 46,131 | 0.000 | 0.005 | -0.104 | -8.923 | -0.019 | 125,870 | 0.005 | -0.109 |
| Other | -0.950 | -9.390 | -0.056 | 46,131 | 0.026 | 0.402 | -0.006 | -0.080 | 0.049 | 125,870 | 0.000 | 0.149 |
| Short Sellers | -0.136 | -2.874 | 0.008 | 46,131 | 0.003 | 0.074 | -0.320 | -7.028 | 0.118 | 125,870 | 0.019 | -0.115 |
| Firms | 0.185 | 6.132 | -0.025 | 46,131 | 0.002 | -0.114 | -0.953 | -16.097 | -0.089 | 125,870 | 0.047 | -0.864 |
| Total | -1.002 | | -0.000 | | | 0.483 | -1.001 | | 0.001 | | | -0.685 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The sample is split based on whether Index Funds were net buyers in a given stock in a given quarter (positive passive position change) or whether they were net sellers in a given stock in a given quarter (negative passive position change). The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

3.1.2.

Lastly, the residual Other group does not, on average, need to do much to clear the market for passive buying. The estimates for the Other group is -0.091 and is statistically indistinguishable from zero, suggesting that for Index Fund buying the groups that are constructed directly from the data nearly clear the market amongst themselves. On the other hand, Index Fund selling is nearly completely bought by our residual Other group. One explanation for this is that net selling by index funds may be the result of stocks leaving the investable universe (e.g., leaving the Russell or S&P universes), thus making other, non-US institutional investors (e.g., retail and foreign investors) more natural buyers.

## B.3 Value-Weighted Regressions

The baseline regressions give each stock-quarter an equal weight. We re-estimate the set of regressions but on a value-weighted basis, by giving each observation a weight within each quarter proportional to its share of total market capitalization at the end of quarter $t-1$ (i.e., the beginning of quarter $t$). The baseline regressions give a sense of who typically clears the market for the average stock. The value-weighted regressions give a sense of who typically clears the market for the average dollar each quarter. These regressions also highlight the difference in equal- and value-weighted average Firm activity, which we expand on in Appendix D.

Table 8 provides the alpha and beta estimates for each group, as well as the average change per group, $\overline{q}_j$. The value-weighted regressions tell a similar story to the equal-weighted regressions above in terms of the

**Table 8:** Regression Estimates: Value-Weighted Regressions

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\hat{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.233 | 4.089 | -0.142 | -7.424 | 172,807 | 0.006 | -0.142 |
| Other Funds | 0.188 | 2.516 | 0.054 | 2.605 | 172,807 | 0.026 | 0.054 |
| Pension Funds | 0.032 | 5.646 | -0.023 | -2.419 | 172,807 | 0.005 | -0.023 |
| Insurance | 0.090 | 3.910 | -0.058 | -3.597 | 172,807 | 0.005 | -0.058 |
| Financial Institutions | -0.382 | -3.103 | -0.048 | -0.423 | 172,807 | 0.005 | -0.048 |
| Insiders | -0.081 | -6.470 | -0.034 | -3.116 | 172,807 | 0.004 | -0.034 |
| Other | -0.182 | -1.893 | -0.086 | -0.760 | 172,807 | 0.001 | -0.086 |
| Short Sellers | -0.162 | -6.990 | 0.022 | 1.055 | 172,807 | 0.006 | 0.022 |
| Firms | -0.736 | -15.000 | 0.315 | 7.080 | 172,807 | 0.038 | 0.315 |
| Total | -1.000 | | -0.000 | | | | -0.000 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$ with each observation given a weight proportional to stock $i$'s share of total stock market capitalization in quarter $t-1$. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

sensitivities of each group: Firms and Short Sellers together have a sensitivity of less than -1. In fact, Firms are a bit more sensitive on a value-weighted basis (-0.774 vs. the equal-weighted estimate of -0.694). The most significant difference is the role of Financial Institutions, which has a statistically significant estimate of -0.360, compared to the equal-weighted estimate of -0.054. That is, Financial Institutions are more responsive in dollar terms than percent ownership terms in clearing the market for Index Fund demand. To offset the increased responsiveness of Financial Institutions, Other and Short Sellers have a small decrease in their beta magnitudes. All of the groups with positive equal-weighted betas also have positive betas in dollar terms.

The value-weighted estimates of $\overline{q}_j$ show substantial differences from their equal-weighted counterparts. The group that takes the majority of the other side of Index Funds for the average dollar is not Firms, but rather Financial Institutions and Active Funds. This is consistent with outflows from hedge funds (captured within Financial Institutions) and Active Funds in aggregate. In addition, Firms have a $\overline{q}_j$ that is positive, consistent with significant buyback programs instituted by large public companies.

The results in Table 8 suggest that each of these three groups – Firms, Financial Institutions, and Active Funds – have slightly more nuanced roles in clearing the market for passive demand than suggested by their beta estimates alone. While in value-weighted terms Firms buy back shares on average, they buy back significantly fewer and/or issue more shares as Index Funds increase their demand. Further, the average dollar demand by Financial Institutions offsets Index Fund demand, and they continue to supply additional dollars as Index Funds demand more, though at about half the rate of Firms. Finally, the average dollar demand for Active Funds also offsets Index Fund demand, but they respond to greater passive demand by selling less/buying more of the same stocks.

The equal- and value-weighted estimates collectively point to a single overarching story. Firms are *by far* the most responsive to changes in Index Fund demand at both the stock and dollar level – when Index Funds buy more, Firms provide more shares. In addition, Short Sellers and Insiders also provide more shares as Index Funds demand more shares. All 13F groups besides Financial Institutions – Active Funds, Other Funds, Insurance, and Pension Funds – increase their demand for shares when Index Funds demand more shares. These statements all speak to the *relative* responsiveness of these groups.

**Value-Weighted Binscatter Plots** We provide binscatter plots that correspond to the value-weighted regressions in Table 8. Figure 9 presents the data in a value-weighted binscatter plot. We divide the data into 100 bins, and all observations in each bin are weighted by the firm's share of total market capitalization at the end of the previous quarter. The plots show very similar patterns to the equally-weighted plots in Section 3.1.2. The biggest change is with Financial Institutions, which look to play a more definitive role in clearing the market for passive for all but the largest Index Fund ownership increases. The other significant change is with the Other group, which looks much smaller and noisier, consistent with both retail investors focusing on small-cap stocks and a greater chance of data errors in small-cap stocks.

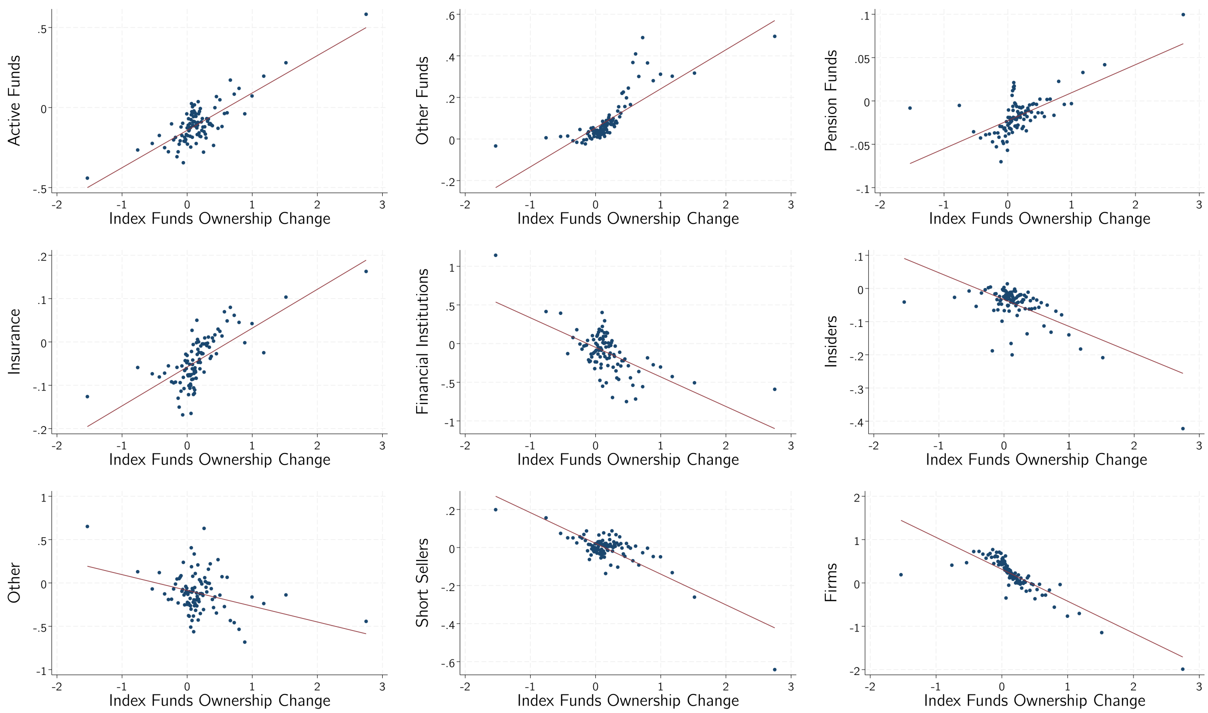
## B.4 Year-over-Year Estimates

Since we are interested in which groups ultimately clear the market in the long run, we estimate our baseline equal-weighted regressions but with year-over-year changes. This helps address the possibility that some groups may act as intermediaries over periods as long as two months, and our quarterly analysis misses the "final" buyer or seller. For example, a hedge fund may sell shares to an index mutual fund a few weeks before the end of a quarter, then the hedge fund may buy the shares back from an active mutual fund manager two months later. Therefore, our quarterly data may miss, e.g., a long-run transaction between an index mutual fund and an active mutual fund intermediated by the hedge fund.

We estimate a series of regressions with overlapping time periods. The regression we estimate for each group in our sample is

$$q_{i,j,t\to t+4} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t\to t+4} + \varepsilon_{i,j,t}, \tag{18}$$

**Figure 9:** Value-Weighted Binscatter by Group




*Notes.* Each panel presents a value-weighted binscatter of net demand by each investor group – $q_{i,j,t}$ – against net demand by Index Funds – $q_{i,\mathrm{IDX},t}$. The unit of observation is security-year-quarter.

where the ownership change for each group $j$ or the Index Fund group is computed over four quarters (i.e., from the end of quarter $t$ to the end of quarter $t+4$). Because we have overlapping observations, we adjust standard errors for clustering at the stock- and year-quarter-level, as well as for autocorrelation up to 6 quarters, i.e., we follow the standard practice of including lags equal to 1.5$\times$ the number of overlapping observations.

Table 9 presents the estimates. If anything, we find that Firms play an even bigger role in clearing the market for passive demand. For the average passive demand, Firms provide *all* of the shares on a year-over-year basis, including not only the shares demanded by Index Funds, but also all of the shares the other groups demand alongside Index Funds. This is captured by the series of $\overline{q}_j$, which show that Other Funds, Financial Institutions, and Other all demand shares on average along side Index funds. Essentially all of those shares are provided by Firms.

In addition, Firms are by far the most responsive group, with a beta of -1.433. This beta magnitude dwarfs all others. The closest other groups in magnitude are Financial Institutions and Active Funds, but both of these groups have the opposite sign. That is, these groups demand *more* shares when Index Funds demand more shares. This means that Firms not only respond to provide shares for passive investors as they demand more shares, but for the position changes of other groups that mirror the demand of Index Funds.

The other notable difference in Table 9 relative to the baseline quarterly results is with Financial Institutions. The quarterly estimates suggest that Financial Institutions take the other side of Index Fund demand, but, over the long run, the year-over-year regressions show that they actually adjust holdings in the same direction as Index Funds. Moreover, the magnitude of the year-over-year beta estimate is larger than any other group. This suggests that these Financial Institutions may indeed act more as relatively short-run intermediaries but, in the long-run, are ultimately buying the same stocks as Index Funds.

## B.5 Sample Selection Robustness

**Sample Selection** We show that our baseline findings are even stronger when examining subsamples of the data. We report regression estimates from 2009 to 2021 in Table 10 and from 2015 to 2021 in Table 11.

**Treatment of Outliers** Section 2.2 outlines our reasons for trimming outliers for the regressions reported in Section 3.1. We repeat our main analyses by using the raw data and by Winsorizing outliers at the 0.5 and 99.5 percentiles. We report the regression estimates with these alternative approaches for handling outliers

**Table 9:** Beta Estimates: Year-over-Year Regressions

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\hat{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.325 | 7.913 | -0.411 | -3.413 | 151,744 | 0.024 | -0.006 |
| Other Funds | 0.150 | 5.153 | 0.170 | 2.364 | 151,744 | 0.046 | 0.357 |
| Pension Funds | 0.042 | 7.520 | -0.077 | -2.751 | 151,744 | 0.020 | -0.025 |
| Insurance | 0.099 | 12.640 | -0.156 | -3.668 | 151,744 | 0.024 | -0.033 |
| Financial Institutions | 0.471 | 4.138 | -0.086 | -0.265 | 151,744 | 0.017 | 0.501 |
| Insiders | -0.148 | -7.128 | -0.069 | -1.458 | 151,744 | 0.009 | -0.253 |
| Other | -0.252 | -1.462 | 1.137 | 3.099 | 151,744 | 0.003 | 0.823 |
| Short Sellers | -0.256 | -11.193 | 0.136 | 0.979 | 151,744 | 0.025 | -0.183 |
| Firms | -1.433 | -14.584 | -0.645 | -2.257 | 151,744 | 0.100 | -2.431 |
| Total | -1.002 | | -0.001 | | | | -1.249 |

*Notes.* Estimates from the year-over-year version of our baseline regression specification:

$$q_{i,j,t\to t+4} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t\to t+4} + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are adjusted for clustering at the stock-level, year-quarter-level, as well as for autocorrelation up to 6 quarters.

**Table 10:** Regression Estimates: 2009-2021

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\overline{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.203 | 6.932 | -0.066 | -2.455 | 137,431 | 0.011 | 0.004 |
| Other Funds | 0.091 | 3.172 | 0.070 | 2.940 | 137,431 | 0.015 | 0.101 |
| Pension Funds | 0.019 | 6.078 | -0.004 | -0.813 | 137,431 | 0.004 | 0.003 |
| Insurance | 0.070 | 8.816 | -0.023 | -1.578 | 137,431 | 0.015 | 0.001 |
| Financial Institutions | -0.054 | -0.803 | 0.115 | 1.421 | 137,431 | 0.000 | 0.096 |
| Insiders | -0.101 | -11.170 | -0.039 | -4.252 | 137,431 | 0.006 | -0.074 |
| Other | -0.272 | -4.765 | 0.347 | 4.557 | 137,431 | 0.004 | 0.254 |
| Short Sellers | -0.263 | -9.198 | 0.037 | 0.848 | 137,431 | 0.023 | -0.053 |
| Firms | -0.694 | -15.837 | -0.435 | -10.099 | 137,431 | 0.037 | -0.673 |
| Total | -1.001 | | 0.002 | | | | -0.341 |

Notes. The table provides estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t},$$

for each investor group $j$. $q_{i,j,t}$ is the quarterly holdings change in stock $i$ for group $j$ in year-quarter $t$ in units of percent ownership of the company. $q_{i,\text{IDX},t}$ is the ownership change for Index Funds. T-statistics are computed from standard errors double clustered at the stock and time (year-quarter) level. The last column reports the average quantity change for each group across all stocks and quarters ($\overline{q}_j$). See Section 2.2 for more details on the investor groups, Section 2.3 for the empirical methodology, and Section 3.1 for more details on the table.

**Table 11:** Regression Estimates: 2015-2021

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\bar{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.179 | 5.198 | -0.088 | -2.419 | 80,729 | 0.013 | -0.014 |
| Other Funds | 0.105 | 2.890 | 0.094 | 2.328 | 80,729 | 0.022 | 0.137 |
| Pension Funds | 0.021 | 5.682 | -0.007 | -1.224 | 80,729 | 0.009 | 0.002 |
| Insurance | 0.072 | 7.672 | -0.011 | -0.564 | 80,729 | 0.027 | 0.019 |
| Financial Institutions | 0.053 | 0.704 | -0.013 | -0.152 | 80,729 | 0.000 | 0.009 |
| Insiders | -0.093 | -10.669 | 0.000 | -0.041 | 80,729 | 0.007 | -0.038 |
| Other | -0.264 | -3.657 | 0.490 | 6.942 | 80,729 | 0.005 | 0.381 |
| Short Sellers | -0.301 | -9.488 | 0.074 | 1.681 | 80,729 | 0.038 | -0.050 |
| Firms | -0.771 | -12.621 | -0.537 | -9.103 | 80,729 | 0.044 | -0.855 |
| Total | -0.999 | | 0.002 | | | | -0.410 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. Run on the subsample from 2015-2021. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

in Tables 12 and 13.

## B.6 Data Errors

One concern arises from the nature of the data. As described in Section 2 and documented in detail in Sammon and Shim (2023), the S12 data is littered with many types of errors, some of which involve staleness in reported holdings. We also find evidence of data errors in the Thomson 13F data. We address two types of data errors: (1) general data errors where a group appears to increase or decrease its ownership of a stock but does not in reality, and (2) stale data.

General data errors result in an inability to clear the market amongst the groups in our sample, and force the residual Other group to take a position that mechanically clears the market. This will have the effect of attenuating the beta estimates of each of the non-residual groups, and push the beta estimate of the Other group toward -1. In this sense, the data errors have the same effect as the well-known attenuation bias due to measurement error. This suggests that our estimates for the non-residual groups in Table 1 are an underestimate.

**Table 12:** Beta Estimates: 2002-2021, Raw Data

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\hat{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.516 | 4.965 | -0.166 | -2.914 | 185,799 | 0.068 | 0.014 |
| Other Funds | 0.132 | 6.669 | 0.049 | 1.897 | 185,799 | 0.034 | 0.095 |
| Pension Funds | 0.052 | 3.724 | -0.019 | -1.747 | 185,799 | 0.012 | -0.001 |
| Insurance | 0.104 | 7.616 | -0.044 | -2.622 | 185,799 | 0.023 | -0.008 |
| Financial Institutions | 0.632 | 2.170 | -0.011 | -0.067 | 185,799 | 0.021 | 0.210 |
| Insiders | -0.114 | -6.481 | -0.092 | -5.193 | 185,799 | 0.002 | -0.132 |
| Other | -1.325 | -2.605 | 0.867 | 3.303 | 185,799 | 0.055 | 0.405 |
| Short Sellers | -0.221 | -6.066 | 0.009 | 0.155 | 185,799 | 0.017 | -0.068 |
| Firms | -0.775 | -7.638 | -0.593 | -8.060 | 185,799 | 0.035 | -0.863 |
| Total | -0.999 | | -0.000 | | | | -0.349 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. Unlike in Table 1, the data is not trimmed for each group at the 0.5 and 99.5 percentiles. Instead, we include the raw (untrimmed, unwinsorized) data. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

**Table 13:** Beta Estimates: 2002-2021, Winsorized Data

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\hat{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.193 | 7.015 | -0.052 | -2.209 | 172,807 | 0.007 | 0.011 |
| Other Funds | 0.086 | 3.148 | 0.060 | 3.361 | 172,807 | 0.012 | 0.088 |
| Pension Funds | 0.022 | 7.423 | -0.007 | -1.243 | 172,807 | 0.004 | 0.000 |
| Insurance | 0.065 | 7.894 | -0.029 | -2.281 | 172,807 | 0.008 | -0.008 |
| Financial Institutions | -0.152 | -2.265 | 0.123 | 1.445 | 172,807 | 0.001 | 0.073 |
| Insiders | -0.084 | -10.920 | -0.033 | -4.684 | 172,807 | 0.005 | -0.060 |
| Other | -0.250 | -4.674 | 0.228 | 2.918 | 172,807 | 0.003 | 0.146 |
| Short Sellers | -0.238 | -6.433 | 0.037 | 0.802 | 172,807 | 0.015 | -0.041 |
| Firms | -0.642 | -16.196 | -0.327 | -8.757 | 172,807 | 0.032 | -0.537 |
| Total | -1.000 | | 0.000 | | | | -0.327 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. Unlike in Table 1, the data is not trimmed for each group at the 0.5 and 99.5 percentiles. Instead, we Winsorize at the 0.5 and 99.5 percentiles. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

### B.6.1 Minimal Residual Demand

All of our measures of net demand come directly from the data except our residual Other category, which is necessary to ensure market clearing holds in our data. The Other category by construction makes a position change that all other investors that are not included in our data must have done in order for the market to clear. There are a few ways to look at this. One is that our Other category is actually picking up some combination of foreign investors, small institutions, and retail traders (in Appendix A.2.3, we provide evidence that our Other category is related to measures of retail trading activity). An alternative interpretation is that this Other category represents an aggregation of data errors for the investors we do have in our data. This can be due to issues with mistiming (e.g., reporting delays, stale data) or just pure data errors, and are all collected in this residual group. Last, it could be that there is some kind of dark matter that clears the market, but we do not know who they are.

As a check, we examine only the observations where the market clears or nearly clears amongst the investor groups we can directly measure in our data. That is, we filter out all observations where our residual group are required to trade more than 0.50% of shares outstanding to make the market clear. Table 14 presents the estimates.

While this exercise omits a significant fraction of the data, the general message is consistent with the baseline results – Firms and Short Sellers collectively account for a significant fraction of the shares demanded by passive. The magnitude is a bit lower than the baseline estimates: Firms have a beta of -0.404 (vs. -0.642 in the baseline) and Short Sellers have a beta of -0.17 (vs. -0.238). Financial Institutions play a much larger role in clearing the market in this sample, with a beta estimate of -0.514 (vs. -0.152). The changes in point estimates are likely driven by observations with large positive or negative changes in Index Fund ownership, which we discuss and explore more in Appendix B.2.

### B.6.2 Stale Data

Stale data may impact the results in an a way that is different than more general data errors. Imagine a scenario where, in reality, Index Funds buy a stock in period $t$ and Active Funds sell to them, but Active Funds' sales are erroneously not recorded in the data and stale holdings from the previous quarter are recorded instead. In order to clear the market, the residual Other group will then be responsible for clearing the market in period $t$. In addition, the Active Fund sale, which is now recorded with a delay in $t + 1$, will be cleared by the Other group's buying. That is, the Other group will look as if it acts as an intermediary

**Table 14:** Regression Estimates: Sample with Low Other Group Activity

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\hat{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.028 | 0.996 | -0.062 | -3.121 | 36,785 | 0.000 | -0.055 |
| Other Funds | 0.059 | 3.327 | 0.055 | 3.599 | 36,785 | 0.005 | 0.070 |
| Pension Funds | 0.013 | 4.126 | -0.008 | -1.624 | 36,785 | 0.001 | -0.005 |
| Insurance | 0.041 | 7.334 | -0.026 | -2.379 | 36,785 | 0.003 | -0.016 |
| Financial Institutions | -0.514 | -9.750 | -0.029 | -0.669 | 36,785 | 0.026 | -0.155 |
| Insiders | -0.046 | -6.410 | 0.011 | 1.969 | 36,785 | 0.004 | -0.000 |
| Other | -0.008 | -2.791 | -0.012 | -4.741 | 36,785 | 0.000 | -0.014 |
| Short Sellers | -0.170 | -4.690 | 0.083 | 2.412 | 36,785 | 0.009 | 0.041 |
| Firms | -0.404 | -12.503 | -0.012 | -0.622 | 36,785 | 0.035 | -0.111 |
| Total | -1.001 | | -0.000 | | | | -0.246 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. Run on the subsample of observations where the absolute value of $q_{i,j,t}$ is less than 0.50%. t-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

between groups over quarters.

To quantify how problematic stale data is, we test the degree to which passive changes in quarter $t$ are related to other groups' position changes in the same stock but in quarter $t+1$. That is, we estimate

$$q_{i,j,t+1} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t+1} \tag{19}$$

for each group in our sample. The notable difference is the left-hand side variable, $q_{i,j,t+1}$, is in quarter $t+1$, not $t$.

If the sign of the beta estimate for the Other group flips sign from $t$ (the baseline regressions) to $t+1$, that would be consistent with, though not definitive proof of, the data being stale. We report these beta estimates in Table 15.

The table shows a beta estimate for the residual group of 0.097. That is, Index Fund changes at $t$ are negatively related to the residual group's position change in $t$ but positively related in $t+1$. This is consistent with at least some portion of the data being stale. However, the degree to which stale data affects the findings seems low, given the magnitude is significantly smaller in $t+1$ than in $t$ (0.124 vs -0.271). In addition, the groups most likely to have stale data appear are those that have the opposite of Other in $t+1$, which, again, is Firms.

This robustness test also yields more insight into the role of Firms and Short Sellers. Table 15 shows that

**Table 15:** Beta Estimates with Future Group Changes

| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | Obs. | $R^2$ | $\hat{q}_j$ |
|---|---|---|---|---|---|---|---|
| Active Funds | 0.065 | 5.731 | -0.050 | -1.845 | 162,315 | 0.001 | -0.027 |
| Other Funds | 0.032 | 6.267 | 0.069 | 2.769 | 162,315 | 0.002 | 0.080 |
| Pension Funds | 0.008 | 3.696 | -0.010 | -1.704 | 162,315 | 0.001 | -0.007 |
| Insurance | 0.019 | 5.162 | -0.019 | -1.432 | 162,315 | 0.001 | -0.012 |
| Financial Institutions | 0.025 | 0.662 | 0.050 | 0.586 | 162,315 | 0.000 | 0.059 |
| Insiders | -0.047 | -9.559 | -0.034 | -4.757 | 162,315 | 0.002 | -0.050 |
| Other | 0.097 | 2.681 | 0.105 | 1.399 | 162,315 | 0.001 | 0.139 |
| Short Sellers | -0.081 | -3.313 | 0.012 | 0.256 | 162,315 | 0.002 | -0.016 |
| Firms | -0.260 | -13.778 | -0.368 | -9.386 | 162,315 | 0.006 | -0.458 |
| Total | -0.142 | | -0.245 | | | | -0.294 |

*Notes.* Estimates a modified version of our baseline regression specification where we compare investors' net demand in quarter $t+1$ against passive demand in quarter $t$:

$$q_{i,j,t+1} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

these groups adjust positions in $t+1$ in the opposite direction of Index Fund changes in $t$. That is, if Index Funds buy a stock in quarter $t$, Firms and Short Sellers tend to sell in both quarters $t$ and $t+1$. The estimates for Firms and Short Sellers, respectively, are -0.26 and -0.081. While the betas for future changes are about a third of the magnitude of betas for contemporaneous changes, they are the most economically significant coefficients for quarter $t+1$.

## B.7 By Stock Characteristics

### B.7.1 By Industry

One reason why Firms issue shares is for employee compensation. Employees may receive shares, then sell them in the market, which could end up in the hands of Index Funds. Note, however, that this does not include current executives, which are captured by our Insiders group. This type of compensation appears to be more prevalent with technology firms and young firms.

In this appendix, we examine whether the tendency for some industries to issue more stock (perhaps via compensation) are more responsive to Index Fund demand. This may also capture general tendencies in primary market activity by industry, including how responsive primary market activity in an industry (buybacks and issuance) might be to secondary market activity (Ma, 2019).

Table 16 presents the beta estimates in Panel (A) and the average position change in Panel (B) for each of the Fama-French 10 industries.[22] The table shows that beta estimates vary the most for Firms and Financial Institutions, and they tend to substitute for one another, just as in non-linear analysis in Section 3.1.2 and Appendix B.2, and the over-time analysis in Appendix B.8. Firms do indeed tend to be more responsive to Index Fund changes in industries known for more stock compensation (high-technology and healthcare), but also are active in other more surprising industries (energy and "other", which contains the financial sector). Regardless, Firms appear to at least be somewhat responsive to Index Funds for *every* industry. On the other hand, Financial Institutions have estimates close to zero or positive for several industries. The other seven groups have largely similar estimates, regardless of industry.

Panel (B) also reports the average position change by group. This shows that Firms and Insiders have the most consistently negative estimates across industries. On the other hand, Financial Institutions have just as many negative industry averages as positive. Taken together, Panels (A) and (B) suggest that Firms most steadily provide shares to Index Funds and are the most responsive in issuing more shares with greater Index Fund demand, regardless of industry. There is also heterogeneity across industries – energy, high tech, healthcare have both the most issuance on average and are the most responsive in issuing as Index Funds demand more.

## B.8 Beta Estimates over Time

Given the time trend in passive ownership, significant events that likely adjusted portfolio allocations like COVID-19, and improvements in data quality, we further study how our series of beta estimates vary over time. To check for time trends in who is on the other side of passive trades, we estimate the baseline regressions for each group, but do so separately for each quarter in our data. That is, we estimate a series of cross-sectional regressions, which recover a beta estimate per non-passive group per quarter.

Figure 10 visually presents the beta estimates for each group using an 8-quarter moving average to smooth out estimation errors and get a better sense of the general patterns.

There are two significant trends, each going in the opposite direction. Firms have beta estimates that are growing in absolute value, going from around -0.5 at the beginning of the sample and steadily declining starting from around 2008 to end the sample with estimates of nearly -1. Over the same period, Financial Institutions are steadily increasing, from an estimate of around -0.5 at the beginning of the sample to around 0.1 at the end.

[22] See Ken French's data library for details on industry classifications.

**Table 16:** Estimates by Industry

Panel A: Betas

| Investor Group | Nondurables | Durables | Manuf. | Energy | High Tech | Telecom | Shops | Health | Utilities | Other |
|---|---|---|---|---|---|---|---|---|---|---|
| Active Funds | 0.091 | 0.103 | 0.080 | 0.239 | 0.153 | 0.129 | 0.175 | 0.227 | 0.051 | 0.239 |
| Other Funds | 0.063 | 0.071 | 0.067 | 0.076 | 0.089 | 0.052 | 0.072 | 0.080 | 0.090 | 0.097 |
| Pension Funds | 0.022 | 0.018 | 0.011 | 0.019 | 0.022 | 0.021 | 0.025 | 0.024 | 0.007 | 0.024 |
| Insurance | 0.045 | 0.096 | 0.057 | 0.066 | 0.057 | 0.043 | 0.063 | 0.074 | 0.064 | 0.069 |
| Fin. Insts. Ex. Funds | -0.428 | -0.479 | -0.470 | -0.022 | -0.280 | -0.400 | -0.490 | -0.005 | -0.373 | 0.027 |
| Insiders | -0.044 | -0.084 | -0.052 | -0.051 | -0.071 | -0.074 | -0.099 | -0.043 | -0.054 | -0.111 |
| Other | -0.289 | -0.300 | -0.195 | -0.392 | -0.247 | -0.165 | -0.297 | -0.353 | -0.227 | -0.229 |
| Short Sellers | -0.198 | -0.191 | -0.139 | -0.288 | -0.184 | -0.166 | -0.181 | -0.308 | -0.125 | -0.278 |
| Firms | -0.262 | -0.233 | -0.360 | -0.647 | -0.539 | -0.439 | -0.267 | -0.697 | -0.432 | -0.838 |
| Total | -1.000 | -0.999 | -1.001 | -1.000 | -1.000 | -0.999 | -0.999 | -1.001 | -0.999 | -1.000 |

Panel B: $\hat{q}_j$

| Investor Group | Nondurables | Durables | Manuf. | Energy | High Tech | Telecom | Shops | Health | Utilities | Other |
|---|---|---|---|---|---|---|---|---|---|---|
| Active Funds | 0.010 | -0.062 | -0.024 | 0.013 | -0.020 | -0.098 | -0.126 | 0.069 | 0.072 | 0.063 |
| Other Funds | 0.071 | 0.067 | 0.079 | 0.067 | 0.097 | 0.041 | 0.058 | 0.091 | 0.101 | 0.101 |
| Pension Funds | -0.013 | -0.009 | -0.009 | 0.012 | -0.000 | 0.002 | -0.015 | 0.005 | -0.001 | 0.006 |
| Insurance | -0.033 | -0.018 | -0.015 | -0.013 | -0.010 | -0.027 | -0.027 | 0.002 | 0.003 | 0.005 |
| Fin. Insts. Ex. Funds | -0.060 | -0.120 | -0.089 | 0.094 | 0.133 | 0.032 | -0.162 | 0.259 | 0.177 | 0.141 |
| Insiders | -0.050 | -0.013 | -0.019 | -0.067 | -0.047 | -0.047 | -0.081 | -0.014 | -0.015 | -0.094 |
| Other | 0.040 | 0.000 | -0.019 | 0.509 | 0.076 | 0.195 | 0.050 | 0.382 | -0.052 | 0.200 |
| Short Sellers | -0.033 | 0.002 | 0.006 | -0.139 | -0.008 | -0.019 | -0.018 | -0.071 | -0.030 | -0.062 |
| Firms | -0.181 | -0.101 | -0.183 | -0.879 | -0.531 | -0.348 | 0.073 | -1.066 | -0.582 | -0.738 |
| Total | -0.249 | -0.254 | -0.273 | -0.403 | -0.312 | -0.270 | -0.249 | -0.342 | -0.326 | -0.377 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The regression is estimated separately, in each column only including firms which are a member of each of the listed industries. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

These patterns suggest that the void left in responsiveness by Financial Institutions, who by the end of the sample are, on average, buying and selling more when Index Funds buy or sell more, are offset by Firms who are becoming increasingly more responsive.

Our residual Other group's beta also appears to be trending somewhat toward zero for most of the sample, with an estimate around -0.4 within the first few years of our sample ending up around -0.1 at the very end of out sample. This is consistent with the Other group offsetting some data errors, but data quality improving over time. This could also relate to retail investors increasingly trading in a way that is orthogonal to Index Fund demand.

Many of the smoothed estimates are relatively stable over time. For example, Active Funds, Short Sellers, Insiders, and Insurance all have beta estimates that look roughly unchanged over the entire sample period.

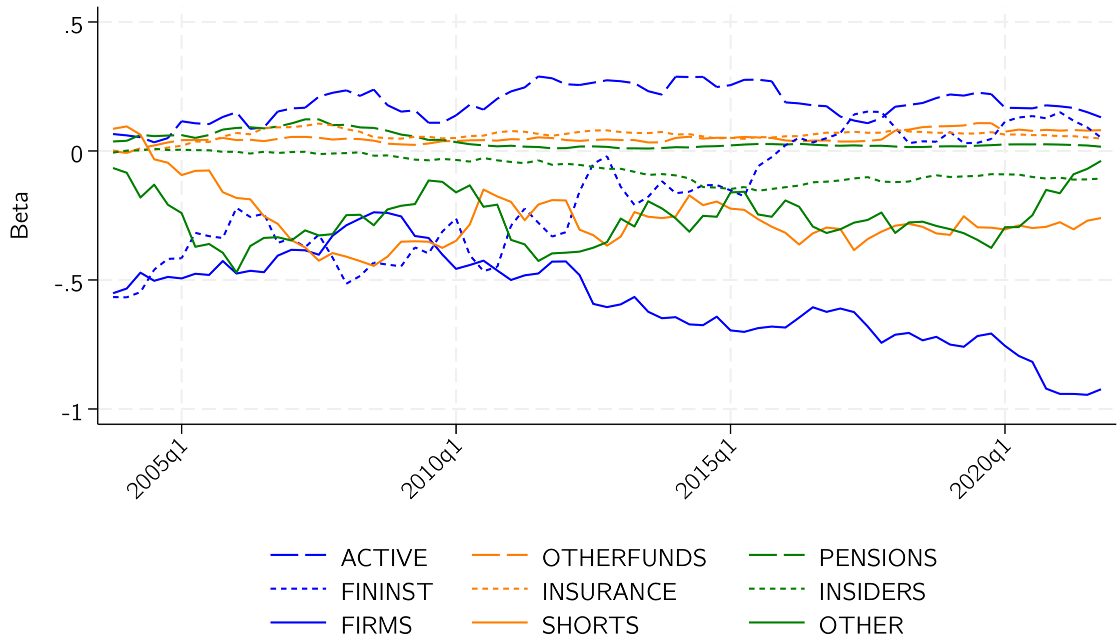
These facts collectively support two non-mutually exclusive themes. First, as passive has continued to grow over time, Firms have continued to take a larger and larger role in making shares available as Index Funds demand more shares, while other institutions and mutual funds have had little to no change in their responsiveness. Second, as the quality of our data has improved, one of our main conclusions, that share suppliers (Firms and Short Sellers) are more responsive in providing shares when Index Funds demand more, appears to gain more support.

## B.9 Broad-Based vs. Style Funds

The largest individual index funds by AUM are broad-based funds that track a large swath of stocks across a range of industries and characteristics like Vanguard's Total Market Fund, VTI. But especially over the last twenty years, there has been a significant growth in what we call style index funds, which track indices based on specific industries and factors, ranging from a "value" version of the S&P 500 (IVE) to the pharmaceutical industry (e.g., PPH) and stocks with high free cash flow yield (e.g., COWZ). Figure 11 shows that these style index funds are collectively large – with similar total AUM to the broad-based funds – although they are individually much smaller, as there more than 20 times as many style funds as there are broad-based funds.

Table 17 presents the regression estimates separately with demand only from broad-based Index Funds as the focal group and again with style-based Index Funds. In each set of regressions, we also include the omitted Index Fund group as a separate investor group $j$. The estimates show that Firms have the most

**Figure 10:** Betas Estimates over Time




*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. This regression is run separately each quarter, and the lines represent an 8-quarter moving average of the betas estimated each quarter.

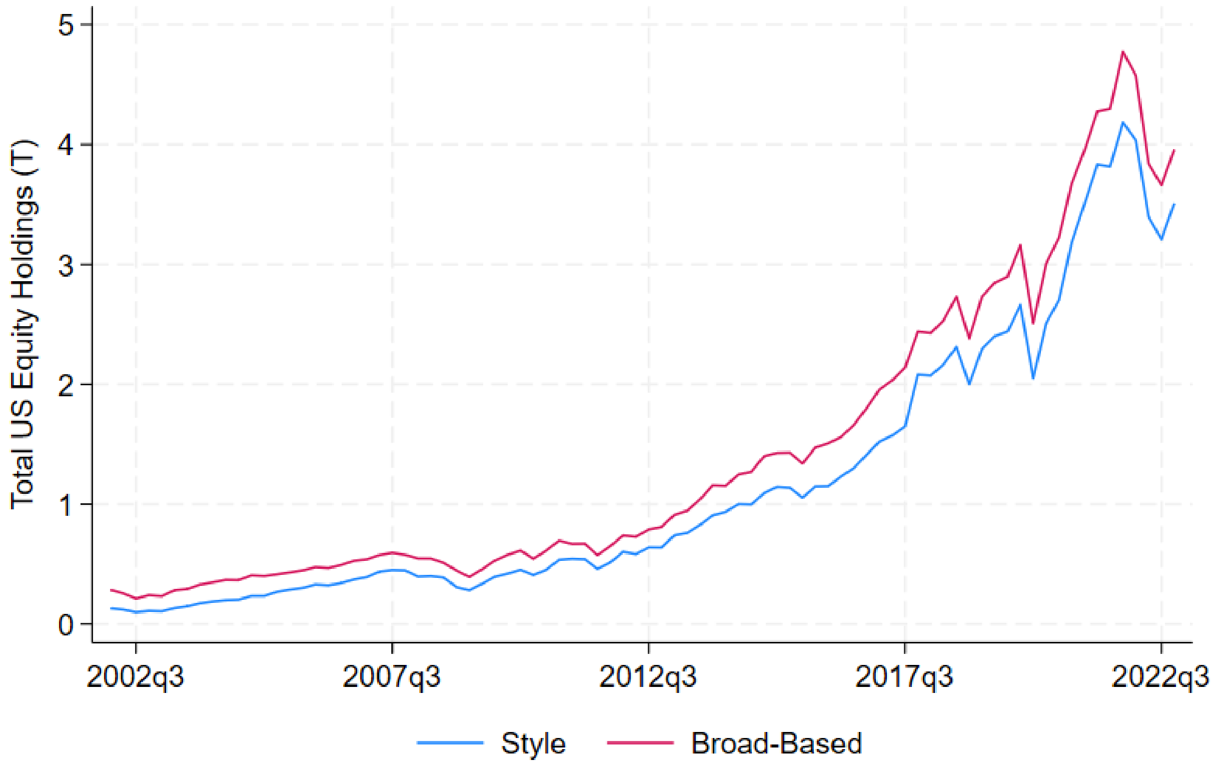
negative beta regardless of whether we focus on broad-based or style Index Fund demand. Firms are much more responsive to broad-based than style Index Funds, partially because there is "more to clear" in the sense that style funds tend to mimic the demand of broad-based funds. That is, when broad-based funds buy, style funds tend to buy more of the same stock and Firms must accommodate not only the broad-based fund demand but also the style fund demand that typically accompanies it.

## B.10 Index Effects

There is a long literature that studies index additions and deletions as a source of inelsatic demand shocks that originates from Shleifer (1986) and Harris and Gurel (1986). In that spirit, we test whether there are differences in who trades with passive when stocks have switched indices (i.e., "switchers") or have not had any change in any major index it belongs to (i.e., "stayers").

Collectively, the tests point to a story where intermediaries can facilitate trading between Index Funds and other institutions if there is a clear signal or event where passive will demand shares. This is similar to

**Figure 11:** Total Assets Managed by Broad-Based Index Funds and Style Index Funds




*Notes.* Total dollars of US equity holdings by Style and Broad-Based passive funds.

the role intermediation might play in facilitating the trades required from index changes, as documented in Chinco and Sammon (2023). Without this signal for intermediaries, the supply side plays the largest role in clearing the market.

#### B.10.1 Index Switchers vs. Stayers

We examine stocks that did or did not switch membership in major indices. Each quarter, we define a stock as an index switcher ("switcher") if it moved into or moved between any of the major indices (e.g, S&P 500, 400, 600 and the Russell 1000 and 2000; see Section A.1 for the full list). Otherwise, the stock is categorized as a "stayer." We split the sample based on this designation.[23] Table 18 presents the estimates for each subsample.

The index stayers tell a largely similar story to the overarching theme of the paper: Firms account for most

[23]Relative to the size of all stocks held by all indices in our data, switching is rare. We still think that this is an important exercise given the salience of index switching and the attention it receives from investors and academics.

**Table 17:** Beta Estimates by Index Fund Type

| | Broad-Based Index Funds | | | | Style Index Funds | | | |
|---|---|---|---|---|---|---|---|---|
| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | $t(\alpha_j)$ |
| Active Funds | 0.364 | 6.051 | -0.039 | -1.493 | 0.241 | 5.106 | -0.032 | -1.430 |
| Other Funds | 0.139 | 3.673 | 0.067 | 3.137 | 0.122 | 2.316 | 0.064 | 3.798 |
| Pension Funds | 0.042 | 4.432 | -0.006 | -0.969 | 0.025 | 6.031 | -0.005 | -0.851 |
| Insurance | 0.113 | 5.863 | -0.023 | -1.805 | 0.084 | 5.693 | -0.023 | -1.825 |
| Financial Institutions | 0.017 | 0.089 | 0.064 | 0.742 | -0.305 | -3.214 | 0.122 | 1.407 |
| Insiders | -0.197 | -9.534 | -0.033 | -4.479 | -0.104 | -8.434 | -0.041 | -5.671 |
| Other | -0.201 | -1.505 | 0.147 | 1.913 | -0.249 | -3.368 | 0.165 | 2.176 |
| Short Sellers | -0.242 | -1.466 | 0.002 | 0.051 | -0.309 | -9.470 | 0.025 | 0.518 |
| Firms | -1.532 | -11.956 | -0.292 | -7.521 | -0.704 | -15.208 | -0.377 | -10.523 |
| Index Funds (Broad) | | | | | 0.199 | 14.063 | 0.103 | 8.917 |
| Index Funds (Style) | 0.498 | 15.340 | 0.112 | 8.564 | | | | |
| Total | -0.999 | | -0.001 | | -0.999 | | 0.000 | |

Notes. The table provides estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t},$$

for each investor group $j$. We separately estimate the regression where $q_{i,\mathrm{IDX},t}$ is the ownership change for broad-based Index Funds (e.g., S&P 500) and style Index Funds (e.g., value funds and industry funds). $q_{i,j,t}$ is the quarterly holdings change in stock $i$ for group $j$ in year-quarter $t$ in units of percent ownership of the company. When testing broad-based Index Funds, we include style Index Funds as another group $j$ (and vice versa). T-statistics are computed from standard errors double clustered at the stock and time (year-quarter) level. See Section 2.2 for more details on the investor groups, Section 2.3 for the empirical methodology.

of the other side of passive (both on average and in terms of their responsiveness), with Short Sellers and Insiders also consistently contributing to clearing the market for Index Funds.

Market clearing is quite different for index switchers. Most notably, Firms play a much larger *average* role: $\overline{q}_j$ for Firms is -1.082pp, nearly all of the 1.102pp average Index Fund demand for index switchers. However, Firms are much less *responsive*, with a beta estimate of -0.201 (as opposed to -0.763 for stayers).[24] This compliments the results in Tamburelli (2024), which shows that firms accommodate inelastic demand from low tracking-error funds around S&P 500 index inclusion events. It is also consistent with the anecdotal high profile cases of equity issuance around S&P 500 index inclusions by e.g., Facebook (Meta) and Tesla.

In order for the market to clear, which groups take a larger role on the other side of increased passive demand? The biggest differences come from Financial Institutions, Active Funds, and Other Funds. In fact, Active Funds and Other Funds appear to be breaking a persistent pattern we have seen throughout the paper: instead of trading in the same direction as Index Funds, they actually accommodate increased Index Fund ownership changes.

[24] One reason for this difference could be large differences within switchers based on whether a stock is coming from outside of a major index to inside a major index (and vice versa) relative to a stock that is switching indices within the same index family. For the former, many of the shares must come from outside of Index Funds. For the latter, many of the shares come from within the Index Fund group see, e.g., Greenwood and Sammon (2022).

**Table 18:** Beta Estimates: Index Switchers vs. Stayers

| | Index Switchers | | | | | | Index Stayers | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Investor Group | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | Obs. | $R^2$ | $\overline{q}_j$ | $\beta_j$ | $t(\beta_j)$ | $\alpha_j$ | Obs. | $R^2$ | $\overline{q}_j$ |
| Active Funds | -0.024 | -0.841 | -0.054 | 1,424 | 0.001 | -0.080 | 0.235 | 7.164 | -0.073 | 134,394 | 0.013 | 0.002 |
| Other Funds | -0.033 | -2.083 | 0.255 | 1,424 | 0.008 | 0.219 | 0.104 | 3.060 | 0.065 | 134,394 | 0.018 | 0.098 |
| Pension Funds | 0.013 | 2.010 | 0.047 | 1,424 | 0.008 | 0.061 | 0.017 | 4.795 | -0.005 | 134,394 | 0.003 | 0.000 |
| Insurance | 0.089 | 8.529 | 0.017 | 1,424 | 0.094 | 0.115 | 0.064 | 6.775 | -0.024 | 134,394 | 0.011 | -0.004 |
| Financial Institutions | -0.147 | -1.897 | 0.728 | 1,424 | 0.008 | 0.566 | -0.068 | -0.926 | 0.103 | 134,394 | 0.000 | 0.081 |
| Insiders | 0.001 | 0.123 | -0.108 | 1,424 | 0.000 | -0.107 | -0.114 | -10.580 | -0.035 | 134,394 | 0.007 | -0.071 |
| Other | -0.328 | -4.413 | 0.437 | 1,424 | 0.030 | 0.075 | -0.249 | -3.759 | 0.341 | 134,394 | 0.003 | 0.262 |
| Short Sellers | -0.370 | -7.681 | -0.461 | 1,424 | 0.146 | -0.869 | -0.225 | -8.835 | 0.042 | 134,394 | 0.015 | -0.030 |
| Firms | -0.201 | -2.609 | -0.860 | 1,424 | 0.015 | -1.082 | -0.763 | -16.627 | -0.414 | 134,394 | 0.039 | -0.657 |
| Total | -1.000 | | 0.001 | | | -1.102 | -0.999 | | -0.000 | | | -0.319 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$. The sample is split into index switchers – defined as those that are added to, dropped from or switch between the Russell 1000, Russel 2000, S&P 500, S&P 400, S&P 600, Nasdaq 100 or the CRSP Total Market. The unit of observation is security-year-quarter. T-statistics are computed based on standard errors which are double clustered at the stock and time (year-quarter) level.

This points to a story where a salient event, like a stock switching indices, may help facilitate the transfer of shares amongst funds and institutional investors. Financial intermediaries may be able to use the attention these stocks garner to get demand-side institutions to adjust portfolio allocations to clear the market (Greenwood and Sammon, 2022). Firms also provide an important role in market clearing by providing a significant fraction of the shares on average, but are much less responsive to deviations from the average Index Fund demand than in the index stayers sample (and, for that matter, in most other subsamples).

In Appendix B.4, we repeat the analysis above but for year-over-year changes. The year-over-year estimates show that the lack of Firm responsiveness for switchers is largely temporary – over the course of a year, Firms are the most sensitive in providing shares to Index Funds, with a beta estimate of -0.701. In addition, the negative beta estimates for Active Funds and Financial Institutions observed in the quarterly regressions flip to positive in the year-over-year regressions, reinforcing the role that these groups may play in short-term intermediation, especially when a salient event, like switching an index, draws attention to possible uninformed trading by indexers.

### B.10.2 Index Migrations and Direct Additions/Deletions

We further study index switchers in light of how different market clearing looks relative to the baseline results. In addition, better understanding these stocks helps provide greater context to index switching events, which are the focus of a long literature. We separate index switchers into several categories. First, we examine direct additions, which are stocks that moved from outside of an index fund family to within it.

For example, a stock that was added to the S&P 500 that was not previously held by the S&P 500, 400, or 600 would be a direct addition. Similarly, we also separately examine direct deletions.

Second, we examine stocks that migrated from one index to another related index (e.g., Russell 1000 to 2000). These events typically have a much smaller change in overall (net) Index Fund ownership because many of the shares that need to be sold or bought are exchanged between Index Funds. That is, in these cases, the market clears for a non-trivial fraction (and, in many cases, the majority) of shares within the Index Fund group. We separately study migrations that led to net buying or net selling by Index Funds. The alpha, beta and $\overline{q}_j$ estimates for our series of regressions for each subsample are presented in Table 19. The table also shows the number of observations for each sample to get a sense of how rare these events are, and also presents the average ownership change for Index Funds and for all institutions (in percentage points) to get a sense of magnitudes for each type of event.

Panel A of Table 19 focuses on direct additions and deletions – showing that who clears the market in each of these cases is quite different. As in our baseline results, for additions, Firms and Short Sellers tend to be the most responsive in supplying shares (this time with Short Sellers supplying relatively *more* shares for each additional unit of Index Fund demand than Firms), with Financial Institutions also playing an important role. For deletions, Short Sellers are about as active in clearing the market (by reducing their short positions) and Firms are not responsive at all. Instead, Other is the most responsive. This is consistent with these stocks leaving the investable universe and finding buyers in retail investors, as well as small institutional investors who may have fewer mandates preventing them from buying such stocks.

However, the betas alone undersell the role of Firms and Short Sellers in accommodating passive demand around index additions. The alphas for these groups are large, and as a result, the associated $\overline{q}_j$s are large as well. Further, the alpha for Financial Institutions is large and positive, evidence that despite a negative beta, this group is predicted to be a net buyer around the typical index addition event, in a way that is not correlated to the size of the demand shock by Index Funds. Similarly, for deletions, we see that the alpha for Other is large, further evidence that foreign institutions, small institutions and retail investors are needed to clear the market when a firm leaves the investable universe and is unlikely to be held by most U.S. institutions.

Panel B of Table 19 focuses on migrations. In the cases of migrations where Index Funds are net buyers of shares (e.g., firms switching from the Russell 1000 to the Russell 2000), Financial Institutions are the most responsive, with a beta of -0.51. Short Sellers and Active are the next most responsive groups, with betas of -0.30 and -0.20, respectively. While Firms have a positive beta, they have a large negative alpha, and therefore a moderately negative $\overline{q}_j$, suggesting that they do play an important role in clearing passive demand around migration events with net passive buying, however more through what they tend to do on average, rather than how they respond to the amount of passive demand.

When Index Funds are net sellers of shares around migrations (e.g., the case of firms migrating from the Russell 2000 to the Russell 1000), Financial Institutions and Firms are the most responsive. That being said, around such events, the alpha for Firms is large and negative, leading to a negative $\overline{q}_j$ for this group. In fact, the $\overline{q}_j$s suggest that our Other group, as well as Financial Institutions and non-passive mutual funds, tend to be the most important groups for clearing the market on average in this subsample.

**Table 19:** Beta Estimates: Index Additions, Deletions, and Migrations

| | Panel A: Direct Adds/Drops | | | | | |
|---|---|---|---|---|---|---|
| | Betas | | Alphas | | $\overline{q}_j$ | |
| Investor Group | Additions | Deletions | Additions | Deletions | Additions | Deletions |
| Active Funds | 0.039 | 0.290 | 0.108 | -0.307 | 0.200 | -0.552 |
| Other Funds | 0.075 | -0.059 | 0.036 | 0.037 | 0.213 | 0.087 |
| Pension Funds | 0.035 | 0.040 | 0.080 | -0.044 | 0.162 | -0.078 |
| Insurance | 0.074 | 0.087 | 0.144 | -0.187 | 0.319 | -0.261 |
| Financial Institutions | -0.201 | -0.098 | 1.552 | -0.792 | 1.079 | -0.709 |
| Insiders | -0.044 | -0.106 | -0.056 | -0.225 | -0.159 | -0.135 |
| Other | -0.385 | -0.980 | 0.307 | 2.046 | -0.601 | 2.876 |
| Short Sellers | -0.374 | -0.215 | -0.706 | -0.480 | -1.587 | -0.298 |
| Firms | -0.220 | 0.040 | -1.465 | -0.051 | -1.984 | -0.084 |
| Total | -1.000 | -1.000 | 0.000 | 0.000 | -2.359 | 0.847 |
| # Obs. | | | | | 1,739 | 103 |
| Avg. Index Fund Ownership Chg. (pp) | | | | | 2.359 | -0.847 |
| Avg. Institutions Ownership Chg. (pp) | | | | | 4.331 | -2.359 |

| | Panel B: Migrations | | | | | |
|---|---|---|---|---|---|---|
| | Betas | | Alphas | | $\overline{q}_j$ | |
| Investor Group | Net Buying | Net Selling | Net Buying | Net Selling | Net Buying | Net Selling |
| Active Funds | -0.204 | 0.085 | 0.167 | 0.286 | -0.261 | 0.143 |
| Other Funds | -0.042 | -0.075 | 0.098 | 0.312 | 0.010 | 0.438 |
| Pension Funds | -0.005 | -0.027 | -0.100 | 0.048 | -0.110 | 0.093 |
| Insurance | 0.081 | 0.028 | -0.012 | -0.065 | 0.158 | -0.112 |
| Financial Institutions | -0.507 | -0.354 | 1.222 | 0.190 | 0.160 | 0.784 |
| Insiders | 0.015 | -0.015 | -0.155 | -0.193 | -0.123 | -0.168 |
| Other | -0.202 | -0.307 | 0.005 | 0.555 | -0.417 | 1.071 |
| Short Sellers | -0.297 | -0.033 | -0.363 | 0.097 | -0.986 | 0.153 |
| Firms | 0.162 | -0.301 | -0.863 | -1.231 | -0.524 | -0.726 |
| Total | -1.000 | -1.000 | 0.000 | 0.000 | -2.094 | 1.678 |
| # Obs. | | | | | 503 | 515 |
| Avg. Index Fund Ownership Chg. (pp) | | | | | 2.094 | -1.678 |
| Avg. Institutions Ownership Chg. (pp) | | | | | 0.846 | 0.846 |

*Notes.* Estimates from our baseline regression specification:

$$q_{i,j,t} = \alpha_j + \beta_j \cdot q_{i,\mathrm{IDX},t} + \varepsilon_{i,j,t}$$

for each investor group $j$ estimated on each subsample. In Panel A, the sample is split into direct additions and deletions. Direct additions are defined as securities that are added to the S&P 500, S&P 400, S&P 600 from outside the S&P 1500 universe, securities added to the Russell 1000 or 2000 from outside the Russell 3000 universe and securities which are added to the Nasdaq 100 or CRSP Total Market Indices. Direct deletions are defined similarly, as firms that leave these indices to outside of their respective index universes. In Panel B, the sample includes only migrations, and is split based on whether there is net buying or net selling by passive funds. An example of a migration typically associated with net buying is a migration from the Russell 1000 to the Russell 2000, while an example of a migration typically associated with net selling is a migration from the Russell 2000 to the Russell 1000. The unit of observation is security-year-quarter.

# C Mechanism and Identification Appendix

## C.1 Instrumental Variable and Market Clearing

In the body of the paper, we focused on Firms' demand as the key outcome variable of interest in our IV regression. In Table 20, we repeat the analysis in Table 2, but for all our investor groups. We omit the first stage regression results here, as it would be the same for each group (see Section 4.1 for the first stage estimates). Broadly, our takeaway from Table 20 is that the results from the OLS regressions in the main body of the paper are mostly qualitatively unchanged in the IV setting. Firms are still the most responsive to Index Fund demand, as they have the most negative IV estimate of any group. Active Funds, Insurance, and Pensions still have positive coefficients, i.e., trade in the same direction as Index Fund demand.

Two notable differences with the IV results from the OLS results are Short Sellers and Other. In Table 20, the coefficient on Short Sellers is positive and insignificant, while in the baseline OLS specification the coefficient is negative. One explanation for this change is that our IV identifies "non-fundamental" demand shocks, and Short Sellers, as smart money, focus on betting against fundamental shocks (Hanson and Sunderam, 2014). An additional change is that the coefficient for Other has become much larger in magnitude, going from -0.25 to -0.85. One possible explanation for this change is that clearing the market for non-fundamental demand shocks relies more heavily on the Other category. However, we must highlight that with the IV, there is no observation-by-observation enforcement of market clearing, as the instrumented Index Fund demand is not used to compute the Other category. Therefore, residuals in the first stage may be correlated with our Other category – explaining the larger (in magnitude) coefficient in the IV setting.

## C.2 IV with only past returns

As mentioned in the main body of the paper, one may be concerned about our use of both lagged ($t-1$) and contemporaneous ($t$) returns to predict Index Fund demand in quarter $t$. One possible problem is that the "exogenous" (i.e., instrumented) Index Fund demand causes high returns in the focal stock $i$ at time $t$, which leads to flows into the funds $k \in K$ which hold stock $i$. Then, it is possible that these flows (caused by the flow-based price pressure on the focal stock) put price pressure on the unrelated co-holdings and lead to even more flows into funds $k \in K$. In this scenario, the instrument may be partly picking up reverse causality, violating the exclusion restriction in our IV setup.

**Table 20:** Instrumental Variables Specification – All Investor Groups

| | Act. Fnds. (1) | Oth. Fnds. (2) | Pens. (3) | Ins. (4) | Fin Inst. (5) | Insiders (6) | Others (7) | Shorts (8) | Firms (9) |
|---|---|---|---|---|---|---|---|---|---|
| $q_{i,IDX,t}$ | 0.477*** | 0.0989 | 0.0479** | 0.108** | -0.0627 | -0.213*** | -0.835*** | 0.243 | -0.864** |
| | (0.123) | (0.113) | (0.023) | (0.045) | (0.412) | (0.055) | (0.311) | (0.205) | (0.389) |
| $SUE_{i,t}$ | 0.0130*** | 0.00142** | 0.000805* | 0.00571*** | 0.0240*** | -0.00354** | -0.0550*** | 0.003 | 0.0105** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.006) | (0.001) | (0.014) | (0.006) | (0.005) |
| $SUE_{i,t-1}$ | 0.00700** | 0.00179*** | 0.000 | 0.000 | 0.002 | -0.00356*** | -0.019 | -0.003 | 0.0145** |
| | (0.003) | (0.001) | (0.000) | (0.000) | (0.006) | (0.001) | (0.011) | (0.002) | (0.006) |
| Observations | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 |
| R-squared | -0.017 | 0 | -0.004 | 0.004 | 0.001 | -0.012 | -0.014 | -0.042 | 0.026 |
| Fixed Effects | FF 10 by YQ | | | | | | | | |

*Notes.* The table provides estimates from the second stage of our IV regression:

$$q_{i,j,t} = \beta \cdot \hat{q}_{i,IDX,t} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + \varepsilon_{i,j,t},$$

where $\hat{q}_{i,IDX,t}$ is the instrumented Index Fund demand based on weighted average leave-out co-holdings returns in quarters $t-1$ and $t$. $SUE_{i,t}$ is the standardized unexpected earnings (SUE) of stock $i$ in quarter $t$. FE are a set of Fama-French 10 industry-by-year-quarter fixed effects. Standard errors are double-clustered at the stock and time (year-quarter) level.

To construct a version of the IV insulated from this reverse causality problem, we replicate the results in Table 2, but modify the first stage regression to only use "leave-out" co-holdings returns at $t-1$. Even if these lagged returns lead to price pressure on the focal stock $i$ in quarter $t$, which in turn feeds back into the unrelated co-holdings' returns in $t$, the IV will not use this variation to predict Index Fund demand in stock $i$ in quarter $t$. We still control for SUE at $t$ because it is not clear how flows would have a causal effect on contemporaneous SUE (i.e., how the reverse causality would be relevant for the focal firm's *fundamentals*). Before discussing the results, we want to highlight that this is a much higher bar than our original IV design. As we show in Appendix C.3, the flow-performance relationship is significant for both current and past returns. Omitting contemporaneous returns, therefore, introduces noise, weakening the instrument.

With that said, we report the results of this alternative IV design in Table 21. For ease of comparison, the left panel replicates the results in Table 2. The right panel contains the results when omitting the contemporaneous leave-out co-holdings returns. By omitting the contemporaneous co-holdings returns, the first stage becomes weaker, with the F-statistic dropping from 23.3 to 11.7. However, it remains above the standard threshold of 10 for weak instruments. The estimated IV coefficient becomes larger in magnitude but less statistically significant. The increase in the magnitude of the IV coefficient is likely due to the instrument becoming weaker (i.e., because the denominator in $X'Y/X'Z$ is closer to zero). However, the IV estimate remains marginally significant. Finally, the reduced form still retains marginal significance and predicts Firm demand with the expected sign. In all, Table 21 shows our IV results are robust to this higher bar of using only lagged leave-out co-holdings returns to predict Index Fund demand.

**Table 21:** Instrumental Variables Specification: Robustness to Using Only Past Returns

| | Baseline IV specification | | | IV using only coholdings returns at $t-1$ | | |
|---|---|---|---|---|---|---|
| | First Stage | IV | RF | First Stage | IV | RF |
| $\overline{r}^{coholdings}_{i,t-1}$ | 1.045*** | | -1.433* | 1.034*** | | -1.427* |
| | (0.253) | | (0.842) | (0.303) | | (0.854) |
| $\overline{r}^{coholdings}_{i,t}$ | 1.567*** | | -1.004 | | | |
| | (0.330) | | (0.758) | | | |
| $SUE_{i,t}$ | 0.00463*** | 0.0105** | 0.006 | 0.00477*** | 0.0129** | 0.006 |
| | (0.001) | (0.005) | (0.005) | (0.001) | (0.005) | (0.005) |
| $SUE_{i,t-1}$ | 0.00373*** | 0.0145** | 0.0112** | 0.00388*** | 0.0164** | 0.0111** |
| | (0.001) | (0.006) | (0.005) | (0.001) | (0.007) | (0.005) |
| $q_{i,IDX,t}$ | | -0.864** | | | -1.379* | |
| | | (0.389) | | | (0.782) | |
| Observations | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 | 130,794 |
| R-squared | 0.1 | 0.026 | 0.033 | 0.096 | -0.013 | 0.033 |
| F-statistic | 23.26 | | | 11.66 | | |
| Fixed Effects | | FF 10 by YQ | | | FF 10 by YQ | |

*Notes.* The table provides our baseline IV estimates, as well as estimates from the first and second stages of our IV regression using only past returns, along with the associated reduced form regressions:

$$q_{i,IDX,t} = \gamma_{t-1} \cdot \overline{r}^{coholdings}_{i,t-1} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + e_{i,t},$$

$$q_{i,Firm,t} = \beta \cdot \hat{q}_{i,IDX,t} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + \varepsilon_{i,Firm,t},$$

where $\overline{r}^{coholdings}_{i,t-1}$ is the weighted average leave-out co-holdings returns across all funds $k \in K$ that held stock $i$ at the end of quarter $t$. $SUE_{i,t}$ is the standardized unexpected earnings (SUE) of stock $i$ in quarter $t$. FE are a set of Fama-French 10 industry-by-year-quarter fixed effects. Standard errors are double-clustered at the stock and time (year-quarter) level. In each panel, the first column reports the first stage regression, and the last row reports the associated F-statistic. The second column reports the IV specification, while the third column reports the reduced form regression. The left panel replicates our baseline IV results, while the right panel uses only past leave-out co-holdings returns.

## C.3 Flow-Performance Relationship

As highlighted in Section 4.1, our baseline IV essentially takes the flow performance relationship as given. In this subsection, we aim to refine that result by explicitly estimating the flow performance relationship for each fund. To this end, we run the following regression separately for each fund $k$:

$$flow_{k,t} = \alpha_k + \beta_{1,k} r_{k,t} + \beta_{2,k} r_{k,t-1} + e_{k,t}, \tag{20}$$

where $flow_{k,t}$ is the percentage flow – i.e., flow as a percentage of lagged AUM (Barber et al., 2016) – into fund $k$ between the end of quarter $t-1$ and the end of quarter $t$. $r_{k,t}$ and $r_{k,t-1}$ are the quarter $t$ and quarter $t-1$ returns for fund $k$.

Empirically, there is significant variation in $\beta_{1,k}$ and $\beta_{2,k}$ across funds. To visualize this, in Figure 12 we split the sample by "pure" (or broad-based) and "systematic" (or style) funds, and then further split each sample into 20 groups based on the number of holdings in each fund. The data points on the left represent funds with a small number of holdings, while the data points on the right represent funds with a large number of holdings. Pure or broad-based index funds are defined as the broad capitalization-based funds for the S&P 500, 400, 600, the Russell 1000, 2000, and the CRSP-Capitalization-Based indices used by Vanguard. Systematic or style funds, include, e.g., sector and factor ETFs.

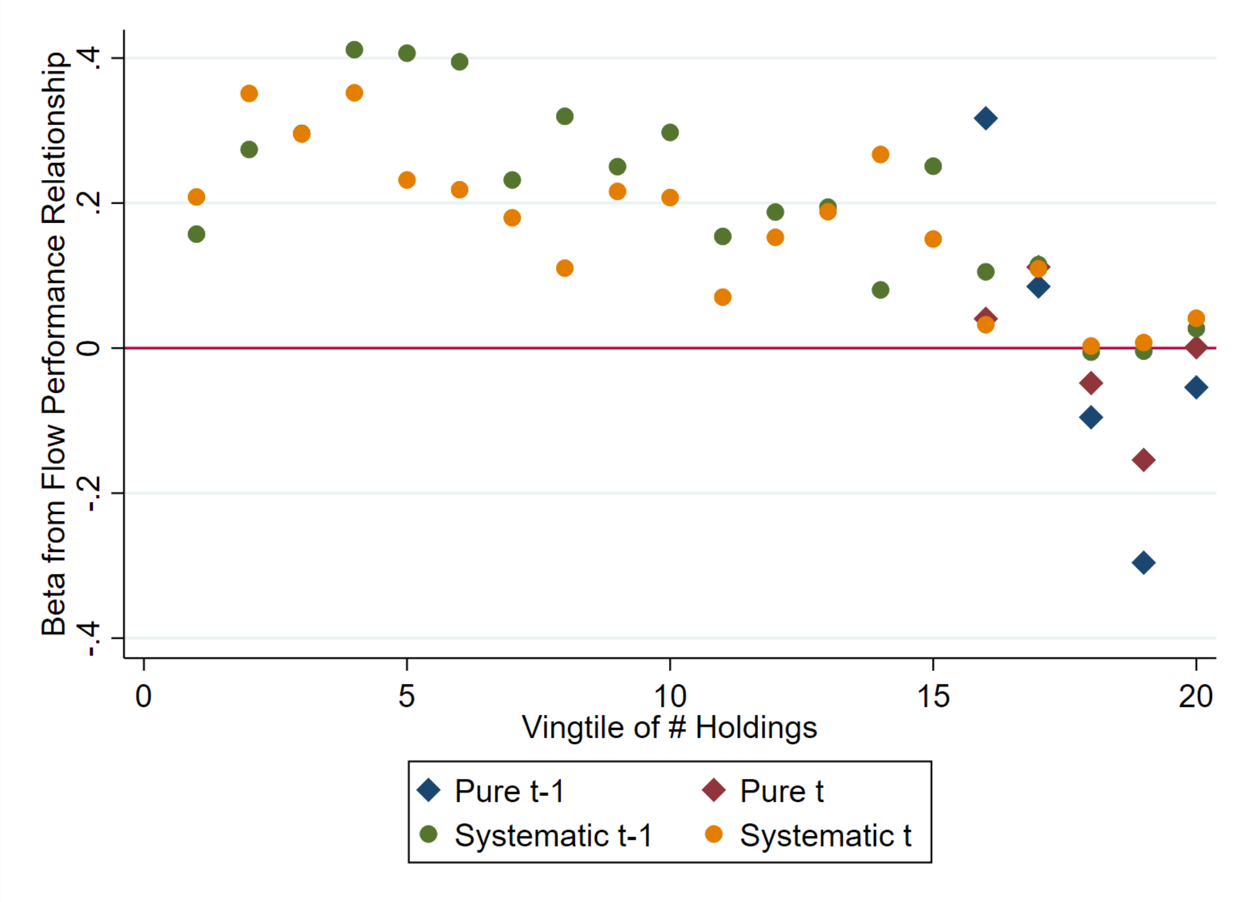
Three important patterns emerge in Figure 12. First, the flow performance relationship is stronger in funds with fewer holdings. Second, for pure capitalization-based funds, the flow performance relationship is non-existent or even negative. This latter fact may at first seem bizarre, but, as a reminder, these are *slopes*. The growth of, e.g., VTI may come through regular employee 401K contributions which are unrelated to VTI's performance, and therefore may show up in the $\alpha_{VTI}$ of Equation 20. Finally, $\beta_{1,k}$ and $\beta_{2,k}$ are of similar magnitudes, suggesting that both last quarters' and this quarters' performance are relevant for predicting flows.

We then use $\beta_{1,k}$ and $\beta_{2,k}$, along with the leave-out co-holdings returns, to predict the part of flows coming from the performance of stock $i$'s unrelated co-holdings. Specifically, for stock $i$ in quarter $t$ for fund $k$, the predicted flow is

$$\widehat{flow}_{i,k,(t-1,t)} = \beta_{1,k} \cdot r^{coholdings}_{i,k,t} + \beta_{2,k} \cdot r^{coholdings}_{i,k,t-1}.$$

Note, that even though each fund only has one estimated flow performance relationship ($\beta_{1,k}$ and $\beta_{2,k}$), predicted flows $flow_{i,\hat{k},(t-1,t)}$ for each focal stock $i$ and quarter $t$ will be different because it will be estimated

**Figure 12:** Flow Performance Relationship by Vigintile of Holdings




*Notes.* The figure shows the estimated flow performance relationship ($\beta_{1,k}$ and $\beta_{2,k}$) by vigintile of fund holdings. Vigintile 1 represents funds with fewer holdings, while the vigintile 20 represents funds with the largest number of holdings.

with different leave-out co-holdings returns (which depend on a stock and its characteristics and industry).

We then use these estimated flows to predict trading by each fund – which, given that value-weighted index funds scale up and down positions in proportion to flows, should be $\widehat{flow}_{k,(t-1,t)} \cdot \text{Shares Held}_{i,k,t-1}$. Finally, we aggregate across all funds $k \in K$ that held stock $i$ at the end of quarter $t$, and normalize by lagged shares outstanding to make the units consistent with $q_{i,IDX,t}$:

$$\tilde{q}_{i,t} = \sum_{k \in K} \widehat{flow}_{i,k,(t-1,t)} \cdot \text{Shares Held}_{i,k,t-1} / \text{Shares Out.}_{i,t-1} \tag{21}$$

To reduce the effect of outliers, we Winsorize $\tilde{q}_{i,t}$ at the 0.5% and 99.5% level. As an additional robustness check – to avoid the possible reverse causality issue discussed in Section C.2 – we also construct a version of predicted flows in Equation 20 using only past leave-out co-holdings returns.

Ultimately, we want to run a regression of Firm demand on our predicted measure of Index Fund demand, or

$$q_{i,Firm,t} = \beta \cdot \tilde{q}_{i,IDX,t} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + \varepsilon_{i,Firm,t}. \tag{22}$$

We report the results in Table 22. The left panel uses both co-holdings returns at $t-1$ and $t$ in the flow performance relationship. The first column reports the first stage, showing that our measure of predicted flow-based trading is strongly correlated with actual Index Fund demand. The coefficient of 0.857 implies that for every 1 percentage point of market capitalization we expect to be demanded by Index Funds, they actually demand 86 basis points. Also note that the F-statistic is larger than in Table 2 – suggesting that leveraging the observed flow performance relationship allows us to more accurately predict Index Fund demand.

In this setting, we don't use a traditional two-stage least-squares estimate, as our measure of expected index fund demand is already a prediction using our co-holdings returns as a type of instrument, and is in the same units as $q_{i,IDX,t}$. In this sense, Equation 22 is our "second-stage IV" estimate but looks like a reduced form regression – in that we are directly regress Firm demand on our measure of predicted Index Fund demand – which is why we label these columns RF. These results are in the second column. We find a coefficient of -0.88, which is very similar in magnitude to our OLS regression. This confirms the results in 2, and further allays concerns of the IV estimate being distorted by the difference in units between returns and Index Fund demand.

Finally, in the right panel, we repeat this exercise but only estimate the flow performance relationship using each fund's $t-1$ returns, and then predict flows given leave-out co-holdings returns at $t-1$. Unsurprisingly, we find a weaker first stage, but again, this "alternative" IV-style regression is quantitatively unchanged. This further suggests that reverse causality in quarter $t$ is not driving the strength of our IV results.

## C.4 Ruling Out Mechanical Stories

As discussed in Section 4.3, one concern with our OLS results is reverse causality due to the way Index Funds mechanically must rebalance in response to Firm issuance and buyback activity. We have already presented several pieces of evidence against this story, including:

1. The IV results in Section 4.1, which leverage the *flow*-based part of Index Fund demand and are not designed to pick up the mechanical part of Index Fund demand in response to Firm activity, i.e., the source of the reverse causality problem.

**Table 22:** Predicting Index Fund Flows Using Fund-Level Flow Performance Relationship

| | $t$ and $t-1$ Returns | | $t-1$ Returns | |
|---|---|---|---|---|
| | First Stage | "IV" | First Stage | "IV" |
| $\tilde{q}_{i,IDX,t}$ | 0.857*** | -0.883*** | 0.875*** | -0.852** |
| | (0.123) | (0.247) | (0.265) | (0.357) |
| $SUE_{i,t}$ | 0.00449*** | 0.00669 | 0.00447*** | 0.0067 |
| | (0.001) | (0.005) | (0.001) | (0.005) |
| $SUE_{i,t-1}$ | 0.00382*** | 0.0112** | 0.00386*** | 0.0111** |
| | (0.001) | (0.005) | (0.001) | (0.005) |
| Observations | 130,794 | 130,794 | 130,794 | 130,794 |
| R-squared | 0.099 | 0.033 | 0.096 | 0.033 |
| F-statistic | 48.26 | | 10.88 | |
| Fixed Effects | FF 10 by YQ | | FF 10 by YQ | |

*Notes.* The table provides estimates from the first stage and reduced form regressions using predicted passive flows based on fund-level flow performance relationships:

$$q_{i,Firm,t} = \beta \cdot \tilde{q}_{i,IDX,t} + \psi_1 \cdot SUE_{i,t} + \psi_2 \cdot SUE_{i,t-1} + \text{FE} + \varepsilon_{i,Firm,t},$$

where $\tilde{q}_{i,IDX,t}$ is the predicted Index Fund demand based on the estimated flow performance relationships and leave-out co-holdings returns. $SUE_{i,t}$ is the standardized unexpected earnings (SUE) of stock $i$ in quarter $t$. FE are a set of Fama-French 10 industry-by-year-quarter fixed effects. Standard errors are double-clustered at the stock and time (year-quarter) level. The left panel uses both contemporaneous and lagged co-holdings returns to predict flows, while the right panel uses only lagged co-holdings returns to predict flows.

2. The time series increase (in magnitude) of our estimated $\beta_{Firm}$ in Figure 10, which goes against the expected decrease (in magnitude) in the mechanical/reverse causality effect over time due to the growth of average passive ownership.

3. The asymmetry in Figure 4, as the mechanical effect should be completely symmetric with respect to buybacks and issuance.

In this section, we perform a decomposition exercise designed to separate the piece of Index Fund demand coming from flows and the part of Index Fund demand coming from the mechanical response of Index Funds to Firm activity. Specifically, for each stock $i$ at each time $t$, we identify all the funds $k \in K$ that held stock $i$ at time $t-1$. We then use the *ex-post observed* flows between time $t-1$ and time $t$ multiplied by lagged shares held as a predictor of flow-based Index Fund demand. As in Section C.3, we are leveraging the fact that value-weighted index funds are expected to scale up their holdings proportionally in response to flows. In an equation, we predict Index Fund demand from flows by estimating

$$q_{i,IDX,t}^{flows} = \sum_{k \in K} \text{flow}_{k,t} \times \frac{\text{Shares Held}_{i,k,t-1}}{\text{Shares Out.}_{i,t-1}}. \tag{23}$$

We normalize the flow-based expected demand shock by Shares Out.$_{i,t-1}$ to make the units consistent with

**Table 23:** Decomposition of Index Fund Demand

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $q_{i,IDX,t}$ | -0.668*** | | | |
| | (0.040) | | | |
| $q^{flows}_{i,IDX,t}$ | | -0.210** | | -0.374*** |
| | | (0.098) | | (0.111) |
| $q^{residual}_{i,IDX,t}$ | | | -0.713*** | -0.724*** |
| | | | (0.046) | (0.045) |
| Observations | 109,707 | 109,378 | 109,378 | 109,378 |
| R-squared | 0.047 | 0.001 | 0.046 | 0.048 |

*Notes.* The table provides estimates from our baseline regression specification:

$$q_{i,Firm,t} = \alpha_j + \beta_j \cdot q_{i,\text{IDX},t} + \varepsilon_{i,Firm,t},$$

where we decompose Index Fund demand into a flows-based component:

$$q^{flows}_{i,IDX,t} = \sum_{k \in K} \text{flow}_{k,t} \times \frac{\text{Shares Held}_{i,k,t-1}}{\text{Shares Out.}_{i,t-1}},$$

and a residual component:

$$q^{residual}_{i,IDX,t} = q_{i,IDX,t} - q^{flows}_{i,IDX,t}.$$

Standard errors double clustered at the stock and time (year-quarter) level.

our baseline measure of $q_{i,IDX,t}$. To reduce the effect of outliers, we Winsorize $q^{flows}_{i,IDX,t}$ at the 0.5% and 99.5% level.

The main difference between this and our baseline IV strategy in Section 4.1 is that rather than trying to *predict* flows given performance, we are using *observed* flows. To ensure variation in $q^{flows}_{i,IDX,t}$ and to avoid a large number of zero values, we require that stocks were in the Russell 3000 in the current and previous quarter. The logic is that such stocks have at least some passive ownership, so Shares $\text{Held}_{i,k,t-1}$ by the passive funds will not be zero or close to zero.

In Table 23, we show that, although the baseline OLS effect is attenuated when using our flow-based measure of expected Index Fund demand, our main finding of Firms responding to passive buying by issuing holds up using $q^{flows}_{i,IDX,t}$. Specifically, the coefficient decreases in magnitude from -0.67 to -0.21 or to -0.374, depending on the specification. One explanation for this attenuation is that, as shown in the numerical example above, the mechanical coefficients are an order of magnitude larger than our OLS effect. The attenuation could also, however, be due to noise in $q_{i,IDX,t}$, in the classical measurement error sense. Specifically, $\text{flow}_{k,(t-1,t)}$ is only an estimate of net quarterly flows, as the true flows are not directly observable (Barber et al., 2016). Further, $q^{flows}_{i,IDX,t}$ assumes all rebalancing and re-investment of flows occurs at the end of the quarter, which may not hold in the data.

A final natural question is why we don't perform the decomposition by first computing the expected mechanical trading of Index Funds in response to Firm buyback and issuance activity and then attribute the residual to flows. In the simple numerical example in Section 4.3, it seems like this should be straightforward, but issuance does not happen in a vacuum. For example, an Index Fund may only need to mechanically buy a specific stock if that stock had high issuance relative to the other stocks in the fund.

More broadly, the required rebalancing in the face of heterogeneous issuance/buyback behavior is a fixed-point problem for each fund. Each fund needs to choose the number of shares held of each stock such that they hold a constant percentage of the float of every stock and that their AUM is fully invested (Sammon and Shim, 2023). To fully account for the net mechanical effect, one would need to account for the size of the issuance of the focal stock relative to all other stocks for each fund that holds it. This illustrates one way that the knock-on effects of mechanical rebalancing are hard to predict, as discussed in Chinco and Fos (2021).

# D Stylized Facts on Buyback vs. Issuance

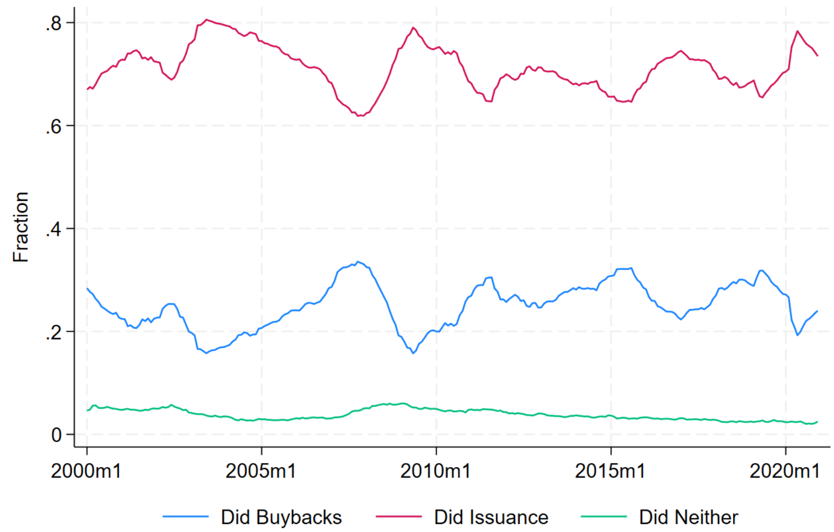
As we show in Sections 3.1 and in Appendix B.3, Firms' responsiveness to Index Fund demand is similar on both an equal-weighted and value-weighted basis. At first glance, this seems hard to square with two well-known trends. The first is the rise of passive ownership, which went from nearly nothing in the early 1990s to owning nearly 17% of the US stock market in 2021 (Investment Company Institute, 2023). The second trend is the substantial increase in the dollar value of buybacks over the past 20 years (see e.g., see the Financial Times, Guru Focus, and New York Life). The apparent conflict is that if, on a value-weighted average basis, Firms have been buying back shares *and* Index Funds have been buying shares, how can Firms' demand be the most responsive to Index Funds on average? Further, it seems difficult to reconcile these trends with our main results that – on an equal-weighted basis – Firms have been the largest overall supplier of shares to Index Funds.

In this appendix, we resolve this conflict by highlighting two important stylized facts. First, in every year in our sample, a significantly larger *number* of firms issue shares than buy back shares. Second, the average size of issuance is larger than the average size of buybacks. Together, these facts imply that the equal-weighted average firm in our sample has actually issued shares, and thus could have been the largest provider of shares to Index Funds who demanded shares on average.

The first natural question we answer is which *fraction* of firms conducts buybacks or issuance each year. In a given month, we classify a firm as doing buybacks over the next year if its split-adjusted shares outstanding has declined 12 months in the future. Similarly, we classify a firm as doing issuance over the next year if its split-adjusted shares outstanding are higher 12 months in the future. Finally, we say that a firm has done neither if split adjusted shares outstanding are constant 12 months in the future. We use a 12-month horizon – instead of, say, a quarterly horizon – to reduce the noise inherent in using possibly stale split-adjusted shares outstanding data in CRSP and to account for seasonalities.

Figure 13 plots the fraction of firms in each of these three categories since 2000. The first salient feature of this figure is that the fraction of firms doing neither buybacks nor issuance has been steadily declining. At the same time, there has been a slight upward trend in the fraction of firms doing buybacks or issuance. Another striking takeaway from this figure is that the fraction of firms doing buybacks is relatively small, hovering between 20 and 30 percent over the last few years. So, while the firms doing buybacks may have been doing more and more in dollar terms, they are still a relatively small fraction of the universe of firms we consider in this paper.

**Figure 13:** Fraction of Firms Doing Buybacks and Issuance



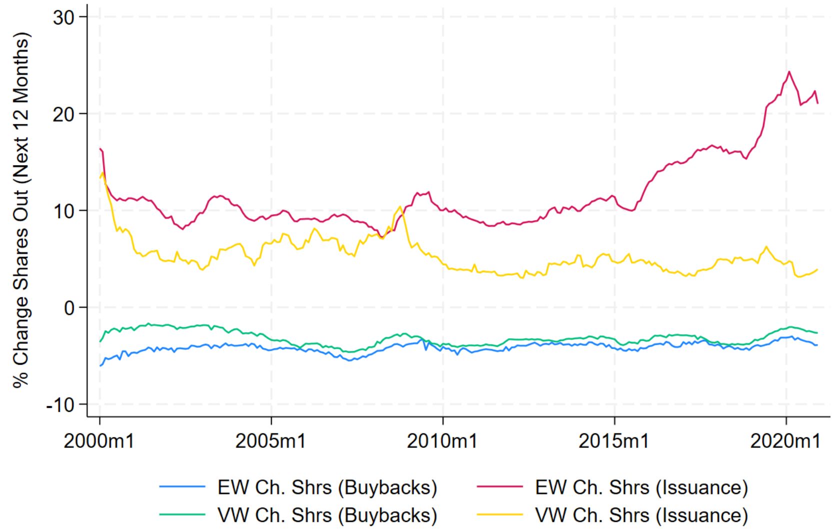

*Notes.* Fraction of firms which, over the next year, will do net buybacks, net issuance, or neither. A firm is classified as having net issued equity if it has a year-over-year increase in split-adjusted shares. A firm is classified as having done a net buyback if it has a year-over-year decrease in split-adjusted shares. A firm is classified as having done neither if there has been no change in split-adjusted shares.

Figure 13 says nothing about the relative magnitudes of these phenomena. Specifically, it could be that each year, many firms are issuing relatively small amounts of equity, while a few firms are buying back a significant amount of equity, so the overall net effect (on a value-weighted basis) is towards fewer shares outstanding. To examine this, we plot the equal-weighted and value-weighted percentage change in shares outstanding for firms that issue or buyback in Figure 14. Perhaps surprisingly, not only is the fraction of firms doing issuance larger than the fraction of firms doing buybacks, the average magnitude of issuance is larger than the average magnitude of buybacks – especially on an equal-weighted average basis.

Putting together the results in Figures 13 and 14, the final natural question we study is whether there is issuance or buybacks *in aggregate*. To make buybacks and issuance comparable across firms, we first redefine these quantities to be in dollar terms. To do this, we need an assumption about the price firms paid for buybacks/received for issuance, so we assume that firms transact at the dollar volume weighted average split-adjusted price over the next 12 months. We then multiply the change in split-adjusted shares outstanding over the next 12 months by this average price to get an estimate of buybacks/issuance in dollars. Finally, we add this up across firms for each month to get a measure of net dollar issuance or buybacks.[25]

[25] One might be concerned that there is some bias in using the average price over the next 12 months to estimate dollar buybacks/issuance. Specifically, one might think that we are overstating net buybacks if firms that do buybacks have stock price increases, and firms that do issuance have stock price decreases. Our methodology, however, yields almost identical estimates for the amount of buybacks by S&P 500 firms in this chart, which is based on data directly from S&P Dow Jones.

**Figure 14:** Magnitudes of Buybacks and Issuance (Relative to Shares Outstanding)

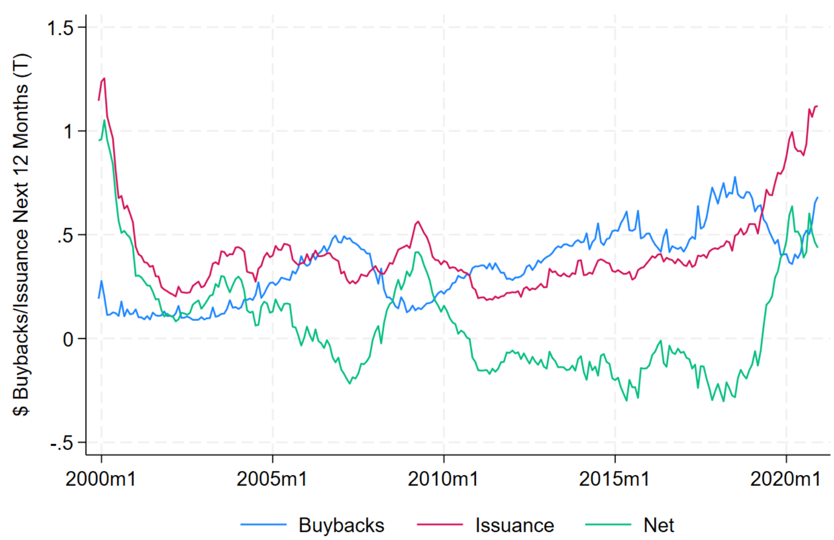



*Notes.* Firms are split into groups based on whether they net issued shares (i.e., had an increase in split-adjusted shares outstanding) or net bought back shares (i.e., had a decrease in split-adjusted shares outstanding) over the next year. For each group, we plot the equal-weighted and value-weighted average percentage change in shares outstanding over the next 12 months. For the value-weighted averages, each firm's weight is proportional to that firms' share of total market capitalization *in the buyback or issuance group* in quarter $t-1$.

Figure 15 plots the total dollar amount for all firms that issued shares, bought back shares, and a signed net (aggregate buybacks minus aggregate issuance). Leading into the Global Financial Crisis, buybacks grew significantly, becoming larger than issuance. This temporarily reversed during and after the Global Financial Crisis, and then switched back to a buyback-heavy regime from around 2011 to 2019. Finally, many firms issued equity during the COVID-19 crisis.

At first glance, Figure 15 might deepen the puzzle outlined at the beginning of this section. Passive ownership grew significantly, i.e., passive funds had to buy a significant fraction of each firm's shares outstanding, and at the same time, there were net aggregate dollar buybacks by firms. These trends still seem hard to reconcile with our main results, which show that firms have provided liquidity to passive ownership. Recall, however, that the bulk of our analysis is effectively equal-weighted. This is one explanation for how, in dollar terms, buybacks may have dominated issuance, but for the (equal-weighted) average firm, issuance has been more prevalent than buybacks. This is supported by Figure 13, which shows that more firms issue than buy back shares, and Figure 14, which shows that the average size of issuance is larger than buybacks.

**Figure 15:** Total Dollar Value of Buybacks and Issuance




*Notes.* Firms are split into groups based on whether they net issued shares (i.e., had an increase in split-adjusted shares outstanding) or net bought back shares (i.e., had a decrease in split-adjusted shares outstanding) over the next year. We assume that firms do buybacks and issuance at the volume-weighted average split-adjusted price over the next year. The lines represent the estimated total dollar value of buybacks and issuance over the next 12 months for each group.