Do Active Funds Do Better in What They Trade?*

Marco Sammon† John J. Shim‡

December 2023

Abstract

We develop two new measures to quantify active fund decisions at the position level. Our measures are designed to separate flow-based passive scaling from active rebalancing decisions. We find that additive active rebalancing – both for existing and new positions – predicts higher stock-level alpha over the following quarter. We show our results are not driven by mechanical price pressure, and provide evidence that funds may trade on forecasts for future earnings. Finally, we aggregate our stock-level measure to the portfolio level and show that actively adding to positions translates to outsized returns for fund investors.

Keywords: Mutual Funds, Active Management

JEL Classification: G11, G23

*We thank Malcolm Baker, Shane Corwin, Martijn Cremers, Roger Edelen, Tim Loughran, Emily McComb, Max Miller, Chris Murray, Andre Perold, Philippe van der Beck, Chen Wang and seminar participants at the Harvard Finance Lunch Workshop, Notre Dame and Virginia Tech for helpful comments.

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

‡Mendoza College of Business, University of Notre Dame; jshim2@nd.edu

1 Introduction

Active mutual funds and ETFs, who manage portfolios with the aim of delivering outsized returns to their investors, manage around $14 trillion in U.S. equities as of 2022. Whether these managers have skill in selecting stocks with high future alpha is a long-running debate, with mixed evidence.[1] While performance is one way to identify skill, another strand of literature has aimed to identify skill by quantifying the "activeness" of managers at the *fund-quarter* level. For example, Cremers and Petajisto (2009) and Pástor et al. (2017) show that managers whose portfolio weights deviate more from their benchmark's weights and managers with more portfolio turnover tend to outperform their benchmarks. Because these measures are only defined at the fund-quarter level, however, it is not obvious whether the activeness *itself* is generating the outsized returns. It could be, for example, that when managers churn their portfolio, the positions they leave untouched are the ones that outperform.

In this paper, we examine whether individual position changes of active funds are informative about future returns *in those stocks*. That is, if a fund increases its position in a stock this quarter, does that stock have higher or lower returns next quarter?

One concern with position changes is that they are not solely determined by considerations about future returns. Perhaps the most significant driver of trading for open-ended mutual fund managers is flows. For example, when a fund receives an outflow, it must raise cash by selling positions, even if the fund has strong convictions in its holdings. This presents an empirical challenge, as we need to distinguish between position changes that stem from beliefs about returns and those that are a response to flows.[2]

To separate return-based and flow-based motivations for trade in active funds, we document and use an empirical observation from passive index funds. Because these funds seek to track (and not beat) an index, they provide a suitable benchmark for how funds change positions to handle flows. Empirically, index funds nearly perfectly scale up and down their positions in proportion to the number of shares held previously.[3]

[1]For example, see Sharpe (1991), Jensen (1968), and Fama and French (2010) for papers which argue the average active manager cannot or does not beat the market. On the other hand, many papers (e.g., Baks et al. (2001), Pástor and Stambaugh (2002), Avramov and Wermers (2006), Kosowski et al. (2006) and Barras et al. (2010)) argue that it is possible to identify funds with persistent outperformance.

[2]We analyze position changes in terms of shares held instead of, say, portfolio weights because the latter are affected by relative returns within a portfolio. For example, consider a fund which held 10 stocks, did zero trading, had no flows, and 9 of the stocks had a return of 0%, while the 10^{th} stock had a return of 10%. While one stock's weight increased and all other stocks' weights decreased, it was not necessarily a consequence of the fund's belief about future returns. Position changes in share terms are unaffected by relative returns, as a position can only change if a fund takes action to change it. See Section 2.1 for more details.

[3]There are, of course, exceptions. Two common reasons to deviate from perfect scaling are float adjustments and index additions and deletions. Even with these exceptions, index fund position changes still nearly resemble perfect scaling with some small adjustments. See Section 2.5 for more details.

For example, if an S&P 500 index fund receives an inflow of 1% of assets under management (AUM), it deploys the cash by increasing the number of shares held in every position by 1%. This approach is not a coincidence. Index funds – by definition – hold a constant percentage of each constituent's (float-adjusted) shares outstanding. And, proportionally adjusting positions based on shares held ensures that the index fund can deploy cash while preserving value weights (we provide detailed examples in Section 2.1 and Section 2.4). In fact, scaling positions in proportion to lagged shares held preserves relative portfolio weights *from any initial allocation* (see Section 2.5 for details). So, the benchmark of perfect scaling applies to active management, even though such funds' allocations may deviate substantially from what an index fund would do.

Using these insights, we decompose quarterly share changes for *each stock held by each active fund* into a passive rebalancing component and an active rebalancing component. The passive rebalancing measure, or *rebalscale*, is how an index fund would trade in response to flows. For example, if an active fund received a 1% of AUM outflow in quarter t, the predicted share change is -0.01 times the shares held at the end of quarter $t-1$. The active component, what we refer to as intensive-margin rebalancing or *rebalin*, is the difference between the observed position change in a stock and the predicted passive rebalancing in that stock. For stocks newly added or completely sold in a quarter, we define another measure called extensive-margin rebalancing, or *rebalout*, which is simply equal to the number of shares added or dropped.

To see how our measures capture activeness, consider several examples. First, if a fund were to receive a 1% of AUM inflow over a quarter and perfectly scale up its existing positions without adding or dropping any stocks, *rebalin* and *rebalout* would be exactly zero, and all the activity would be captured by *rebalscale*.

Second, consider a similar example fund with the same 1% of AUM inflow that deviates from perfect scaling. Specifically, the fund initially holds an equal-weighted portfolio of ten stocks, but believes stock A will outperform the others and stock B will underperform the others. Thus, the fund increases shares held in A by 3% and sells 1% of shares in B. Outside of these two stocks, the fund increases shares held in all other holdings by 1%. Then, *rebalin* would be 2% for stock A, -2% for B and zero for all other positions. That is, *rebalin* captures the fund's beliefs of future performance in specific stocks by the way it deviates from perfect scaling.

Both of these examples capture only intensive-margin rebalancing. Consider a third example fund, which also initially holds an equal-weighted portfolio of ten stocks but receives a 10% of AUM inflow and decides to add a new stock, which it believes will outperform the others. To operationalize this view, the fund allocates half of the inflow to the new stock and half to scaling up its existing positions. In this case, *rebalout* would be positive for the new stock and *rebalin* would be -5% for each of the ten previously-held stocks. These measures represent that the fund has a more positive outlook on the new stock than all the previously-held stocks.

We construct our measures using the Thomson Reuters Mutual Fund (S12) dataset. Since our measures require position-by-position changes per fund and quarter, we develop a detailed data-cleaning methodology to address many issues with the data that are unique to our setting. Essentially all studies use some form of aggregation (e.g., averaging or summing over stocks within a fund or across funds that hold a given stock), which may implicitly or explicitly address many of the issues at the more granular stock-fund-quarter level. We carefully document each step of our data cleaning methodology in Section 2.6 and Appendix B to ensure replicability of our results and to provide a guide for future research.

We use our measures to test whether active rebalancing in quarter t is associated with future returns in $t + 1$. We conduct tests at the stock-fund-quarter level by regressing two measures of future stock returns, four-factor alpha (computed from daily returns in $t + 1$) and benchmark-adjusted returns, on our measures of active rebalancing. In order to ensure comparability across managers and across time, we re-scale all the rebalancing measures to be fractions of AUM by multiplying the change in shares by the end-of-quarter price and dividing by end-of-quarter AUM. We further split each of our measures into positive and negative components. For example, we split *rebalin* into two variables: one that captures positive signals (where *rebalin* is positive, zero otherwise) and one that captures negative signals (where *rebalin* is negative, zero otherwise).

We find that larger positive signals (i.e., positive active rebalancing), both on the intensive- and extensive-margins, predict higher future alphas. On the intensive-margin, actively rebalancing 1% of AUM towards a stock predicts 22 basis points to 36 basis points of quarterly alpha in the following quarter. On the extensive-margin, adding 1% of AUM in a new stock holding predicts 17 basis points to 21 basis points of quarterly alpha the next quarter.

We find weaker and less significant results for sales. In fact, we find evidence that more selling is associated with higher future returns. Further, we do not find a statistically significant relation between any active rebalancing measures and benchmark-adjusted returns. Finally, these findings are not sensitive to the inclusion of various fixed effects (fund and quarter, fund-by-quarter, and no fixed effects). We interpret the consistency across fixed effects specifications as evidence that the relation between active rebalancing and future alpha holds in three ways: (1) unconditionally, (2) between more and less active funds, and (3) relative to other stocks in a fund's portfolio.

To rule out mechanical price impact explanations for our results, we perform two additional tests. The first is testing whether the flow-predicted rebalancing component of position changes, $rebalscale$, is also associated with future price changes. We find no evidence that the passive rebalancing measure predicts future alphas or benchmark-adjusted returns. In addition, we also examine index mutual funds/ETF, which provide a useful placebo test, as they do not trade on expectations about future performance. That is, if there is a relation between future alphas and our active rebalancing measures for index funds, it is more likely to be from a mechanical impact of trading on future prices. We find no evidence of similar results between active and passive funds. While we believe our baseline results for active funds are difficult to reconcile with a price-pressure story (given purchases lead to continuation and sales lead to reversal), these additional tests provide greater assurance.

We then explore possible drivers of the relation between active purchases and future alphas. We find that the relation is significantly stronger – both in terms of magnitudes and statistical significance – in quarters with inflows relative to quarters with outflows. We interpret this as evidence that portfolio adjustment costs, such as transaction costs/price impact of selling existing positions, are an important factor in managers' rebalancing decisions since such constraints may be relaxed by inflows.

We further split our sample based on other stock- and fund-related variables to better understand the relation between active rebalancing and alpha. We find several patterns. First, the relation between positive active rebalancing and future alpha is strongest among positions with embedded gains and short holding periods. We argue this is evidence that managers with limited attention may focus their information gathering on a small subset of stocks (Cohen et al. (2009), Pomorski (2009)). This leads to outperformance in these stocks – which is why they have relatively high embedded gains – and managers better actively adjust these positions going forward. It is also evidence that managers may have a saliency bias, focusing more on stocks that have been added recently (Barberis, 2018).

In addition, we find our results on positive intensive-margin rebalancing are stronger among funds that hold fewer individual positions, perhaps because these managers are focused on making stock-level bets instead of sector- or factor-level bets. On the other hand, funds that hold many positions tend to outperform funds with few positions on positive extensive-margin rebalancing (i.e., adding new stocks). This suggests that when managers hold many stocks, they have greater predictive power for adding new stocks; if managers hold few stocks, they have greater predictive power in adding to existing stocks. We do not find substantially different results for large and small funds (based on AUM).

To speak directly to the relation between active managers' rebalancing decisions and information gathering, we examine the relation between active rebalancing and future earnings surprises. The logic is that if managers are gathering fundamental information, we would expect their trades to predict future fundamental news. We find a significant relation between both positive intensive-margin and extensive-margin rebalancing and future standardized unexpected earnings (SUE). Further, this is concentrated in the idiosyncratic component (rather than the market-wide or industry-specific component) of SUE, evidence that active managers are gathering firm-specific information.

Finally, to better understand the rebalancing measures themselves we run a series of regressions to evaluate the importance of other motivations for trade like transaction costs, past returns, recent rebalancing, and embedded gains. We note several interesting patterns. First, we find that active rebalancing decisions are somewhat autocorrelated. For example, we show that managers are more likely to add to positions they've been building and more likely to trim positions they've been cutting in the recent past. This is consistent with managers slowly building and trimming positions, rather than aggressively trading in and out of stocks based on short-lived information.

That being said, the magnitudes of the autocorrelation in active rebalancing decisions are economically small. For example, when a firm has positive intensive-margin rebalancing of 1% of AUM in quarter t, this predicts intensive-margin rebalancing in quarter $t+1$ of 0.12% of AUM in the same stock. In addition, past rebalancing activity in general is associated with active rebalancing, regardless of sign. For example, past active negative rebalancing in a stock is statistically related to both positive and negative rebalancing in the future. That is, the more funds actively rebalance a stock, the more likely they are to actively rebalance that stock in *some* direction in the future. These results further allay concerns of autocorrelation in active rebalancing mechanically driving the relation between rebalancing and future returns through price pressure from continued buying or selling.

Second, funds are less likely to add to positions that have embedded gains or have performed better than other stocks in their portfolio, but are more likely to trim these positions or sell them all together. This is consistent with the disposition effect (Odean (1998), Frazzini (2006)), which describes the tendency to sell out of positions with embedded gains too quickly, and hold on to losing positions for too long.

Third, funds are more likely to actively rebalance stocks with high recent returns, but the effect is much greater with negative rebalancing. That is, funds are much more likely to trim or completely sell stocks than add to existing positions with high returns. In addition, funds do not actively rebalance toward stocks with positive alpha but do engage in negative rebalancing away from stocks with high realized positive alphas. Lastly, funds are sensitive to trading costs: funds do less active rebalancing on both the intensive- and extensive-margins when price impact is high.

We note that our results at the stock-fund-quarter level do not necessarily mean that these funds deliver high returns to investors. While we find high future alphas in the stocks that funds are more actively rebalancing towards, our findings capture two things: (1) the association between alphas and the size of a *rebalancing decision*, not the size of a position, and (2) the return of a stock if the fund were to hold the position for the entirety of the quarter (without accounting for the weight of that stock). Because our measures are built on quarter-over-quarter changes in holdings, we are not able to account for active trading within a quarter, other expenses that the fund may have incurred, or whether alphas in one actively traded stock are offset by other less-well-performing stocks held by the fund.

We can, however, speak to fund-investor returns by aggregating up our measures at the fund-quarter level. That is, we can ask whether more positive and negative active rebalancing is related to returns as reported by the fund itself – which will account for within quarter trading and the returns to positions which were less actively rebalanced. We find that larger portfolio-level positive active rebalancing is associated with higher future alphas, as well as higher future benchmark-adjusted fund returns (both gross and net of fees) – although without strong statistical significance.

Contribution We make three sets of contributions to the literature.

The first contribution is to develop a data cleaning methodology for the Thomson Reuters Mutual Fund (S12) dataset to study individual fund positions and their changes. This requires careful, painstaking data examination since data errors and mistaken assumptions will result in completely erroneous inferred position changes. We carefully document our data cleaning methodology in Section 2.6 and Appendix B. The data cleaning required is extensive, and may be a reason why most, if not all, papers utilizing the S12 data have focused on fund-level or portfolios-of-fund-level analyses. We hope the data cleaning approach documented in this paper will encourage more research that uses the full scope of the S12 data.

The second contribution is to document empirical facts on a passive scaling benchmark used by passive index funds in response to fund flows, and to use that benchmark to develop a simple methodology to separate active rebalancing decisions from passive scaling at the individual position level. Our contribution in this area is significant in several ways: (1) Passive index funds have very predictable trading (in response to flows), (2) We are able to separate passive from active trading decisions of fund managers, and (3) We are able to aggregate our position-level measures up to the fund level to assess the activeness of managers overall.

The third contribution uses these data and methodological innovations to document new facts on the actual activity of fund managers, which we described earlier in the introduction. Our methodologies are critical to our analysis: including flow-induced positions changes into measures of fund activity do not provide clear signals for future position-level performance. It is exactly the active rebalancing component of position changes that is associated with future alphas.

1.1 Related Literature

Our paper contributes to several strands of literature on active management. First, there is a long literature that studies whether active managers have skill and how we might identify them (Jensen (1968), Sharpe (1991), Wermers (2000), Berk and Green (2004), Kosowski et al. (2006), Carhart (1997), Fama and French (2010), Barras et al. (2010), Berk and Van Binsbergen (2015), Giglio et al. (2021), Kaniel et al. (2023)). Our paper is different, as we are focused on stock-level performance after changes in manager's *individual positions*, rather than portfolio-level performance.

Our paper is related to the literature linking manager activity to future fund performance. For example, Cremers and Petajisto (2009) and Petajisto (2013) study whether managers whose holdings deviate more from benchmark weights (i.e., managers with higher "active share") are more likely to outperform those benchmarks. Our rebalancing measures are different from active share along two dimensions: (1) our measures are defined at the position level, rather than at the fund-quarter level, and (2) active share is a persistent fund-level characteristic, while our measure varies significantly within a fund over time and across positions within a given fund-quarter. Another paper in this literature is Pástor et al. (2017), who show that portfolio level turnover predicts fund returns going forward. Again, our innovation is twofold: (1) our measures speak to whether the relation from portfolio turnover is due to the positions which are actually adjusted i.e., whether the activeness itself is driving future returns, and (2) our measure separates flow-induced activity from discretionary activity.

A closely related paper to ours in this literature is Alexander et al. (2007), who compare the returns of funds which are buying large quantities of stocks in the face of outflows against funds which are buying few stocks in the face of inflows. The logic is that a manager buying in the face of outflows likely has a strong belief that such stocks will outperform, while a manager who doesn't buy in the face of inflows is likely out of ideas. The authors find that managers buying in the face out outflows outperform managers who buy little in the face of inflows – consistent with information-motivated trades outperforming liquidity-motivated trades.

This builds on a broader literature (see e.g., Edelen (1999)) which shows that funds perform poorly when they receive inflows. Our paper is different, in that we are looking at deviations from expected trading in response to flows at the individual position level. In fact, we find that even when managers receive inflows, the individual positions they add *more than predicted by scaling up* tend to outperform over the following quarter.

Yet another related strand of literature has focused on evaluating the performance of managers' individual positions and "best ideas" (Cohen et al. (2009), Pomorski (2009), Lewellen (2011), Hartzmark (2015)). Our paper has a different focus, in that we examine the largest *changes* in positions, while these papers are mostly focused on the biggest deviations from benchmark weights. We argue that a manager could have a very large position, but scale that position up and down in response to flows, and this may have different implications for future returns than large changes in such positions.

While papers like Cremers and Petajisto (2009) and Pástor et al. (2017) are focused on aggregation at the *fund-quarter* level, there is another literature which has aggregated mutual fund activity to the *stock-quarter* level. For example, Chen et al. (2000) examine whether changes in the fraction of a stock's shares outstanding held by all active mutual funds predicts future returns. They find that aggregate buying by active funds predicts high returns going forward, while aggregate selling has little predictive power for future returns. Our paper is different in two ways: (1) we are focused on changes at the individual stock-fund-quarter level, although we do present stock-quarter aggregation results in Appendix A.6, and (2) Chen et al. (2000) acknowledge that their measure of aggregate buying is more likely to be positive when the active mutual fund industry has inflows i.e., their measure does not account for the mechanical effect portfolio changes in response to flows.

Finally, we believe one of the closest paper to ours is Akepanidtaworn et al. (2021), who show that investment managers appear to have skill in buying decisions, but not selling decisions. This seems to mirror our results that managers' positive intensive and extensive-margin rebalancing predicts high future alpha, while negative rebalancing does not predict low alpha going forward. We believe our contribution is different in several ways. The first is methodological: Akepanidtaworn et al. (2021) focus on separately managed accounts (SMAs), for which flows are less of a concern than for open-ended equity mutual funds. One of our contributions is to construct a methodology which accounts for the mechanical part of rebalancing managers must engage in to account for flows. In fact, in their setting, flows are relatively rare, so to raise money for new ideas, SMA managers will almost always have to sell existing positions. We are able to split our sample on fund-quarter observations with inflows and outflows, and show that relaxing this constraint (i.e., needing to sell in order to add new positions) makes managers' trades more predictive of future performance on average.

Another advantage of our setting is that we have a natural counterfactual test using passive funds. Akepanidtaworn et al. (2021) focus on a set of skilled managers who earn positive alpha on average, while we have the universe of active managers who arguably earn zero alpha on average (Fama and French, 2010). By comparing our active managers to passive funds, we can rule out several mechanical explanations for our results like price pressure. Finally, and more broadly, we believe our paper has a different focus. Akepanidtaworn et al. (2021) are documenting behavioral biases that lead to differences between buying and selling decisions. Our paper is focused on a different question, namely whether deviations from a passive rebalancing benchmark in response to flows predict alpha going forward.

2 Methodology & Data

2.1 Motivation

We aim to measure active trading in individual stocks by fund managers, then ultimately combine those measures to describe overall fund activeness. The challenge is to separate out passive trading in response to flows with active decision-making designed to generate outperformance. In this setting, our benchmark for a totally passive strategy is a value-weighted index fund (VWIF). First, suppose this VWIF receives no flows. Some stocks may perform better than others, but in the absence of changes to the index (in the form of e.g., additions, deletions and changes in float-adjustments) the fund will not have to trade. When such a fund receives inflows, it will scale up all its positions in proportion to the number of shares it held previously. The same fund will scale down positions proportionally when it receives outflows.

This is because a value-weighted index fund should hold the same fraction of each of the index constituent's (float-adjusted) shares outstanding. Our measures are motivated by the logic that funds with a larger tendency to scale up and down existing positions in response to flows are more like a VWIF, and therefore its trading is more passive.

To illustrate the mechanics of passive rebalancing, consider an example value-weighted index fund in Table 1 that holds Facebook, Apple, Amazon, Netflix and Google and has $6,000 assets under management (AUM). Initially, the fund's AUM is 10 basis points of total index capitalization and therefore it holds 10 basis points of each constituent's shares outstanding. Now suppose that at the end of the quarter, this fund receives

Table 1: Example Value-Weighted Index Fund, Flows are 10% of AUM

Stock	$shares_t$	$shrout_t$	p_{t+1}	Pre-Flow $dheld_{t+1}$	Pre-Flow $pct.float_{t+1}$	Post-Flow $\Delta\$$	Post-Flow $\Delta shares$	Post-Flow $shares_{t+1}$	Post-Flow $dheld_{t+1}$	Post-Flow $pct.float_{t+1}$
FB	100	100,000	$ 22.50	$ 2,250.00	0.10%	$ 225.00	10	110	$ 2,475.00	0.11%
AAPL	35	35,000	$ 57.35	$ 2,007.25	0.10%	$ 200.73	3.5	38.5	$ 2,207.98	0.11%
AMZN	10	10,000	$ 19.99	$ 199.90	0.10%	$ 19.99	1	11	$ 219.89	0.11%
NFLX	20	20,000	$ 36.23	$ 724.60	0.10%	$ 72.46	2	22	$ 797.06	0.11%
GOOG	20	20,000	$ 40.91	$ 818.25	0.10%	$ 81.83	2	22	$ 900.08	0.11%
Total				$6,000					$ 6,600	

Notes. This table provides an example of a hypothetical value-weighted index fund that holds 5 stocks. The table provides the shares held by the fund (*shares*), the shares outstanding in the market (*shrout*), and the price of the stock (p). The table illustrates how an index-fund adjusts holdings in each stock in response to an inflow of 10% of AUM. We provide the dollars held (*dheld* and the percentage of shares outstanding held ($pct.float$), as well as the position change in dollars and shares ($\Delta\$$ and $\Delta shares$) to show that when a value-weighted index fund receives flows, it scales its holdings in proportion to shares held previously to ensure it continues to hold a constant fraction of each stock's shares outstanding.

$600 of inflows i.e., inflows equal to 10% of its AUM. To ensure it continues to hold an equal fraction of each constituent's shares outstanding, the fund can scale up the number of shares it holds of each stock by 10%. After investing the flows, the fund owns 11 basis points of each constituent's shares outstanding.

One way to quantify this type of passivity is to run the following regression for each fund j in each quarter t:

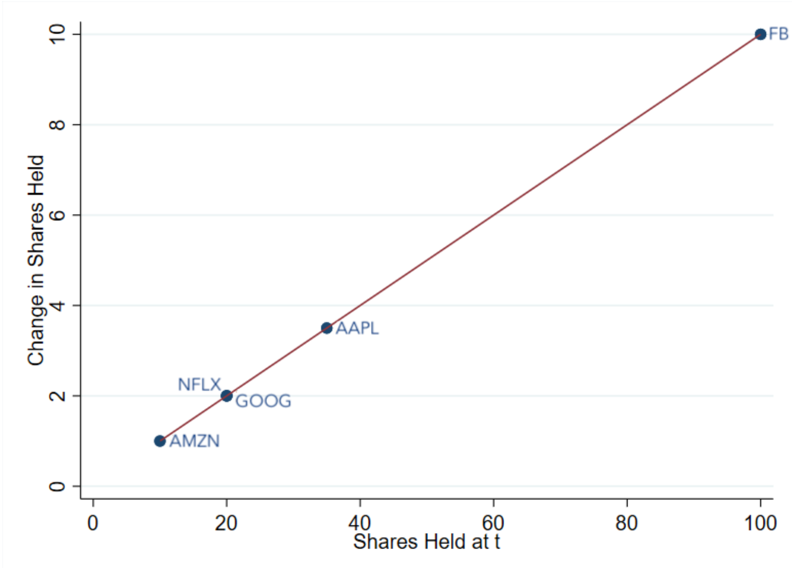
$$\Delta shares_{i,j,t} = \beta_{j,t} shares_{i,j,t-1} + \epsilon_{i,j,t} \tag{1}$$

where $shares_{i,j,t}$ is the number of shares held by fund j of stock i in quarter t and $\Delta shares_{i,j,t}$ is the quarter-over-quarter change in shares held. As mentioned above, if the underlying index doesn't change, the VWIF's change in shares held should be a perfect linear function of lagged shares held for each position, i.e., the R-squared of this regression should be 1. A constant term is not included in Equation 1 because for a fund that perfectly scales up and down its existing positions, the best-fit line will pass through the origin. Finally, if observations are weighted in proportion to portfolio weights in quarter $t-1$, $\beta_{j,t}$ will be equal to flows, as a fraction of total AUM.

Figure 1 shows this is indeed the case for our example VWIF, as the change in shares held is a perfect linear function of lagged shares held. Further, the estimated slope, $\beta_{j,t}$ is 0.1, which is equal to flows as a fraction of AUM.

At first glance, it seems as though there are other ways our example fund could invest $600 of inflows to preserve value weighting, e.g., by scaling up positions in proportion to dollars previously held. More broadly, a natural question is why we are focused on shares held, rather than portfolio weights or dollars held. The reason is because working in shares naturally preserves value weighting even in the face of differences in returns across holdings and flows.

Figure 1: Example Index Fund, Change in Shares vs. Shares Held



Notes. This figure plots the change in shares held against shares held by the example value-weighted index fund that received an inflow of 10% of AUM described in Table 1. The figure illustrates that an index fund increases shares held in proportion to shares held to ensure it continues to hold a constant fraction of each stock's shares outstanding.

We illustrate this in Table 2. In this table, our example VWIF received no flows. Recall that a VWIF holds a constant fraction of each constituent's shares outstanding, so if the fund receives no flows, the fund will not have to trade. Despite this, the weights in each position changed from quarter t to $t+1$ due to differences in relative returns. We believe that working with shares is more straightforward than dollars/weights because it accounts both for differences in relative returns (because of value-weighting) and has a straightforward way to account for inflows and outflows. Said differently, weights can change from returns or trading; share changes can only occur with action by the fund.

More broadly, we would like to highlight that from the perspective of a VWIF, weights are something that is determined ex-post *after* the fund buys the required percentage of the constituent firm's shares outstanding. This means that the size of the buying "shock" a stock experiences from being added to a value-weighted index is only a function of the difference between (1) the ratio of the dollars tracking the index it was added to relative to that index's total capitalization and (2) the ratio of the dollars tracking the index it moved from relative to that index's total capitalization (if it was simultaneously dropped from another index). The weight of the stock, either in the index it moved to or moved from, has no bearing on what percentage of that firm's shares are purchased after index addition.[4] For example, at the time of writing, a larger fraction

[4] As discussed in Appel et al. (2016), when a firm switches from, e.g., the Russell 1000 to the Russell 2000, its relative portfolio weight increases (going from the smallest stock in an index of large stocks, to the largest stock in an index of small stocks). So, after switching, for every dollar invested in the *index it belongs to*, a relatively larger share of that dollar gets invested in that stock. While this is true in dollar terms, it is not true in terms of the percentage of each index member's shares outstanding. When a given dollar amount flows into a value-weighted index fund, that goes to purchasing the same fraction of

Table 2: Example Value-Weighted Index Fund, No Flows

			Before Returns Realized			After Returns Realized		
Stock	$shares_t$	$shrout_t$	p_t	$dheld_t$	wt_t	p_{t+1}	$dheld_{t+1}$	wt_{t+1}
FB	100	100,000	\$ 22.50	\$ 2,250.00	38%	\$ 24.75	\$ 2,475.00	39%
AAPL	35	35,000	\$ 57.35	\$ 2,007.25	33%	\$ 48.75	\$ 1,706.16	27%
AMZN	10	10,000	\$ 19.99	\$ 199.90	3%	\$ 14.39	\$ 143.93	2%
NFLX	20	20,000	\$ 36.23	\$ 724.60	12%	\$ 54.35	\$ 1,086.90	17%
GOOG	20	20,000	\$ 40.91	\$ 818.25	14%	\$ 42.96	\$ 859.16	14%
Total				\$6,000			\$ 6,271	

Notes. This table provides an example of a hypothetical value-weighted index fund that holds 5 stocks. The table provides the shares held by the fund (*shares*), the shares outstanding in the market (*shrout*), the price of the stock (p), dollars held (*dheld*), and the portfolio weight (*wt*). The table illustrates how an index fund's weights change with returns but shares held do not. In other words, weights can change from trading or differences in relative returns, while changes in shares held can only come from trading.

of the Russell 2000 is owned by passive funds than the Russell 1000 (Pavlova and Sikorskaya, 2023). This is why firms switching to the Russell 2000 experience an increase in their passive ownership share, i.e., the percentage of their shares outstanding held by passive funds. It is not due to the fact that firms switching from the Russell 1000 to the Russell 2000 experience an increase in index weight.

More broadly, as we discuss in Section 2.5, scaling up and down shares held preserves value weights from any *initial* weighting scheme. This is an additional benefit of working in shares, as many active funds' portfolio weights differ significantly from market and/or benchmark weights.

2.2 Measure Construction

With this benchmark of passivity in mind, we aim to decompose all trading into the following three components:

$$\Delta shares_{i,j,t} = \underbrace{\text{Passive rebalancing}}_{rebalscale} + \underbrace{\underbrace{\text{Intensive margin rebalancing}}_{rebalin} + \underbrace{\text{Extensive margin rebalancing}}_{rebalout}}_{\text{Active trading}} \tag{2}$$

where $\Delta shares_{i,j,t}$ is the change in shares held of stock i from quarter $t-1$ to t by fund j.

We define the first part of this decomposition, passive rebalancing, as follows:

$$rebalscale_{i,j,t} = (flow_{j,t} \times shares_{i,j,t-1}) \times p_{i,j,t}/AUM_{j,t}, \tag{3}$$

each constituent's shares outstanding, regardless of differences in relative weights.

where $flow_{j,t}$ are the flows received by fund j in quarter t as a fraction of AUM, and $shares_{i,j,t-1}$ is the shares held of stock i by fund j at the end of quarter $t-1$. The logic behind $rebalscale$ is that the passive part of rebalancing is a perfect scaling up and down of lagged shares held in proportion to flows received. We then multiply this quantity by the end-of-quarter price, $p_{i,j,t}$, and divide by end of quarter assets under management, $AUM_{j,t}$, to make the units of $rebalscale$ a fraction of the fund's total assets.

Also in line with the framework outlined above, our benchmark for passive investing does not include additions to or deletions from the portfolio. The logic for this decision is that, for a passive fund, additions and deletions are "exogenous" in the sense that they are determined by index rules or a committee decision, and therefore independent of the decision to scale up/down positions in response to flows. So, in our empirical exercises, $rebalscale$, is only calculated for stocks held by fund j in quarter t and quarter $t-1$.

The first type of active trading, based on deviations from the behavior of a value-weighted index fund, is intensive-margin rebalancing, defined as:

$$rebalin_{i,j,t} = (\Delta shares_{i,j,t} - \underbrace{flow_{j,t} \times shares_{i,j,t-1}}_{\text{Predicted change in shares}}) \times p_{i,j,t}/AUM_{j,t}. \tag{4}$$

The logic behind this is as follows: For a VWIF, in a given quarter, the expected change in shares held is $flow_{j,t} \times shares_{i,j,t-1}$, i.e., $rebalscale$. The realized change in shares, however, is $\Delta shares_{i,j,t}$, so the difference between these two quantities is an active decision to deviate from what a VWIF would do. Like $rebalscale$ we only compute this for stocks held by fund j both in quarter $t-1$ and quarter t and we multiply by the stock price and divide by AUM so the units are again a fraction of total assets held.[5]

In terms of interpretation, stocks will have a positive value for $rebalin_{i,j,t}$ when the fund has added to a position more than would be predicted by flows. As an example, this could include receiving an inflow and increasing a position by more than a VWIF would, or receiving an outflow and selling a position but by less than what a VWIF would do.

For a VWIF, which by design perfectly scales up and down its existing positions, intensive-margin rebalancing is expected to be zero for all stocks each quarter. That being said, a mismatch between the portfolio weight of added stocks and dropped stocks can lead to deviations from scaling up and down all positions in proportion to flows. To account for this, we construct a measure of extensive-margin rebalancing for stocks that were

[5] One possible concern with this definition of active rebalancing is that a manager could leave a position unchanged, and yet that position could have a non-zero value for $rebalin$. For example, if a fund receives inflows, and the manager doesn't adjust their holdings of stock i, $rebalin$ will be negative. In Appendix A.1 we show our main results are not driven by manager inactivity being classified as active rebalancing.

added to or dropped from the fund entirely in quarter t:

$$rebalout_{i,j,t} = \Delta shares_{i,j,t} \times p_{i,j,t} / AUM_{j,t}. \tag{5}$$

Based on the way *rebalout* and *rebalin* are defined, a stock cannot have non-zero values for both of these measures.

In our empirical analysis, we further split our rebalancing measures into the positive and negative components. For intensive-margin rebalancing, we call these $rebalin_{i,j,t}^{positive} = \max(rebalin_{i,j,t}, 0)$ and $rebalin_{i,j,t}^{negative} = \min(rebalin_{i,j,t}, 0)$. For extensive-margin rebalancing, we call these $rebalout_{i,j,t}^{add} = \max(rebalout_{i,j,t}, 0)$ and $rebalout_{i,j,t}^{drop} = \min(rebalout_{i,j,t}, 0)$. All 4 of these measures are mutually exclusive, in the sense that for any stock-fund-quarter observation, only one of them can differ from zero.

2.3 Assumptions

Given that our rebalancing measures are constructed using quarterly holdings data, we implicitly rely on several assumptions – all related to the timing of how managers change their portfolios – to ensure that they are valid approximations of the actual trading decisions of managers. First, we assume managers add positions on the last day of the quarter the stock first appears in holdings. This means that when computing returns to a fund's underlying positions, we do not include additions until the first full quarter they appear to be held by the fund. For example, suppose that for a given manager, Apple first appears in their portfolio in 2020q1. We assume the manager bought Apple on the last day of 2020q1, and therefore Apple is not included for the return calculation of 2020q1. The logic is that it could have been bought any time throughout 2020q1 – and we don't know when – so to be conservative, we assume it doesn't count toward 2020q1 returns. For the same reason, we also assume that managers drop positions on the last day of the quarter. We believe this is a conservative assumption, as managers may be tempted to add stocks to their portfolio that have done well over the past quarter and drop poor performers.

The second main assumption, following Barber et al. (2016), is that all flows are received at the end of the quarter. While this assumption seems strong, we believe it is not central in explaining how a VWIF would scale up/down positions in response to flows. Suppose, for example, that between the start of the quarter and the middle of the quarter, stocks in the manager's portfolio have heterogeneous returns. As shown in our example in Table 2, holding a constant fraction of each firm's shares outstanding will take care of value weighting on its own. Further, suppose the manager gets flows in the middle of the quarter. Regardless of when the flows occur, the scaling should be the same, as value weighting has already taken care of the reweighting associated with heterogeneous returns. What this assumption could matter for is fund-quarter level returns, and therefore our estimate of flows. If flows happen early and returns happen late, then our estimate of flows will be over or understated depending on the sign of the return.

2.4 Measure Examples

To clarify our rebalancing measures, we present visual examples from our actual data in Figure 2. In the top left panel, we present data for SPY – an S&P 500 ETF – in 2016 Q4. The x-axis represents shares held by the fund at the end of quarter $t-1$, and the y-axis represents the change in shares held between quarters $t-1$ and t, both in millions. Each blue dot represents an individual stock e.g., the rightmost dot is SPY's holdings of AAPL. In 2016 Q4, SPY received an inflow equal to roughly 10% of AUM. So, in line with equation 3, the solid red line has a slope of 0.1, because the predicted change in shares is equal to 10% of lagged shares held. Recall that this is exactly *rebalscale* which is equal to $flow_{j,t} \times shares_{i,j,t-1}$.

The green dashed line is the estimated slope from equation 1 i.e., a regression of the change in shares on lagged shares held, with weights proportional to the dollars held of each stock at the end of the previous quarter. As discussed above, if the total portfolio weights of additions and deletions are close, the estimated slope will be equal to flows as a fraction of AUM i.e., $\beta_{j,t} = flow_{j,t}$. This implies that the fitted values from the regression, $\beta_{j,t} \times shares_{i,j,t-1}$ will be equal to expected scaling, $flow_{j,t} \times shares_{i,j,t-1}$. This holds for SPY in this quarter, which is why the red line – expected scaling – and the dashed green line – realized value-weighted average scaling – are almost perfectly overlapping. The vertical orange lines represent *rebalin*, which are deviations from expected scaling (i.e., the red line).

A natural question is why a value-weighted index fund like SPY has values of *rebalin* which are not equal to zero. These deviations are mainly caused by differences in portfolio weights between additions and deletions, as well as changes in float-adjustments. For example, if a company does buybacks, the stock's float will typically decrease, and index funds will be forced to sell in the next quarterly rebalance. For SPY in 2016 Q4, however, these deviations are relatively small: the R-squared of the regression is nearly 1.

The bottom left panel has the same setup except it uses data from AWSHX, the largest active mutual fund at the end of our sample. We plot data from 2010 Q3, when AWSHX received outflows equal to 3% of AUM. Further, like for SPY in the top left panel, the average scaling down of positions in this quarter was 3%, which is why the dashed green line is nearly on top of the solid red line. Unlike, SPY, however, there are significant deviations from prefect scaling down among the individual positions. This is why the R-squared of the regression of change in shares on lagged shares is close to zero, and why there are large vertical orange lines, which represent significant intensive-margin rebalancing.

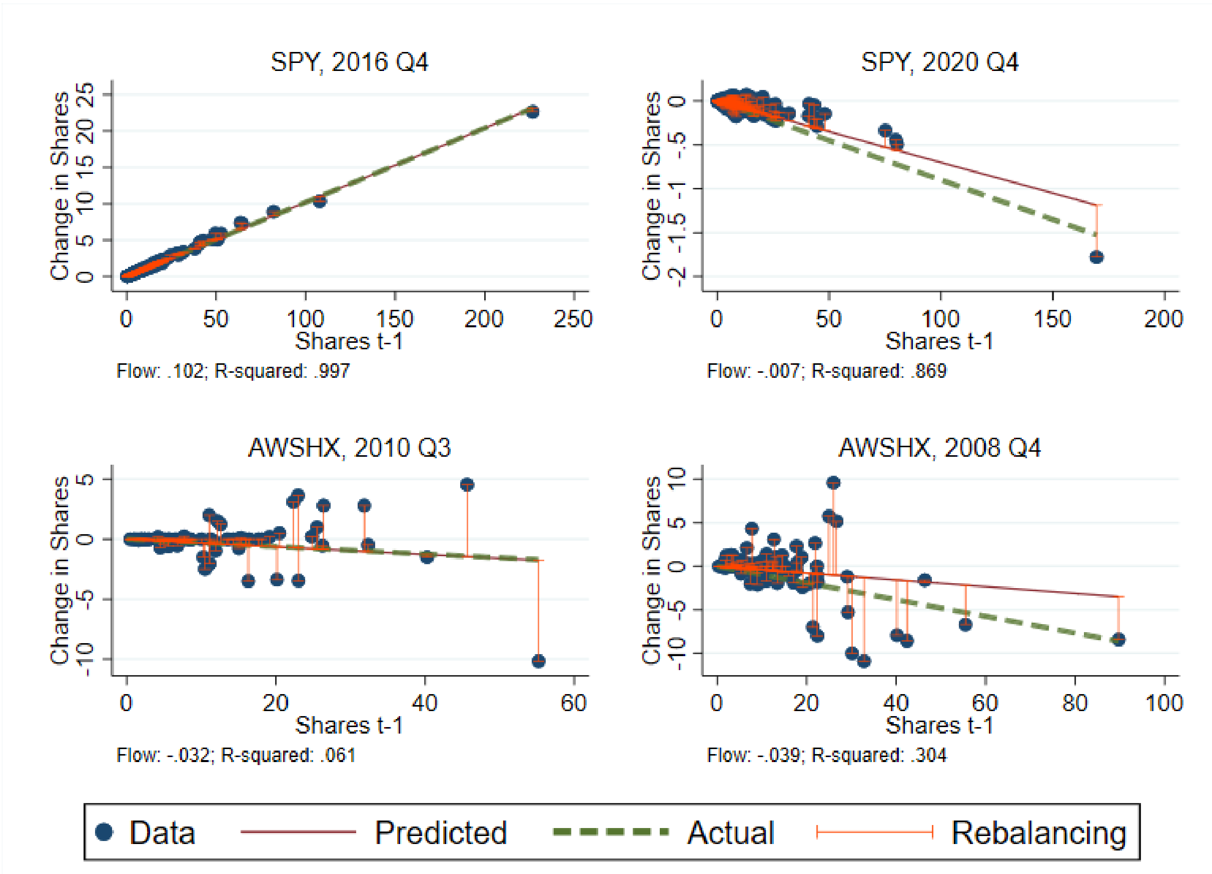
In these first two examples, we have made no mention of extensive-margin rebalancing. This is because in these quarters, the size of additions and deletions (as a fraction of AUM) was roughly equal (i.e., the red line was roughly equal to the dashed green line). The next two panels illustrate cases where this does not hold and the resulting implications for intensive-margin rebalancing. In the top right panel, we plot data from SPY in 2020 Q4. In this quarter, Tesla was added to the S&P 500, with a weight of 1.69%. The corresponding index deletion was Apartment Investment Management, which when it was dropped had a weight of 0.012%. Finally, in this quarter, SPY had outflows equal to 71 basis points of AUM. This creates a mismatch, because the deletion is much smaller than the addition, so if SPY perfectly scaled everything down by 71 basis points (i.e., in proportion to flows), there would not be enough capital to give Tesla its proper weight.

SPY made room for Tesla by scaling down all their other positions by more than would be expected given flow-based scaling. This is why the dashed green line – value-weighted average realized scaling – sits below the red line – expected scaling. The proceeds from this additional scaling down were used to buy enough Tesla to give it an index weight of 1.69%. A striking feature of this panel is the large intensive-margin rebalancing for the single observation at the bottom right of the plot, which is SPY's position in AAPL. Throughout this quarter, AAPL did a significant amount of buybacks, which decreased the float, and thus SPY had to sell a large portion of its position at the end of the quarter.

A consequence of needing to make room for Tesla is that many of the positions for SPY are going to have negative values for intensive-margin rebalancing in 2020 Q4. This is because average scaling sits below expected scaling. While this may seem counterintuitive, we want to reiterate that our rebalancing measures are based on a very specific benchmark of what it means to be passive. Related to this, another striking feature of this panel is that most of the data points seem to lie above the best fit line. In computing the slope of the dashed green line, each position is weighted in proportion to lagged dollars held. So fitting the point associated with AAPL is significantly more important (from a minimizing least squares perspective) than the point associated with, e.g., Domino's Pizza, which had a market capitalization about 1/100th of AAPL's. Even with what seems like a lot of intensive margin rebalancing, the R-squared is still very high, capturing that most of what SPY is doing is capture by passive rebalancing.

This behavior of scaling everything down more than expected to make room for new additions is not specific to passive funds. In the bottom right panel of Figure 2, we show data for AWSHX in 2008 Q4. As with SPY in 2020 Q4, AWSHX scaled down all the positions (on a value-weighted average basis) by more than 4% – which is what would be required to meet redemptions – to free up capital for additions to the portfolio. Also note that even though SPY in 2020 Q4 scaled down everything more than would be expected given flows, it essentially scaled things down proportionally (the R-squared was 0.869). This was not the case for AWSHX in 2008 Q4, with an R-squared from this same regression of just 0.304. This, however, is unsurprising, because AWSHX is an active fund and can freely adjust its existing positions to make room for new additions.

Figure 2: Examples of Rebalancing: Index vs. Active Funds



Notes. In each panel, the x-axis is shares held by the fund in quarter $t-1$ (in millions), and the y-axis is the change in shares from quarter $t-1$ to t (in millions). The red line – expected scaling – goes through the origin and has a slope equal to flows as a percentage of AUM over quarter t. The green line – realized average scaling – is the best fit line from a regression of change in shares on lagged shares held with no intercept and weights proportional to the fund's dollar holding of the stock in quarter $t-1$. Orange vertical lines – intensive-margin rebalancing – represent deviations from perfect scaling. Reported R-squared values at the bottom of each panel are from the estimated regression (i.e., the green line).

2.5 Discussion

In this subsection, we explain additional implications for our measures of rebalancing, as well as how our measures relate to previously studied measures of fund activeness.

Separating Intensive- and Extensive-Margin Rebalancing

A natural question is why we are treating additions and deletions differently from stocks that the fund held in both quarter t and $t-1$. We could, for example, apply equations 3 and 4 to all positions, setting $shares_{i,j,t-1}$ to zero for additions. Our main motivation for this is that, for a VWIF, the decision to scale up and down existing positions is independent of the need to add or drop stocks based on changes to the underlying index. This is apparent in Figure 2, where SPY needed to make room for Tesla, and did so by scaling down all its existing holdings (approximately) in proportion to how much they held at the end of the previous quarter.

Further, we have many reasons to believe that the decision to drop a stock entirely has different information than trimming a position. For example, we interviewed a former portfolio manager at Fidelity who explained, "It's mentally way easier to trim than to drop entirely. Dropping entirely normally takes discipline if it's a stock that has really worked... I would be willing to bet that most of the time when a position goes from whole hog to 0, something really bad has happened." More broadly, the literature has shown that trimming positions has different predictive power for future returns than exiting positions entirely (see e.g., Akepanidtaworn et al. (2021), Ahn et al. (2023)).

Value Weights

Although the scaling up of positions in proportion to lagged shares held applies to value-weighted index funds, it also applies to any fund trying to maintain value weights *from a given starting point*. For example, suppose a manager holds \$50 in Facebook and \$50 in Microsoft. This is not a value-weighted portfolio in the traditional sense, because at the time of writing, Facebook and Microsoft do not have the same market capitalization.

Now, suppose the manager has no flows, but that Facebook has a return of 10% and Microsoft has a return of 0%. By not trading, the manager now holds \$55 of Facebook and \$50 of Microsoft. This is like value weighting from the starting point of a 50-50 portfolio split, in that the increased weight on Facebook accounts for the increase in Facebook's stock price – relative to Microsoft's stock price – since the portfolio was formed.

Alternatively, suppose the manager has flows of \$10 and Facebook and Microsoft both have returns of 0%. Scaling up each in proportion to lagged shares held, the manager holds \$55 of Microsoft and Facebook, again preserving value weights from the initial starting point of a 50-50 portfolio split. Finally, consider the case both where the manager receives \$10 of inflows (at the end of the quarter), Facebook has a return of 10% and Microsoft has a return of 0%. To preserve the initial value weights, the manager should hold \$60.24 of Facebook and \$54.76 of Microsoft.

A different way to explain this logic is to frame active management as running a float-adjusted index fund, where the manager has total discretion over the float-adjustment applied. In this example, in quarter $t-1$, the float-adjustment-factor applied to Facebook and Microsoft makes their market capitalizations equal, which is why the manager holds equal dollar amounts of both. We explore this in more detail in the next subsection.

Active Management & Preserving Value Weights

Active management is typically modeled as a decision over portfolio stocks and weights. Another way to view active management, which more naturally maps into our rebalancing framework, is a dynamic choice of float adjustments for every stock in the manager's investment universe. Under this logic, an active manager will behave just like an index fund, in the sense that they will hold a constant fraction of each stock's float, but the float adjustments are based on beliefs of future performance, not publicly available shares. And, when that manager receives inflows, to preserve value-weights *given their chosen float adjustments*, they would scale up positions in proportion to lagged shares held. Unlike a true index fund, however, an active manager can apply arbitrary changes to the float adjustments each quarter.

To clarify this point, Table 3 starts with the same investment universe as Table 1. Consider an active manager with \$6000 AUM who wants to hold relatively more Google stock than a VWIF, relatively less Netflix stock than a VWIF, and no Facebook stock. This can be achieved by setting Facebook's float to 0, trimming Netflix's float to 1/2 of its true shares outstanding, and increasing Google's float to 1.5 times its true shares outstanding. With these adjustments, the total float of the manager's investment universe is \$3,796,825. And with \$6000 in AUM, they should hold $6000/3796825 = 158$ basis points of each stock's float.

Now, suppose this same fund has a \$600 inflow (i.e., 10% of AUM) at the end of the quarter and all of the stocks have constant prices from quarter t to quarter $t+1$. If the manager wants to preserve value weights given their initial choice of float adjustments, they can scale up each position by 10% of the number of shares

Table 3: Example Active Fund, No Flows

Stock	$shrout_t$	Float Adj.	$float_t$	p_t	$capadj_t$	wt_t	$dheld_t$	$shares_t$	$pct.float_t$
FB	100,000	0	0	\$ 22.50	\$ -	0%	\$0	0	-
AAPL	35,000	1	35,000	\$ 57.35	\$ 2,007,250	53%	\$3,172	55	0.158%
AMZN	10,000	1	10,000	\$ 19.99	\$ 199,900	5%	\$316	16	0.158%
NFLX	20,000	0.5	10,000	\$ 36.23	\$ 362,300	10%	\$573	16	0.158%
GOOG	20,000	1.5	30,000	\$ 40.91	\$ 1,227,375	32%	\$1,940	47	0.158%
Total					\$ 3,796,825		\$6,000		

Notes. This table provides an example of a hypothetical active fund whose investment universe is the same as the value-weighted index fund in Table 1. The table provides the shares outstanding in the market (*shrout*), a float adjustment applied by the fund (*FloatAdj.*) which can be interpreted as another way of stock picking, the float ($float$), the price of the stock (p), the float-adjusted market cap (*capadj*), the portfolio weight (wt), the dollars and shares held (*dheld*, *shares*), and the percent of float (using float adjustments by the fund) held ($pct.float$). The table illustrates how an active fund's portfolio can be reinterpreted as picking a float adjustment.

Table 4: Example Active Fund, Flows but No Returns

				Pre-Flow		Post-Flow				
stock	$shares_t$	$float_t$	p_t/p_{t+1}	$dheld_{t+1}$	$pct.float_{t+1}$	$\Delta\$$	$\Delta shares$	$shares_{t+1}$	$dheld_{t+1}$	$pct.float_{t+1}$
FB	0	0	\$ 22.50	\$0	-	\$0	0	0	\$0	-
AAPL	55	35,000	\$ 57.35	\$3,172	0.16%	\$ 317.20	6	61	\$3,489	0.174%
AMZN	16	10,000	\$ 19.99	\$316	0.16%	\$ 31.59	2	17	\$347	0.174%
NFLX	16	10,000	\$ 36.23	\$573	0.16%	\$ 57.25	2	17	\$630	0.174%
GOOG	47	30,000	\$ 40.91	\$1,940	0.16%	\$ 193.96	5	52	\$2,134	0.174%
Total				\$6,000					\$6,600	

Notes. This table provides an example of a hypothetical active fund whose investment universe is the same as the value-weighted index fund in Table 1. The table provides the shares outstanding in the market (*shrout*), a float adjustment applied by the fund (*FloatAdj.*) which can be interpreted as another way of stock picking, the float ($float$), the price of the stock (p), the float-adjusted market cap (*capadj*), the portfolio weight (wt), the dollars and shares held (*dheld*, *shares*), and the percent of float (using float adjustments by the fund) held ($pct.float$). The table illustrates how an active fund's portfolio can be reinterpreted as picking a float adjustment. If the fund receives inflows, the fund would scale up positions proportionally and would own an equal fraction of the float for each stock.

they held at the end of the previous quarter. Table 4 shows this leads the fund to hold 17.4 basis points of each constituents' float-adjusted shares outstanding (15.8 + 1.58 = 17.4). Of course, the manager's actual portfolio could deviate from this because unlike a VWIF (1) an active manager does not need to preserve value weights and (2) the manager might change the float adjustments applied to each stock between quarter t and $t + 1$.

Finally, suppose that between quarter t and quarter $t + 1$, the stocks have heterogeneous returns. Table 5 shows a specific example of this, where after the returns have been realized, the fund has \$5800 in AUM, and the relative weights of all the stocks in their portfolio have changed. Then, suppose that as before, the fund has an inflow equal to 10% of AUM, and wants to preserve value weights given their initial float adjustment and the realized returns. Table 5 shows that even though the ex-post portfolio weights will be different, this fund will hold the same proportion of each stock's float as in Table 4, where the stocks had no changes in price.

Table 5: Example Active Fund, Flows and Returns

				Pre-Returns/Flow		Post-Returns/Pre-Flow			Post-Returns/Flow		
stock	$shares_t$	$float_t$	p_t	$dheld_t$	$pct.float_t$	p_{t+1}	$dheld_{t+1}$	$pct.float_{t+1}$	$shares_{t+1}$	$dheld_{t+1}$	$pct.float_{t+1}$
FB	0	0	$ 22.50	$0	-	$ 27.00	$0	-	0	$ -	-
AAPL	55	35,000	$ 57.35	$3,172	0.16%	$ 62.00	$3,429	0.16%	61	$ 3,772.10	0.174%
AMZN	16	10,000	$ 19.99	$316	0.16%	$ 15.00	$237	0.16%	17	$ 260.74	0.174%
NFLX	16	10,000	$ 36.23	$573	0.16%	$ 30.00	$474	0.16%	17	$ 521.49	0.174%
GOOG	47	30,000	$ 40.91	$1,940	0.16%	$ 35.00	$1,659	0.16%	52	$ 1,825.21	0.174%
Total				$6,000			$5,800			$ 6,380	

Notes This table provides an example of a hypothetical active fund whose investment universe is the same as the value-weighted index fund in Table 1. The table provides the shares outstanding in the market (*shrout*), a float adjustment applied by the fund (*FloatAdj.*) which can be interpreted as another way of interpreting stock picking, the float (*float*), the price of the stock (*p*), the float-adjusted market cap (*capadj*), the portfolio weight (*wt*), the dollars and shares held (*dheld*, *shares*), and the percent of float (using float adjustments by the fund) held (*pct.float*). The table illustrates how an active fund's portfolio can be reinterpreted as picking a float adjustment. If the fund receives inflows, the fund would scale up positions proportionally and would own an equal fraction of the float for each stock, even in the face of differences in realized returns across portfolio stocks.

Viewing active management as a choice over float adjustments is useful, as it clarifies why our methodology should apply to more than just passive funds. Any fund wishing to preserve value weights from a given initial portfolio allocation will scale up positions in proportion to lagged shares held. And, as shown above, this logic applies equally to active funds as passive funds.

Float Adjustments

As mentioned above, another factor that affects the interpretation of our intensive-margin rebalancing measure is changes in float adjustments applied by index providers. For example, consider a stock with 1,000 shares outstanding at the end of 2020q4. Further, suppose insiders purchase 50 shares before the end of 2021q1. On the 2021q1 reconstitution date (i.e., the date index funds update their index membership and weights), indices that exclude shares held by insiders from the float will adjust the eligible index shares down by 5%, and thus funds tracking these indices will need to sell. Similarly, if a firm does a secondary equity offering (SEO), index funds will need to buy some fraction of the issued shares. While these cases are important – as illustrated in the way they can contribute to large values of *rebalin* in Figure 2 – we do not incorporate float changes into our calculation of *rebalscale*. This is for several reasons. First is a practical one, which is that not every index applies the same float adjustment. Second, we believe that applying a float adjustment is an *active* decision by index funds to deviate from perfect scaling up and down of existing positions, and therefore should be classified as intensive-margin rebalancing.

Another argument for ignoring float adjustments is that intensive-margin rebalancing is describing portfolio adjustments that funds actually do. As a specific example of this, consider SPY in 2019 Q3. Throughout the quarter, Apple had done a significant amount of buybacks, which decreased the number of index-eligible shares. In addition, Oracle had issued a significant amount of stock as part of their employee compensation program (i.e., not through a secondary equity offering). So, at the end of the quarter when the index was reconstituted to reflect these changes, SPY had negative intensive margin rebalancing for Apple and positive intensive margin rebalancing for Oracle.

Comparison to Other Measures of Rebalancing

Our framework of linking changes in shares to lagged shares held similar to the regression used in Lou (2012):

$$\frac{\Delta shares_{i,j,t}}{shares_{i,j,t-1}} = \beta_0 + \beta_1 flow_{j,t} + \gamma_2 X + \gamma_3 flow_{j,t} X + \epsilon_{i,j,t} \tag{6}$$

where X are a set of stock-level characteristics. Ignoring these additional controls, and dividing Equation 1 through by $shares_{i,j,t-1}$, the fitted values from the regression in Equation 6 will be equal to *rebalscale* if (1) the constant term $\beta_0 = 0$ (2) $\beta_{j,t} = flow_{j,t}$ and β_1=1 i.e., all flows go to scaling and (3) we pool across all funds and all quarters. Conditions 1 and 2, however, do not hold empirically, as Table 2 of Lou (2012) shows that β_1 is below 1, especially in quarters with inflows.

Despite the fact that empirically not all flows go to scaling, we still believe *rebalscale* is a reasonable benchmark for a passive decision for a fund. This is because this is just a benchmark – of course, active funds (the focus of our paper and Lou (2012)) do not have to follow this. But it is an *option* that is always available to them. Further, in estimating the regression in Equation 6, Lou (2012) is pooling across all funds and all quarters. As we show, there is substantial heterogeneity in scaling behavior both within and across funds, which our methodology is flexible enough to accommodate.

Another related measure of fund rebalancing is fund-level turnover, used in Pástor et al. (2017). They follow the SEC's definition of turnover:

$$FundTurn_{i,t} = \frac{\min\left(buys_{i,t}, sells_{i,t}\right)}{\text{avg}\left(AUM_{i,t}\right)} \tag{7}$$

In words, *FundTurn* is trading *above and beyond* the minimum amount of trading a fund would need to do to remain fully invested each quarter in the fact of flows. At first glance, it's not clear how this mechanically relates to our measures of rebalancing because of the min operator. For example, suppose a manager receives a flow equal to 10% of AUM. They could either put this into new stocks, scale up existing positions, or a combination of both. If this is all the fund did with the inflow (i.e., the fund did not also sell some existing positions to buy even more new stocks), in both cases $FundTurn_{i,t}$ would be equal to zero. This highlights a distinct advantage of our methodology, as it will distinguish between the scaling up/down of existing positions from adding and dropping positions entirely.

Finally, while it is not a measure of turnover/trading, a natural point of comparison is active share from Cremers and Petajisto (2009). Active share is defined as the total deviation of a fund's portfolio weights from its benchmark weights. For example, a fund like SPY, which exactly tracks the underlying S&P 500 index, will have an active share of 0. For a fund that is benchmarked to the S&P 500 but holds 10 stocks, it will likely be labeled as having an active share of almost 1. Suppose, however, that when this active manager gets inflows, they just scale their existing positions up and down proportionally to how much they held previously. Compare this to when the manager sells all 10 of those stocks and buys 10 new stocks. This change may have little effect on active share, but we argue that there is significant information in this reshuffling of positions, relative to a scaling up and down of existing positions. A benefit of our methodology is that it will identify such changes.

2.6 Data

Since we are using changes in positions to construct our rebalancing measures, we employ an extensive data filtering and cleaning approach. In many ways, changes are less amenable to data issues since the errors often affect the level of positions. Since changes are small relative to levels, erroneously inferring position-level errors as position-level changes can systematically create noise that is an order of magnitude larger than most actual changes. We describe the data methodology below, but provide a comprehensive description and full accounting of our data methodology in Appendix B. The replication code can be found on the authors' websites.

We obtain mutual fund holdings data between 1980 and 2021 from Thompson S12. We merge the individual holdings to CRSP stock data using historical CUSIP. To ensure that holdings are comparable across time, we multiply the raw shares held by CRSP's cumulative factor to adjust shares, and divide prices by CRSP's cumulative factor to adjust prices. Following Greenwood and Sammon (2022), we exclude all stocks which are involved in a merger (either as the acquirer or acquiree) in either quarter t or quarter $t-1$. This filter is designed to remove large changes in shares held that result from delisting or share issuance to fund the merger, and are not driven by decisions at the fund level.

We use the CRSP MF Links database to link S12 fund identifiers ($fundno$) to CRSP fund identifiers ($crsp\_fundno$). Throughout the paper, fund index j identifies data at the S12 $fundno$ level, which groups together all the share classes associated with the individual $crsp_fundno$s. For example, AWSHX is associated with over 10 $crsp_fundno$s, while SPY is only associated with a single $crsp\_fundno$.

Following Barber et al. (2016), we estimate $flow_{j,t}$ as $\frac{AUM_{j,t}}{AUM_{j,t-1}} - (1 + r_{j,t})$ where AUM is fund j's assets under management and $r_{j,t}$ is the fund's total return.[6] Total AUM data comes from CRSP, and is aggregated across all share classes to the S12 $fundno$ level. $r_{j,t}$ is the weighted average return across all $crsp_fundno$s associated with a given S12 fundno, where the weights are proportional to lagged AUM.

While working with the S12 data, we noticed that there are instances where a fund will report holding shares of a given stock in quarter t, holding no shares of that stock in quarter $t + 1$, and holding the same number of shares that they held in quarter t in quarter $t + 2$. We treat these cases as data errors, as it seems unlikely that a fund would entirely drop a stock and then buy back the exact same number of shares one quarter later. We apply the following procedure to account for these cases.

First, we identify instances where a fund holds a stock, then does not hold that same stock, and finally holds the same stock again within 3 quarters. Then, we calculate the change in shares held between the gap in holdings (for most of these observations, the change in shares is zero). If the change in shares is less than 10%, we assume that the start of the gap in holdings is not a drop, and the end of the gap is not an add. But, we set $\Delta shares$ equal to missing for all these quarters, because we don't know what the fund did over this time period.

Our results are not sensitive to applying this specific procedure. We find similar results with two alternative specifications: (1) requiring that the change in shares between the seemingly missing quarters is exactly zero (2) rather than setting the change in shares to missing for all the associated quarters, assuming the trade happens in the last quarter of the gap.

After cleaning the data, we apply several filters at the fund quarter level which are standard in the literature. First, we aim to remove incubation bias, as discussed in Evans (2010), which we do by eliminating observations before the starting year for a fund reported by CRSP as well as the observations with a missing fund name in CRSP. We then further restrict to equity-only funds. We define such funds as those for which we can merge at least 90% of their holdings in s12 to CRSP and for which the value of their total S12 stock holdings is at least 66% of their total AUM in across all associated $crsp_fundno$s (Cremers and Petajisto, 2009). Next we remove small funds, requiring at least \$15M in AUM in 2015 dollars (Pástor et al., 2017). Finally, we identify passive funds following the procedure in Appel et al. (2016), although results are similar

[6] As discussed above, this estimate of flows relies on the assumption that all flows happen at the end of the quarter. One slight departure of our methodology from that in Barber et al. (2016) is that because our holdings data is quarterly, we compute our estimate of $flow_{j,t}$ using quarterly data on returns and AUM, rather than monthly data as in the original paper. Also note that here we use CRSP's TNA variable interchangeably with AUM.

only using the index fund flag in the CRSP mutual fund database.

To account for stale data, we drop fund quarter observations where the filing date (FDATE) increases, but the report date (RDATE) does not, indicating that the data has been carried forward and is likely stale. We also remove the first observation after a quarter of stale data to prevent the inference of quarterly changes when the changes reported are likely from more than one quarter ago. For example, for SPY, until 2008 Q1, holdings data are identical for each quarter within each year except for those filed in December. This is because the fund was only filing one report each year and Thompson forward fills holdings until the next new report is available. This is noticeable in the S12 data because the RDATE is the same for each FDATE for SPY in those years.

Another issue related to gaps between RDATEs is Thompson's re-use of S12 $fundno$s. The documentation explains that if a $fundno$ reappears with more than a year gap between RDATEs it is likely a new fund. So, we get rid of the first observation for a new fund (as everything would be classified as an addition) and the last observation for the old fund (as everything would be classified as a deletion). For the same reason, we also drop the overall first and last observation for every $fundno$.

Finally, we restrict to the subset of fund quarters with non-missing active share (Cremers and Petajisto, 2009), the data for which runs from 1980 to 2019. The authors provide data for the active share of each fund with respect to a variety of benchmarks e.g., the S&P 500, Russell 1000, etc. Each quarter, we take the benchmark which the fund has the minimum active share with respect to the benchmark which the fund is closest to in terms of portfolio weights. In our construction of benchmark-adjusted returns below, for consistency, this is the benchmark we use. In addition, to control for overall portfolio turnover, we restrict to fund-quarters with non-missing turnover in the CRSP mutual fund data (Pástor et al., 2017). After applying all our filters, we are left with 86,000 fund quarters, coming from roughly 3,300 active funds.[7]

2.7 Summary Statistics

Before moving on to the main empirical analysis, we want to present summary statistics on our rebalancing measures both to make sense of the magnitudes of our measures, as well as provide additional validation. To this end, in the spirit of Cremers and Petajisto (2009) and Pástor et al. (2017), we aggregate our measures to the fund/quarter level:

$$\overline{rebalin}_{j,t} = \sum |rebalin_{i,j,t}| \tag{8}$$

[7]This is of similar magnitude to past literature, as Cremers and Petajisto (2009) has 2,647 active funds between 1980 and 2003, while Pástor et al. (2017) have 3,126 active funds between 1979 and 2011.

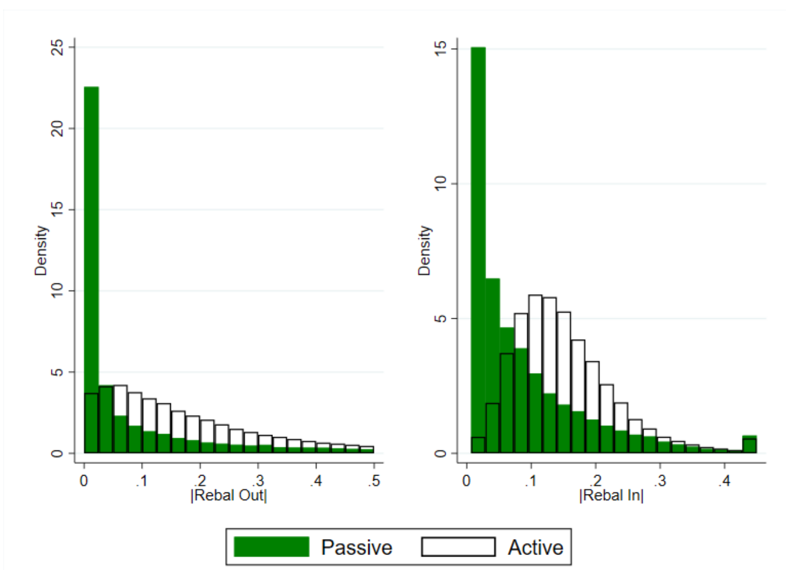
$$\overline{rebalout}_{j,t} = \sum |rebalout_{i,j,t}|. \tag{9}$$

In terms of magnitudes, a value of 0.1 would imply that the sum of all intensive or extensive-margin rebalancing was 10% of AUM. This is likely 2× the actual amount of the portfolio that is turned over, as adding a stock as 5% of AUM, and dropping a stock as 5% of AUM would lead to a total $\overline{rebalout}_{j,t}$ of 0.1.

The left panel of Figure 3 plots a histogram of $\overline{rebalout}_{j,t}$ for active funds in the clear bars, and for passive funds in the green bars. First, there is a striking difference between the active funds and passive funds, with the passive funds doing significantly less adding and dropping of stocks. The picture has been truncated at values of $\overline{rebalout}_{j,t}$ above 0.5, but the long tail is persistently heavier for active funds.

The right panel of Figure 3 mirrors the left panel, except it uses $\overline{rebalin}_{j,t}$. Again, passive funds do much less intensive-margin rebalancing. This is not surprising, however, as our *rebalscale* measure is designed to account for the mechanical scaling up and down that passive funds do, and as we saw in Figure 2, passive funds do little intensive-margin rebalancing aside from mechanical scaling.

Figure 3: Fund-Level Rebalancing Histograms



Notes. This figure provides two histograms, one for extensive-margin rebalancing ($|RebalOut|$) and one for intensive-margin rebalancing ($|RebalIn|$). The unit of observation is a fund-quarter. We aggregate stock-level measures of *rebalout* and *rebalin* by taking the absolute value and adding over all fund positions in a quarter.

3 Active Trading & Fund Performance

In this section, we test whether there is a relation between the measures of active trading defined in Section 2 and future stock-level performance. We find that positive intensive and extensive-margin rebalancing are associated with positive future alpha, while the same does not hold for negative intensive and extensive-margin rebalancing. We then conduct two sets of placebo tests. First, we show that these results do not hold for the active component of *passive* funds' rebalancing, suggesting that mechanical price pressure is not driving our baseline findings. We also show that our baseline regression results do not hold for the mechanical part of rebalancing associated with flows for active funds. This is further evidence that managers' active rebalancing decisions – rather than total changes in portfolio holdings – are the source of predictive power for future alpha in our baseline regressions.

3.1 Stock-Fund Level Results

We regress a measure of performance in quarter $t+1$ for each stock i held or dropped by fund j in quarter t on the active trading in that stock by that fund in that quarter. That is, we ask whether a given type of active trading by a fund in a stock is related to future relation. Specifically, we estimate

$$\begin{aligned}\alpha_{i,j,t+1} = {} & \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} \\ & + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t}.\end{aligned} \tag{10}$$

where $\alpha_{i,j,t+1}$ is the 4-factor alpha for stock i estimated using daily data in quarter $t+1$.[8,9] We include our four measures of active trading, capturing intensive-margin positive and negative trades ($rebalin_{i,j,t}^{positive}$ and $rebalin_{i,j,t}^{negative}$), and extensive-margin additions and deletions ($rebalout_{i,j,t}^{add}$ and $rebalout_{i,j,t}^{drop}$). The measures are *signed* (based on the direction of trade, so $rebalin_{i,j,t}^{negative}$ and $rebalout_{i,j,t}^{drop}$ can only take values less than

[8]There are a subset of observations for which we cannot compute a quarter $t+1$ alpha because the company delists. In Appendix A.4, we show that conservatively setting the alpha of these stocks to -100% does not change our main conclusions.

[9]We would like to highlight that $\alpha_{i,j,t+1}$ is the same for all managers holding stock i in quarter t. It does not vary at the manager j level. Therefore, there must exist a way to re-express the regression in Equation 10 at the stock-quarter level, i.e., the level of variation in the left-hand-side variable. This alternative aggregated version of the regression is not our preferred specification because of our weighting scheme. Specifically, within each fund-quarter, observations are weighted in proportion to how large the positions were relative to manager j's AUM. So, to construct this aggregated regression, when constructing stock-level average rebalancing, we would need to weight each part of that average based on how large those positions in stock i were to every manager's portfolio – all of which are different sizes. We believe that this weighting would complicate the interpretation of our results.

or equal to zero) and are mutually exclusive since a fund can only make one of the four types of trades in a given stock-quarter. In addition to future four-factor alpha, we also use benchmark-adjusted returns as a measure of performance, which is calculated as the difference between stock i's return in quarter $t+1$ and the return to fund j's benchmark index in quarter $t+1$.

The regression in Equation 10 also includes two additional variables that are associated with *fund* outperformance: active share (Cremers and Petajisto (2009), Petajisto (2013)) and fund turnover (Pástor et al., 2017).[10] Both capture a dimension of activeness at the fund-quarter level. Since they do not vary within a fund-quarter, they serve as controls for the overall activeness of a fund in a given quarter. That is, these controls will capture if a fund generates alpha not from rebalancing individual positions, but just from activeness in and of itself.

We estimate Equation 10 as a weighted regression, where stock i's weight is the market value of fund j's holding of stock i divided by total dollars in fund j in quarter t. For stocks that were dropped, we estimate their weight by taking the number of shares held by fund j in quarter $t-1$ and multiplying by stock i's price at the end of quarter t (consistent with our assumption that all trades happen at the end of the quarter). Total fund dollars, therefore, include held and dropped positions among stocks that have non-missing alpha in $t+1$ (i.e., weights for fund j in each quarter t sum to 1).

We acknowledge that the literature has largely focused on alpha and benchmark-adjusted return at the fund level or portfolios-of-funds level. Since we are interested in active trading and its timing, we need more granular measures of fund performance. Our 4-factor alpha for stock i is computed each quarter using daily returns regressed on daily excess market, SMB, HML, and momentum factors. This approach soaks up as much variation as possible through factors. An alternative approach, which may allow for larger residual variation, is to estimate factor loadings over quarter t and/or previous quarters, then use the loadings from t to compute an alpha for $t+1$. The latter approach does not account for how factor loadings may vary over time and may mistakenly categorize factor exposure as alpha. We think of our approach as conservative and working against finding spurious outperformance.

The regression in Equation 10 quantifies the unconditional relation between future performance and fund trading. We test two additional specifications, adding different sets of fixed effects, which change the interpretation of the results. First, we test whether *relative* activeness is associated with performance by including fund and quarter fixed effects. The baseline specification with no fixed effects asks whether more activeness

[10]For consistency with Pástor et al. (2017), we use annual turnover in the CRSP mutual fund database. Even though our estimates are using quarterly changes in holdings, to avoid a look-ahead bias, we use turnover at the end of the previous calendar year.

is associated with greater performance, which could be dominated by funds that are regularly more active than other funds.

By including fund and quarter fixed effects, we ask whether activeness relative to (1) how active a fund typically is and (2) how active all funds have been in the quarter is associated with performance. That is, it allows us to assess whether the relation between trading and performance holds for both a very active fund that becomes even more active and a less active fund that becomes somewhat active. The cost of this specification is that it introduces future information which is implicitly incorporated into the fund fixed effect. Specifically, if funds that do more active rebalancing on average have higher returns, that will be soaked up by the fund fixed effects.

The second additional specification includes fund-by-quarter fixed effects. This specification speaks to an even narrower question: whether a more actively rebalanced position outperforms a less actively rebalanced position *in the same fund and quarter*. This specification omits active share and portfolio-level turnover, which do not vary within a given fund-quarter.

We present the regression estimates in Table 6. We find evidence that larger additive trades in both the intensive- and extensive-margin are statistically related to greater future alpha. This pattern holds for unconditional additions (no fixed effects), additions that are large relative to fund and quarter averages (fund and quarter fixed effects), and additions that are large relative to other additions in the same fund and quarter (fund-by-quarter fixed effects). We also find evidence that dropping a stock is related to greater future alpha, i.e., a large drop misses out on future performance. The economic magnitudes are relatively large – adding 1% of total fund assets to a single existing stock position is associated with 36bps of future quarterly alpha in that stock, or almost 1.5% of annualized alpha. Adding 1% of total assets to a single new stock position forecasts 21bps of future quarterly alpha in that stock.

The regression estimates also show that there is no statistically significant relation between active trading and benchmark-adjusted returns. This suggests that the factors which motivate active rebalancing is a source of alpha, but not a source of beating benchmarks. While there is evidence that managers are concerned with performance relative to their benchmark (e.g., Roll (1992), Jorion (2003)), it may be difficult for managers to precisely link insights into, say, forecasts that a stock will beat earnings expectations with whether that information will lead to alpha or greater benchmark-adjusted return.[11]

[11]For example, Kothari and Sloan (1992) (and many others) have shown that there is an economically strong and statistically significant relation between earnings news and contemporaneous earnings-day returns. As shown in Laarits and Sammon (2022), however, the R-squared of these earnings response regressions is low (less than 10% even when saturated with controls and fixed effects), suggesting that it may be difficult to convert earnings information into alpha on a position-by-position basis.

Table 6: Active Rebalancing and Performance

	Alpha (t+1)			Bench Adj Ret (t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	36.328***	26.531**	22.042*	31.564	18.54	8.993
	(13.73)	(10.56)	(11.91)	(44.32)	(43.86)	(43.88)
Rebal In (negative)	-37.276	-22.429	-17.351	-5.311	4.428	10.88
	(23.89)	(15.97)	(16.62)	(27.61)	(31.18)	(36.80)
Rebal Out (adds)	21.584**	18.855***	17.719***	22.313	20.378	14.626
	(9.04)	(6.28)	(6.28)	(25.48)	(24.39)	(23.63)
Rebal Out (drops)	-16.055**	-16.818***	-17.039***	3.267	3.483	5.589
	(7.38)	(5.27)	(5.44)	(6.90)	(7.32)	(9.43)
Active Share	59.113	-27.153		-17.232	87.217	
	(41.07)	(59.10)		(34.37)	(68.72)	
Turnover	4.149	6.606		-2.562	-0.919	
	(5.41)	(4.08)		(6.92)	(4.72)	
Observations	10,073,284	10,073,284	10,073,284	10,073,284	10,073,284	10,073,284
R-squared	0	0.01	0.048	0	0.008	0.065
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for our sample of active funds. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 for more details.

Table 6 suggests that active funds' purchasing activity is informative about future performance: the more an active fund buys a stock, the better that stock does. However, this does not *necessarily* imply funds themselves will have greater future performance for several reasons. First, our estimates assume that a fund holds positions for a full quarter. That is, we assess how a stock would have contributed to a fund's alpha if that fund did not do any subsequent trading. Second, our sample has a selection bias in that we condition on funds that survived (in order to build a meaningful time series at the fund level), and these funds are more likely to have purchased and held well performing stocks.

We do note that while selection bias may increase the likelihood of alpha on average, it does not necessarily increase the association of a fund's trades with future outperfomance in those stocks. Further, in order to be included in our tests, we require a stock have at least a month of daily return data in quarter $t + 1$ to calculate alphas. This leads to a minor survivorship bias, as stocks that delist in the first month of quarter $t + 1$ are omitted.[12] Lastly, we do not account for fees. In Section 5, we account for all these concerns by aggregating trading activity at the fund level.[13]

In addition, we wish to clarify that although the regression weights by the size of the position in the fund, the regression identifies whether variation in *rebalancing activity* is associated with alpha. While the weighted regression does allow for more weight to be placed on large positions, the variation being exploited is that of the size of the trade, not the size of the *position*. This is different from past work on managers' "best ideas," which is focused on the performance of a fund's largest positions (see, e.g., Cohen et al. (2009) and Pomorski (2009)). The correct interpretation of our findings is that large rebalancing decisions signal larger future alphas.

The results in Table 6 give the impression that there are alphas everywhere, i.e., there are alphas after both positive rebalancing and negative rebalancing. We wish to emphasize three points. First, the regressions capture the relation between *trades* and future alpha, not *positions* and future alpha. Second, one might be concerned that for every share bought by an active fund, it is sold by another active fund. In this case, it must be that positive alpha after buying is matched by positive alpha after selling. In Section 3.2.4, we explain why this is unlikely to explain our results. Finally, we would like to highlight that this is not necessarily alpha earned by fund investors. We explicitly examine portfolio-level alphas in Section 5.

If active buying decisions are informative about future performance, it suggests that active managers possess

[12]In the Appendix, we show that a conservative approach of assigning a return of -100% to such stocks does not affect any of our conclusions.

[13]Wermers (2000) provides additional reasons why a fund that holds outperforming stocks may not necessarily deliver greater fund-level performance.

some skill in adjusting their portfolio holdings. One alternative story is that buying leads to price pressure, mechanically leading to higher returns.[14] We have already shown evidence in Table 6, however, which is inconsistent with this price pressure story. For example, if price pressure was the main explanation for our results, we might expect all the coefficients on our rebalancing measures to be the same sign, which we do not find. In fact, drops have a negative and statistically significant coefficient, suggesting that selling leads to subsequent reversal. In the next subsection, we directly test the price pressure story using data on passive funds.

3.2 Ruling Out Mechanical Explanations

To further rule out price pressure or other stories that generate a mechanical relation between trading in quarter t and returns in $t+1$, we use two placebo settings: passive funds and predictable trading from passive rebalancing. We also discuss why two other possible mechanical mechanisms – autocorrelation in fund rebalancing and trading among active funds – cannot explain our main results.

3.2.1 Passive Funds

We repeat the tests from Section 3.1 but on the sample of stocks traded by index mutual funds. We use the same measures of performance and trading at the stock-fund-quarter level. If a price pressure story or mechanical relation drives the link between trading and future performance, we should expect to see regression estimates that are similar for passive and active funds.

Table 7 shows that trading from passive funds is distinct from that of active funds. The estimates on all four trading measures are mostly statistically insignificant, with the only exception coming from extensive-margin additions (which indicate there is reversal, possibly due to the index addition effect see; e.g., see Greenwood and Sammon (2022)). Moreover, most of the coefficients are of the opposite sign compared to Table 6, providing further evidence against a mechanical trading story.

[14] There are several papers documenting price pressure effects of mechanical buying pressure by mutual funds (see e.g., Coval and Stafford (2007), Lou (2012), Khan et al. (2012)). Recall that in Table 6 we are looking at the rebalancing decisions of *individual funds*. There is some evidence (see e.g., Wermers (1999)) that buying is not highly correlated across funds, allaying this concern. We examine the effects of aggregate rebalancing decisions across funds in Appendix A.6. The stock-quarter aggregation results are inconsistent with mechanical price pressure driving our main results.

Table 7: Placebo Test: Index Fund Rebalancing and Performance

	Alpha (t+1)			Bench Adj Ret (t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	-54.17	-34.546	-55.004	-84.835*	-57.887	-93.474
	(39.90)	(32.90)	(38.79)	(51.07)	(53.95)	(78.26)
Rebal In (negative)	39.3	6.218	12.74	39.859	11.262	12.818
	(44.39)	(31.45)	(42.83)	(62.13)	(68.23)	(91.72)
Rebal Out (adds)	-16.674**	-13.245	-16.784**	-18.685	-15.026	-19.794
	(7.44)	(8.85)	(8.44)	(12.05)	(12.72)	(14.81)
Rebal Out (drops)	-4.845	-6.297	-2.105	1.174	-1.133	4.87
	(7.98)	(8.15)	(8.80)	(9.21)	(10.39)	(13.75)
Active Share	12.673	30.418		-15.261	219.355**	
	(25.41)	(81.38)		(18.32)	(102.58)	
Turnover	-1.002	1.075		0.519	-3.428	
	(3.70)	(5.61)		(4.91)	(6.77)	
Observations	8,920,334	8,920,334	8,920,334	8,920,334	8,920,334	8,920,334
R-squared	0	0.011	0.071	0	0.006	0.081
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for our sample of index funds. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.2 for more details.

Table 8: Placebo Test: Active Fund Predictable Rebalancing and Performance

	Alpha (t+1)			Bench Adj Ret (t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal Scale (positive)	77.343*	33.984	30.862	53.96	28.638	23.012
	(46.38)	(31.82)	(43.62)	(47.25)	(56.13)	(67.20)
Rebal Scale (negative)	-11.936	-22.75	0.832	-13.765	-23.776	29.356
	(23.36)	(18.81)	(39.41)	(35.19)	(44.19)	(92.67)
Active Share	67.081*	-15.422		-14.939	90.97	
	(39.74)	(58.44)		(33.30)	(61.97)	
Turnover	7.31	9.155**		-0.949	0.377	
	(6.26)	(4.23)		(8.36)	(5.22)	
Observations	10,073,284	10,073,284	10,073,284	10,073,284	10,073,284	10,073,284
R-squared	0	0.01	0.048	0	0.008	0.065
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalscale_{i,j,t}^{positive} + \beta_1^n rebalscale_{i,j,t}^{negative} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t}$$

for our sample of active funds. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on predicted positive and negative predicted rebalancing (scaling positions in proportion to shares held). The rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.2 for more details.

3.2.2 Passive Rebalancing

In Section 2.2, we introduced our measures of active rebalancing as the component of portfolio changes that is distinct from the type of scaling up and down that a passive fund would do in response to flows. Flows and predictable trading may lead to price pressure and a mechanical relation between returns and trading (Lou, 2012). As an additional placebo test, we repeat our tests using the set of active mutual funds but regress performance measures on the *predictable* component of trading by mutual funds, or $rebalscale_{i,j,t}$ given in Equation 3. In these tests, we ask whether the predictable trading by an active fund in response to flows is also related to performance.

Table 8 shows that the predictable component has, at best, a weak statistical relation with future performance. These findings further support the notion that it is not trading in response to flows that is associated with future alpha but rather how funds deviate from predictable scaling. This suggests that the predictable component of trading, which is a function of what funds held in the previous quarter, is related to future performance but is not precise enough to be statistically reliable.

3.2.3 Autocorrelation in Active Rebalancing

One might be concerned that the results in Table 6 are due to persistence in active rebalancing decisions. For example: if positive active rebalancing in quarter t predicts positive active rebalancing in quarter $t+1$, it's possible that the higher returns in quarter $t+1$ are due to the continued buying by the fund itself, rather than information about future fundamentals. In subsection 4.5, we show that while such active rebalancing decisions are autocorrelated, the magnitude of these effects are economically small. Further, managers tend to do more active rebalancing in stocks they've been rebalancing actively recently, regardless of the direction. For example, when a firm has *positive* intensive-margin rebalancing at 1% of AUM in quarter t, this predicts *negative* margin rebalancing in quarter $t+1$ of 11 basis points of AUM in the same stock. This makes it even more unlikely that a fund's own price pressure is driving our main results.

As a robustness test, we estimate our baseline regressions but using alpha in quarter $t+2$, $t+3$, and $t+4$. We find that only the quarter immediately following the active rebalancing has alphas that are significantly associated with active rebalancing. This suggests that active rebalancing has a relatively short horizon. It does not necessarily suggest that *funds* have a short horizon, as funds may have long-term expectations for performance but engage in active rebalancing when they expect it to materialize. See Appendix A.7 for the regression estimates and more details.

3.2.4 Trading Among Active Funds

Finally, one might be concerned that the results of positive alpha both after positive intensive/extensive rebalancing *and* negative intensive/extensive margin rebalancing might be mechanical, due to trading among funds. At first pass, this seems supported by the coefficients on positive and negative rebalancing being of similar magnitudes.

To illustrate this possible mechanism, suppose some Manager A has positive rebalancing for a given stock in quarter t which generates alpha in quarter $t+1$. Manager A must have bought those shares from somewhere, and one possibility is that they bought it from another active manager. So, conditional on having positive alpha after e.g., portfolio additions, one might mechanically expect positive alpha to occur after deletions as well.

First, we would like to highlight that this logic is predicated on the result itself. Suppose, for example, that every share bought by one active manager is actually sold by another active manager. In this world, the coefficients on positive and negative rebalancing would be mirror images of each other. The magnitude of these coefficients, however, would still be an empirical question. In fact, it's possible that in this scenario, the coefficients on all the rebalancing measures were zero.[15]

Second, because we are taking out the flow-driven part of rebalancing, there is not a mechanical adding up constraint where every dollar of positive rebalancing is matched by a dollar of negative rebalancing. In fact, a key source of identifying variation in Equation 10 is differences in flows across funds.

Suppose, for example, Fund A has flows of 5% of AUM, while Fund B has outflows of 5% of AUM. Suppose further that both funds have the same initial AUM, but Fund A holds 10 stocks equal-weighted, while Fund B holds 2 stocks equal-weighted. Fund B perfectly scales down its existing positions in response to the flows, selling $5\% \times 0.5 = 2.5\%$ of AUM in each of its existing positions. On the other hand, Fund A buys all of the shares sold by Fund B in these two stocks. Then, for the 2 stocks held by Fund B, Fund A will have positive intensive margin rebalancing, while for the other 8 stocks, Fund A will have negative intensive margin rebalancing. So, in this instance, even in a world with no aggregate flows into active mutual funds *and* active mutual funds only trading with each other, net intensive margin rebalancing is not zero.

Finally, we believe that every share bought by one active manager is not necessarily sold by another active manager. One reason for this is that aggregate flows into the active mutual fund sector is rarely zero. For example, when aggregate flows are positive, active mutual funds have to buy shares from outside the universe of other active funds. Further, *even if* aggregate flows are zero, it still could be that active funds are, e.g., providing liquidity to passive funds and/or retail investors. In this case, the coefficients on positive and negative rebalancing would not need to be the same magnitude.

4 Mechanisms and Drivers of Active Rebalancing

In this section, we conduct several additional tests to better understand the relation between active rebalancing and future stock performance. We start by performing a series of sample splits on flows, embedded gains, holding periods, number of holdings and fund size. We then examine the relation between active rebalancing and earnings surprises to determine if the relation between rebalancing and alpha is due to information about future earnings announcements. Finally, we run a series of regressions to evaluate the importance of other

[15]This, however, oversimplifies our baseline regression in Equation 10, in a way that makes the mechanical explanation seem more concerning than it really is. Within each fund-quarter, each observation is weighted by the fraction of each fund's AUM in that position. So, the adding up constraint in the example above implicitly relies on the funds being the same size and/or the trades being the same size as a fraction of the funds' AUM.

motivations for trade like transaction costs, past returns, and recent rebalancing.

4.1 Active Trading and Flows

In this subsection, we re-examine our main regressions of Section 3.1 by splitting the sample into fund-quarters with inflows and outflows. Flows require that a fund trade to either allocate incoming dollars or raise cash. We present two possibilities for how a fund may choose to actively rebalance – distinct from passively scaling up/down positions – in response to flows.

The first is that a fund may decide to engage in discretionary trading that it would have not otherwise done without flows. For example, a fund may decide to buy a stock only because it has cash on hand, i.e., it would not have bought that stock if it had to sell existing positions. In this scenario, flows could lead to excessive trading and should not be associated (or as strongly associated) with future performance.

Another possibility is that a fund has a list of stocks it would like to purchase, but faces portfolio management constraints. For example, a fund may want to buy a stock it predicts will outperform, but does not have cash on hand and/or faces complexities in adjusting its existing portfolio to make room for this new stock (e.g., large expected price impact or transaction costs). Here, flows would allow the fund to engage in trades that lead to greater future performance.[16]

Table 9 shows the results of the regression in Equation 10, splitting the sample into fund-quarters with positive and negative flows. The regressions show that most of the main results in Table 6 are being driven by fund-quarters with inflows. This is consistent with funds facing portfolio adjustment constraints which are relaxed by the arrival of inflows.

Specifically, the predictive power of both positive intensive and extensive-margin rebalancing are stronger – in terms of both economic magnitude and statistical significance – in the inflows sample (columns 1-3) than the outflows sample (columns 4-6). This is despite the fact that nearly two-thirds of the observations come from fund-quarters with outflows, i.e., it is unlikely to be driven by an issue with statistical power. A natural concern with these results is that skilled managers are more likely to receive inflows, creating a mechanical difference between the inflow and outflow sample. The pattern holds, however, even in the presence of fund fixed effects, at least partially alleviating this concern.

[16] For example, Alexander et al. (2007) show that manager buying in the face of outflows predicts high returns going forward. Our methodology is different, however, because in identifying buying, Alexander et al. (2007) does not account for the expected scaling down of *individual positions* in the face of flows. We would argue that even keeping a position fixed in the face of outflows is evidence that the manager believes that the stock will outperform.

Table 9: Trading, Performance and Flows

	Positive Flows			Negative Flows		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	53.697***	37.734**	34.602**	24.832*	17.868*	13.401
	(18.10)	(14.73)	(16.68)	(13.05)	(10.00)	(10.78)
Rebal In (negative)	-51.051	-19.356	-22.496	-28.173	-19.309	-14.16
	(35.60)	(21.85)	(24.90)	(17.89)	(14.04)	(13.76)
Rebal Out (adds)	35.882***	23.977***	23.090**	11.813*	13.486**	13.921**
	(13.34)	(8.70)	(8.91)	(6.80)	(5.23)	(5.49)
Rebal Out (drops)	-28.049**	-19.553**	-16.877*	-12.908**	-15.015***	-16.496***
	(13.45)	(8.38)	(8.72)	(6.16)	(4.95)	(4.99)
Active Share	12.56	-14.635		81.091**	-9.669	
	(58.91)	(73.73)		(37.65)	(58.78)	
Turnover	13.066	11.086*		-3.56	2.291	
	(8.29)	(5.97)		(4.61)	(4.24)	
	3,904,632	3,904,632	3,904,632	6,168,652	6,168,652	6,168,652
	0	0.016	0.053	0	0.01	0.045
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$\alpha_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for our sample of active funds in fund-quarters with inflows vs. outflows. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. α is quarterly alpha of stock i, where alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Alpha is in quarterly basis points. We regress the return measure on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 4.1 for more details.

4.2 Embedded Gains and Holding Periods

To better understand factors that are associated with active rebalancing and returns, we split our sample based on a measure of embedded gains and holding periods.

Embedded Gains Yet another motivation for trade may be the realization of profits on existing positions. For example, fund managers may sell stocks when they have outperformed and reached their internal price targets. On the other hand, behavioral factors may come into play when considering how embedded gains affect decisions to trade. For example, managers may be hesitant to sell positions with embedded losses, as realizing the loss feels worse than leaving the loss unrealized in their portfolio (Barberis and Xiong, 2012).[17] To proxy for unrealized profits, we use a measure of embedded gains, which we compute for each stock-fund based on the quarter that the stock was added to the fund's portfolio.

Because we do not know exactly when the fund purchased each individual position, we need to rely on several assumptions to compute the embedded gains on each stock they hold. Recall that in constructing *rebalscale*, *rebalin* and *rebalout*, we assume that all trading happens at the end of each quarter. We apply the same logic to compute embedded gains, assuming that the initial cost basis for an added stock is equal to the end of quarter price in the quarter the stock was added. In all quarters where shares held stays the same or decreases, the cost basis is held constant. If shares held increases, the cost basis is adjusted to be a weighted average of the last cost basis and the end of quarter price, where the weights are proportional to how much was previously held vs. how much was newly purchased. Under this construction, embedded gains will not account for dividends received but only capital appreciation.

Holding Periods Another factor in deciding whether to trade may be how long a stock has been in the fund's portfolio, i.e., do managers hold positions for "too long." If funds hold positions for longer periods of time, their active trading may be more incremental rather than a signal that the fund expects high returns in the following quarter. On the other hand, long holding periods could be indicative of continuous information gathering, and therefore one might expect trading in these positions to have more predictive power for future returns than trading in positions they have not held for a long time.

We again regress future alpha on our active rebalancing measures but we split the sample using both embedded gains and holding periods. We create four samples by splitting stocks with (1) embedded gains/losses

[17] Other behavioral factors which may lead embedded gains to affect rebalancing decisions include the disposition effect (Frazzini (2006), An et al. (2022)) and loss aversion/mental accounting Barberis and Huang (2001).

and (2) holding periods above/below the median (e.g., half the stocks in each fund will be split into a low and high holding period group). The splits are done independently. In Appendix A.2, we provide the results for each variable in a one-way split and find similar results.[18] These splits are done within a fund-quarter, so we must make adjustments to our baseline regressions. First, we re-weight positions at the fund-quarter level such that each fund-quarter receives an equal weight within each subsample – which mirrors the weighting scheme in the baseline regressions. Second, we eliminate the rebalancing measure that captures extensive-margin additions because these trades by definition do not have embedded gains and have a holding period of zero.

We report the findings on the relation between active trading and embedded gains in Table 10. First, we would like to highlight that the results for positive intensive-margin rebalancing are stronger in the subsamples with embedded gains than embedded losses, and even more so for positions with short holding periods. One explanation for this difference is that managers have superior information about stocks with embedded gains. Therefore, when managers add further to these positions, it is more likely to be based on information and therefore predictive of future alpha.

Another difference between the subsamples is that managers appear to have worse timing in dropping stocks with embedded gains than embedded losses. Specifically, columns 1 and 2 suggest that if a manager exits a position with embedded gains which was 1% of AUM, that stock is predicted to have about 25 basis points higher alpha the following quarter. Compare this to column 4 which restricts to positions with embedded losses and short holding periods, where the relation is no longer statically significant and the economic magnitude is cut by about half. In addition, trimming positions with positive gains, especially when the holding period is short, is associated with particularly good future performance. This is consistent with evidence on the disposition effect: managers may sell out of positions with embedded gains too quickly, and hold on to losing positions for too long.

The relation between positive intensive-margin rebalancing and future alpha is weakest for stocks with long holding periods and embedded losses – places where managers may be doubling down on poor investments instead of realizing losses. This is consistent with behavior motivated by the law of small numbers, as discussed in Jin and Peng (2023). More broadly, columns 3 and 4 show that active rebalancing in positions with embedded losses is not predictive of future returns regardless of the investment horizon.

[18] In Appendix A.3, we show that the empirical patterns are different when splitting on the portfolio average holding period at the *fund level*, rather than within each fund's holdings. Specifically, we provide evidence that managers seem to specialize in trading at different horizons.

Table 10: Sample Split: Embedded Gains and Holding Period

	Alpha (t+1)			
	Embedded Gains		Embedded Losses	
	Long Holding Per.	Short Holding Per.	Long Holding Per.	Short Holding Per.
	(1)	(2)	(3)	(4)
Rebal In (positive)	32.492**	47.449**	22.1	34.441*
	(15.14)	(22.84)	(15.80)	(18.14)
Rebal In (negative)	-19.494	-87.626**	-5.053	-28.744
	(24.88)	(38.55)	(19.38)	(20.68)
Rebal Out (drops)	-21.884**	-27.566**	-12.754	-13.207
	(9.79)	(11.04)	(10.58)	(9.81)
Active Share	148.313***	88.478*	-127.586*	-84.279
	(42.13)	(51.63)	(72.23)	(52.84)
Turnover	-58.545*	-0.16	-1.633	3.659
	(31.76)	(6.84)	(7.24)	(6.86)
Observations	2,910,056	2,685,231	1,250,478	1,867,847
R-Squared	0.000	0.000	0.000	0.000
Fixed Effects	None	None	None	None

Notes. This table shows coefficients from estimating a weighted regression:

$$\begin{aligned} Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} \\ + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t}, \end{aligned}$$

for four samples formed by splitting our baseline sample of active funds. We assign stocks within a fund-quarter into subsamples based on (1) if the stock has a positive or negative embedded gain, and (2) if the stock is or is not above the median holding period within a fund-quarter. The weights in the regression are equal to the size of the stock as a fraction of AUM of all stocks in fund j and quarter t for that subsample. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section 4.2 for more details.

4.3 Portfolio Count and Fund Size

To further explore the relation between active rebalancing and returns, we reexamine our sample by splitting on two measures of fund size: portfolio count, i.e., the number of holdings in a funds' portfolio, and AUM, i.e., the overall dollar size of the fund.

Portfolio Count We start by considering how many stocks are held by each fund. Holding many stocks may be a signal that a fund is looking to minimize tracking error or engages in more systematic sector or style bets. Holding fewer stocks may signal that a fund is targeting more stock-specific performance. Moreover, a single large trade for a fund that holds 10 stocks may reveal and signal much more than a single large trade for a fund that holds hundreds or thousands of stocks.

AUM Next, we turn to the size of the fund, where many similar arguments could apply. It seems intuitive that skilled managers end up with more AUM than unskilled managers (Berk and Green (2004)). And, even if such managers are shadow indexing a significant portion of their AUM, their marginal trades may be more likely to predict future alpha than managers with less skill. On the other hand, managers with low AUM may be better suited to target micro-cap or small-cap stocks where there may be more opportunities. A fund with large AUM may avoid such stocks because they are not scalable for large trades and positions.

Just as above with embedded gains and holding periods, we split the main sample into four separate samples based on portfolio count and AUM. For portfolio count, we split all funds each quarter based on whether the number of stocks held is above or at/below the median portfolio count for that quarter. Similarly, we independently split the main sample on whether a fund is above or at/below the median AUM in that quarter. We split independently on both variables because portfolio count and fund size may be somewhat mechanically related in the sense that large funds will have more holdings on average. These characteristics, however, have potentially different explanations for performance. For example, if size is an indicator of manager skill, a large fund with few holdings may be focused on stock-specific bets, while a large fund with many holdings may be focused on factor bets. Therefore, from the perspective of future alpha – which by construction removes future factor performance – one might expect the trades of large funds with few holdings to have the strongest predictive power for future alpha.[19]

The estimates from regressing alpha on each of the four rebalancing measures are provided in Table 11.

[19] We provide the results from splitting on each variable separately in Appendix A.2. The takeaway from that analysis is similar.

First, we can see that intensive-margin rebalancing trades have strong predictive power for future alpha among funds with fewer holdings across both small and large funds. This is consistent with funds with fewer holdings focusing their information on stock-specific information for a small set of securities, while funds with many holdings may struggle to tactically reshuffle their existing positions. One explanation for this is that in funds with many holdings, managers' attention may be limited, therefore making intensive-margin rebalancing more difficult when holding more stocks.

On the other hand, extensive-margin rebalancing additions have much stronger predictive power for future alpha in funds with many holdings, and especially so for high AUM funds. This difference survives the inclusion of fund and year fixed effects, making it less likely that a Berk and Green (2004)-type mechanism – where skilled managers amass more AUM and therefore more positions – is driving these results. One interpretation is that if a fund holds many stocks, it spends more time focusing on which new stocks to buy rather than how to optimally position each stock within the fund. A fund that holds few stocks may be able to more carefully rebalance its existing holdings based on how its outlook has changed and thus, the decision to add a stock may be the first step in building a position.

For the extensive margin, completely dropping a position has stronger predictive power for future returns for high AUM funds – with both a large and small number of holdings – than low AUM funds. This could be due to price impact since large funds are likely to exhibit greater price pressure from trading, say, 1% of its fund.

4.4 Active Rebalancing and Earnings Information

Our baseline results in Table 6 show a clear link between active trading – large purchases in particular – and future alpha. One possible reason why large trades predict future alpha is that funds have discovered information about firm fundamentals; once that information is publicly revealed, funds enjoy high returns (Baker et al., 2010).

To test this information channel as a possible explanation for the baseline results in Table 6, we examine whether future earnings surprises are associated with our active rebalancing measures. To quantify earnings surprises, we construct a standardized unexpected earnings (SUE) variable, which is equal to realized earnings-per-share minus the consensus earnings-per-share forecast (in the last IBES statistical period before earnings are actually released), divided by the price the day before the earnings announcement. This gives a measure of the surprise earnings-to-price ratio, and is used in Hartzmark and Shue (2018). Following Glosten

Table 11: Sample Split: Number of Holdings and Fund Size

	Alpha (t+1)			
	High AUM		Low AUM	
	High Holdings	Low Holdings	High Holdings	Low Holdings
	(1)	(2)	(3)	(4)
Rebal In (positive)	33.625*	37.860**	20.151	43.696***
	(20.29)	(16.06)	(20.47)	(14.15)
Rebal In (negative)	-60.227	-38.217*	-45.974	-30.332*
	(48.16)	(23.04)	(36.60)	(17.41)
Rebal Out (adds)	37.438**	13.14	26.080*	20.014**
	(15.91)	(8.46)	(13.55)	(7.85)
Rebal Out (drops)	-40.598**	-19.639**	-12.51	-10.802*
	(17.82)	(8.34)	(8.96)	(6.48)
Active Share	45.556	42.213	53.804*	139.677*
	(30.55)	(99.00)	(30.27)	(72.72)
Turnover	12.06	16.019	-2.882	2.285
	(14.24)	(16.39)	(4.25)	(5.36)
Observations	4,099,415	922,764	3,645,132	1,405,973
R-Squared	0	0	0	0
Fixed Effects	None	None	None	None

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for four samples formed by splitting our baseline sample of active funds. We assign all stocks within a fund-quarter into subsamples based on (1) if the fund is or is not above the median AUM that quarter, and (2) if the fund is or is not above the median number of stocks held. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section 4.3 for more details.

et al. (2021), we also decompose our SUE measure into a systematic and idiosyncratic component, where we regress the firm-level SUE on the market-wide SUE and 2-digit SIC industry SUE in five-year rolling windows. The systematic component is the predicted value from this regression, and the idiosyncratic component is the residual.

Table 12 shows a link between active rebalancing and future earnings surprises. Except for the specification without any fixed effects, the table shows that there is a positive relation between adding stocks on the intensive and extensive margins and future earnings surprises. The magnitudes are relatively large. For example, the median SUE is around 0.04 and the mean is around -0.05; an intensive-margin purchase of 1% of AUM leads to an earnings surprise that is roughly half of the magnitude of the median and mean SUE. In addition, the results are a bit economically and statistically stronger for the idiosyncratic component of earnings surprises. The estimates for extensive-margin additions are similar, though the magnitudes are a bit weaker. We take this as suggestive evidence that funds may buy stocks in anticipation of future positive fundamental news, and that funds focus their research primarily on firm-specific information as opposed to more systematic or industry-related information.

4.5 Other Drivers of Active Rebalancing

So far, we have focused primarily on studying active trading and future returns. In this subsection, we take a different approach: we ask which stock-level characteristics are related to managers' active rebalancing decisions. This is partly motivated by an acknowledgement that not all active rebalancing (as we have already seen) is associated with future returns and may, instead, be driven by contemporaneous or lagged variables. While it is more difficult to make claims about predictability, the regressions nonetheless provide useful insight into possible other drivers of active rebalancing besides future returns.

For these analyses, we replace alpha as the dependent variable in our regressions with each of the four active rebalancing measures. In each regression, we restrict to observations where that measure of rebalancing is not equal to zero. E.g., the first column of Table 13 restricts to observations where intensive margin rebalancing is positive.

We regress each of our measures on a combination of variables, organized by category. The first category uses return-related variables, both contemporaneous and lagged quarterly returns and alphas. The second category is lagged active rebalancing measures and lagged and contemporaneous fund flows. These regressions are meant to capture persistence in active rebalancing decisions and the role flows play in determining the

Table 12: Active Rebalancing and Future Earnings Surprises

	SUE (t+1)			Idio SUE (t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	-0.005	0.016**	0.026**	0.078*	0.022*	0.029**
	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)	(0.02)
Rebal In (negative)	-0.035***	-0.053***	-0.068***	-0.099***	-0.047***	-0.056***
	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)
Rebal Out (adds)	0.007*	0.011**	0.014***	0.040***	0.011**	0.015**
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Rebal Out (drops)	0.003	0.001	-0.001	-0.025***	-0.002	-0.005
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
Active Share	-0.123***	-0.083***		-0.164***	-0.072***	
	(0.02)	(0.03)		(0.04)	(0.03)	
Turnover	0.006***	0.004*		0.029***	0.003	
	(0.00)	(0.00)		(0.01)	(0.00)	
Observations	8,265,777	8,265,777	8,265,777	8,265,777	8,265,777	8,265,777
R-squared	0	0.014	0.045	0.001	0.189	0.211
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$SUE_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for our sample of active funds. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. SUE is a measure of standardized unexpected earnings (SUE) using the surprise in earnings announcements normalized by price (a surprise earnings-to-price ratio) or the idiosyncratic component of SUE (Glosten et al., 2021). We regress the SUE measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section 4.2 for more details.

rebalancing measures themselves. Third, we utilize a combination of portfolio-specific stock-level variables based on holding histories: embedded gains, the embedded gains rank of a stock in the portfolio, and the holding period in number of quarters. Similar to Section 4.2, these variables allow us to see if funds are more likely to rebalance towards or away from stocks with gains and longer holding histories. For the final category, we use a measure of price impact, though the results are similar when substituting this with the effective bid-ask spread. Table 13 presents the results from regressing our four active rebalancing measures on each of the four categories of variables.

We describe several consistent themes by category.

Returns The first column of Table 13 shows that positive intensive margin rebalancing is not very responsive to contemporaneous or past returns – although the coefficient on returns at t is positive and significant, the magnitude is relatively small, especially relative to the estimates in columns 2 to 4. On the other hand, the negative coefficient on contemporaneous returns in column 2 shows that funds are more likely to rebalance away from a position if it has had high returns in the recent past. Column 3 shows that funds are more likely to add a stock which has gone up in the recent past – at least in terms of raw returns – but not in terms of recent past alpha. Finally, column 4 shows that funds are also more likely to drop a stock entirely if it has gone up in the recent past, consistent with column 2.

This loosely implies a type of "lifecycle" for the average stock: the position is initially added with high returns, and that position is trimmed or ultimately dropped completely with sufficient future high returns. This may be consistent with the stock reaching a manager's internal price target, or may be due to well documented behavioral biases, e.g., the utility investors have from realizing embedded gains (Barberis and Xiong, 2012).

Lagged Rebalancing Measures and Flows The first general pattern we would like to highlight is the relationship between flows and rebalancing. Column 1 shows that funds are more likely to add to positions above and beyond what is predicted by flow-based rebalancing when flows are larger. Column 2, however, shows that there is not a strong relationship between contemporaneous flows and trimming positions more than flows alone would predict. Finally, column 4 shows that managers are less likely to drop stocks entirely in the face of inflows (and symmetrically, they are more likely to drop stocks entirely in the face out outflows).

Next we would like to argue that intensive margin rebalancing is persistent, however the sign of recent rebalancing is not as important. For example, in column 1, the coefficient on past positive rebalancing is positive, while the coefficient on past negative rebalancing is negative. This implies that managers are more likely to add to the stock above and beyond what is predicted by flows if they were doing *any* kind of intensive margin rebalancing (positive or negative) in the recent past. Similarly, in column 2 the coefficient on past positive intensive margin rebalancing is negative, and past negative intensive margin rebalancing is positive – again suggesting the recent rebalancing in either direction predicts more rebalancing in that stock going forward. Finally, column 4 shows that managers are more likely to drop stocks entirely that they've either been adding to or trimming from in the recent past.

In addition, we would like to highlight that, as discussed in Section 3.2, these results are further evidence against mechanical price pressure driving our main results, because we do not see uniform autocorrelation in either positive or negative intensive margin rebalancing.

Embedded Gains and Holding Periods Table 13 also shows that active rebalancing is associated with embedded gains and holding periods, and the effect is somewhat asymmetric. For the intensive margin, funds are more likely to sell positions with larger embedded gains by percentage and rank within a portfolio and less likely to add to them. This is broadly consistent with the behavioral finance literature (e.g., Barberis and Xiong (2012), Odean (1998) and Frazzini (2006)). Further, funds are more likely to drop stocks with greater percentage embedded gains and stocks that have had relatively lower gains than the rest of their portfolio. Lastly, stocks with longer holding periods have more intensive-margin rebalancing (both positive and negative) and are less likely to be dropped completely.

Liquidity The final set of regressions use price impact, a measure of illiquidity, as a regressor. Table 13 shows that rebalancing is associated with price impact on the intensive margin: funds do more rebalancing when price impact is smaller. On the extensive margin, larger drops are weakly associated with lower price impact. New additions, on the other hand, tend to have higher price impact. Overall, these regressions suggest that funds are responsive to liquidity except when adding new stocks.

5 Fund-Level Analysis

Our results in Section 3 show that, at the individual trade level, adding stocks predicts future alpha. This is true both for adding to existing positions more than predicted by flows (i.e., on the intensive margin), and entering into new positions. While the magnitude for intensive-margin rebalancing is higher, the effect

Table 13: Drivers of Active Rebalancing

		Rebal In (pos) (1)	Rebal In (neg) (2)	Rebal Out (add) (3)	Rebal Out (drop) (4)
Returns	Ret (t)	0.0288***	-0.118***	0.145***	-0.338***
		(0.007)	(0.020)	(0.032)	(0.051)
	Ret (t-1)	0.0111	-0.0711***	0.0118	-0.181***
		(0.008)	(0.014)	(0.026)	(0.041)
	Alpha(t)	0.00604	-0.00145	0.00926	-0.0772*
		(0.007)	(0.018)	(0.027)	(0.042)
	Alpha(t-1)	0.00322	-0.0219*	0.00574	-0.0758**
		(0.008)	(0.013)	(0.022)	(0.032)
	Observations	2,865,322	2,245,419	822,014	810,686
	R-squared	0.209	0.242	0.554	0.503
Past Rebalancing and Flows	Rebal In (pos,t-1)	0.124***	-0.271***		-1.182***
		(0.008)	(0.008)		(0.020)
	Rebal In (neg,t-1)	-0.112***	0.170***		0.118***
		(0.007)	(0.007)		(0.012)
	Rebal Out (add,t-1)	0.0910***	-0.0357***		-0.267***
		(0.003)	(0.003)		(0.011)
	Flow (t)	0.106***	0.0424**		0.549***
		(0.016)	(0.018)		(0.025)
	Flow(t-1)	0.0555***	0.0672***		0.152***
		(0.009)	(0.013)		(0.017)
	Observations	2,865,322	2,245,419		810,686
	R-squared	0.229	0.264		0.545
Gains & Holding Per.	Gains (pct)	-0.00835***	-0.00395**		-0.0164***
		(0.002)	(0.002)		(0.005)
	Gains (rank)	-0.0203***	-0.0727***		-0.239***
		(0.007)	(0.004)		(0.012)
	Holding Period	0.000157	-0.00110***		-0.00528***
		(0.000)	(0.000)		(0.000)
	Observations	2,865,322	2,245,419		810,686
	R-squared	0.21	0.246		0.508
Liquidity	Price Impact	-0.792***	1.565***	-8.646***	8.435***
		(0.200)	(0.359)	(1.275)	(1.464)
	Observations	2,865,322	2,245,419	822,014	810,686
	R-squared	0.209	0.236	0.553	0.489

Notes. This table shows coefficients from estimating an equal-weighted regression of one of our four *rebal* measures on a host of variables separated into four categories: returns, past rebalancing and flows, embedded gains and holding periods, and liquidity. We estimate separate regressions of each *rebal* variable on each set of variables within a category. We use our baseline sample of active funds. The four rebalancing measures are: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). In each column, we restrict to the subset of observations where each of the respective rebalancing measures are not equal to zero. The active rebalancing measures are signed, and in units of a percent of AUM. The variables for the returns category are current and lagged returns and alphas (in decimal i.e., 1%=0.01); for past rebalancing and flows, the variables are lagged rebalancing measures and current and lagged fund flows; for embedded gains and holdings, the variables are embedded gains in percent and in percentile rank within a portfolio (with 100 being the highest) and holding period in quarters; for liquidity the variable is price impact as measured in Holden and Jacobsen (2014). The regressions all include fund and quarter fixed effects, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section 4.5 for more details.

is slightly more precisely measured for extensive-margin purchases. Further, this holds both when the fund is receiving inflows and being forced to meet redemptions.

In some ways, funds appear to be less skilled in selling decisions. When dropping stocks completely, the larger the drop, the *better* that stock typically does. Intensive margin selling is a bit more complicated. While more selling is associated with better future performance, there are two caveats: this is not completely negative in that intensive margin selling still implies that the fund still holds at least some position in the stock (and thus benefits from high future alpha), and these estimates are not statistically significant. With this in mind, we now use these findings to aggregate our individual trade measures into an aggregated measure at the fund level.

We construct two measures of active trading at the fund level: total active buying, or $rebal^{add}_{j,t}$, and total active selling, or $rebal^{drop}_{j,t}$. We combine the two measures associated with buying because they tell a similar story (analogously with selling). We define our measures as the sum of all buying and selling respectively, or

$$\begin{aligned} rebal^{add}_{j,t} &= \sum_i w_{i,j,t} rebalin^{positive}_{i,j,t} + \sum_i w_{i,j,t} rebalout^{add}_{i,j,t} \\ rebal^{drop}_{j,t} &= \sum_i w_{i,j,t} rebalin^{negative}_{i,j,t} + \sum_i w_{i,j,t} rebalout^{drop}_{i,j,t}, \end{aligned} \tag{11}$$

where $w_{i,j,t}$ is the weight of stock i in fund j in quarter t following Section 3.1 (which assigns a weight for dropped stocks as well). We take a weighted average to better mimic the stock-fund-quarter findings. Economically, our measures capture the value-weighted average activity, with more weight placed on activity in larger positions.[20]

We regress measures of fund performance on $rebal^{add}_{j,t}$ and $rebal^{drop}_{j,t}$. Our measures of fund performance are $alpha_{j,t+1}$ (the value-weighted individual position-level alphas in quarter $t+1$ based only on positions held at the end of quarter t)[21], $rmb^{net}_{j,t+1}$ and $rmb^{gross}_{j,t+1}$ (the net and gross fund return minus the benchmark return in quarter t). We also include fund-quarter-level active share (Cremers and Petajisto, 2009) and turnover (Pástor et al., 2017) in the regressions. This allows us to control for known measures of fund activeness and helps articulate whether our measures capture something different from what was previously documented in the literature.

[20] Measures constructed via a simple sum or using an equally-weighted average are not strongly associated with future performance as $rebalout^{add}_{i,j,t}$ and $rebal^{drop}_{j,t}$. This suggests that managers' performance comes disproportionately from activity in its largest positions. That is, large trades in relatively small positions may not impact fund performance as much as smaller trades in large positions. We illustrate the mechanics of $rebal^{add}_{j,t}$ and $rebal^{drop}_{j,t}$ through a series of examples in Appendix A.5.

[21] To clarify, this is not the alpha for fund investors (Grinblatt and Titman, 1989). This is the hypothetical alpha the fund investors would receive if the fund did not trade and perfectly scaled up and down positions throughout quarter $t+1$.

Section 3.1 discussed why our regressions do not necessarily imply higher fund level returns. We address those concerns here. The net and gross returns we use in these tests are actual returns that fund investors received as reported by the fund in the CRSP mutual fund database. Regressing fund performance in $t+1$ on active trading measures on t exactly tests whether more activity forecasts higher *fund* returns. This is also why these regressions are not simply a tautology given the results in Table 6 – we are in essence aggregating up the right-hand side variables from Section 3.1, but replacing the left-hand side variables with actual realized fund returns. We also include specifications with and without fund and quarter fixed effects to test whether activity in general and/or activity relative to a fund's average level forecasts future performance.

We present the regression estimates in Table 14. The regressions show that funds have greater alpha next quarter when they engage in more active purchasing. This is somewhat unsurprising given the previous findings at the stock-fund-quarter level. In addition, the coefficients on additive rebalancing for both the net and gross benchmark-adjusted returns are positive but statistically insignificant for many specifications, but are significant with the inclusion of only fund fixed effects (specifications 6 and 11). This suggests that stock-picking ability may in fact accrue to actual fund investors through future fund performance but with some caveats.

First, funds deliver higher benchmark-adjusted returns when funds are more actively adding stocks relative to how active a fund typically is (specifications with fund fixed effects in columns 6 and 10). Second, it is important to control for how active a fund is on average – specifications without fixed effects (columns 5 and 9) yield insignificant coefficients. Third, when active funds are unusually active, they tend to generate alpha in the same quarters. This is supported by specifications with quarter fixed effects (columns 7/8 and 11/12), which also lead to statistically insignificant estimates.

The coefficient magnitudes on additive active rebalancing are relatively large. For positive rebalancing that is about 1% across the fund, investors receive a benchmark-adjusted return of between 40 and 90 bps per quarter, or 1.5% to over 3.5% per year.

Lastly, we note that more active selling activity is not associated with future benchmark-adjusted returns (positive or negative). The other two measures of fund activeness, Active Share and Turnover, are also mostly not associated with future returns. The exception is Active Share in the presence of quarter fixed effects (columns 7-8 and 11-12). We would like to note, however, that Active Share does not vary much within a fund over time. Thus, taking out fund and quarter fixed effects leaves little residual variation (which is why the estimated slope is relatively large). What little variation is left is associated with higher future returns, though it is unlikely that this residual variation captures the spirit of Active Share as described in

Cremers and Petajisto (2009).

As a placebo test for these results, we also repeat the analysis above but with the set of passive index funds. Table 15 provides the regression estimates. The table confirms that our measures of active trading are not associated with future alpha for passive funds, again ruling out a purely mechanical explanation for our findings. If anything, more positive active rebalancing is associated with negative benchmark-adjusted returns (Panels B and C).

We also note that Turnover appears positive and statistically significant. Unlike active share, Turnover can vary within a fund over time, even for passive funds. For example, if a passive fund receives many days of inflows and outflows, the fund will have to engage in both buying and selling, increasing the Turnover variable (see Section 2.2 for more details on how Turnover is defined). The regression estimates suggest that unusually high turnover is associated with high future performance for passive funds. Given the definition of Turnover and the fact that we control for our more active measures of trading, Turnover could be capturing scaling trades or passive rebalancing and may inadvertently be describing an association between (the persistence of) flows and future performance.

6 Conclusion

In this paper, we develop a new way to decompose changes in active managers' portfolio positions into three pieces. The first is a passive benchmark, which is the expected scaling up and down of shares held in proportion to lagged shares held in response to flows. The second is intensive-margin rebalancing, which quantifies active deviations from perfect scaling up and down among stocks the manager held in both the current quarter and previous quarter. The third is extensive-margin rebalancing, which accounts for new additions to and total deletions from a manager's portfolio. We argue that separating passive scaling from position changes is crucial when evaluating the rebalancing decisions of active managers.

Our main empirical finding is that positive active rebalancing predicts returns going forward. Specifically, both positive intensive-margin rebalancing and portfolio additions predict high stock-level alpha going forward. We show that the same does not hold when (1) using data exclusively from passive mutual funds/ETFs, and (2) focusing on the passive part of rebalancing – i.e., expected scaling in response to flows. This allays concerns that our results are driven by mechanical price pressure, which should apply equally to passive funds and passive scaling as it does to active rebalancing.

Table 14: Fund-Level Rebalancing and Performance

	Panel A: Alpha (t+1)			
	(1)	(2)	(3)	(4)
Rebal Add	41.732*	85.418***	5.821	31.527**
	(22.421)	(25.088)	(14.672)	(13.958)
Rebal Drop	-14.865	-63.858*	21.622	-16.019
	(32.399)	(34.745)	(15.382)	(14.626)
Active Share	71.284	-87.826	127.904***	-25.007
	(44.239)	(100.004)	(32.439)	(57.471)
Turnover	1.994	14.900**	0.293	5.833*
	(4.851)	(6.145)	(3.537)	(2.992)

	Panel B: Bench Adj. (t+1), Net			
	(5)	(6)	(7)	(8)
Rebal Add	62.253	92.649**	41.463	56.119
	(41.988)	(46.799)	(31.576)	(34.679)
Rebal Drop	25.202	-17.079	38.810*	4.584
	(28.182)	(28.417)	(22.272)	(22.875)
Active Share	1.495	45.050	31.171	100.308*
	(27.904)	(57.462)	(27.590)	(57.744)
Turnover	-8.187	1.327	-8.906	-3.622
	(5.138)	(4.761)	(5.435)	(4.110)

	Panel C: Bench Adj. (t+1), Gross			
	(9)	(10)	(11)	(12)
Rebal Add	66.394	94.740**	44.194	56.998
	(42.081)	(46.787)	(31.613)	(34.651)
Rebal Drop	21.842	-18.516	36.223	3.659
	(28.236)	(28.412)	(22.252)	(22.870)
Active Share	15.931	45.475	48.806*	103.820*
	(27.846)	(57.379)	(27.506)	(57.692)
Turnover	-6.449	2.132	-7.585	-3.459
	(5.139)	(4.756)	(5.451)	(4.116)
Fixed Effects	None	Fund	YQ	YQ/Fund

Notes. This table shows coefficients from estimating

$$Return_{j,t+1} = \beta_1^a rebal_{j,t}^{add} + \beta_1^d rebal_{j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{j,t},$$

for our sample of active funds. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against SMB, HML, and UMD using daily data in quarter $t+1$, then aggregated using portfolio weights based on holding at the end of quarter t. Benchmark-adjusted return is the return of fund j minus the total return of the fund's benchmark index. We use benchmark-adjusted returns net and gross of fees. All return measures are in quarterly basis points. We regress the return measures on two measures of aggregated active rebalancing. $rebal^{add}$ is the AUM-weighted average of positive rebalancing measures on the intensive and extensive margin. $rebal^{drop}$ is the AUM-weighted average of negative rebalancing measures (on both margins). The aggregated active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). We include specifications with no fixed effects, with fund fixed effects, with quarter fixed effects, and fund and quarter fixed effects. Standard errors are double clustered at the fund and quarter level. See Section 5 for more details.

Table 15: Placebo Test: Fund-Level Rebalancing and Performance, Index Funds

	Panel A: Alpha (t+1)			
	(1)	(2)	(3)	(4)
Rebal Add	-24.745	19.915	-29.854	-1.849
	(23.759)	(29.042)	(28.314)	(35.682)
Rebal Drop	1.719	-28.567	22.112	0.412
	(30.077)	(35.697)	(28.773)	(31.756)
Active Share	9.271	286.100	17.875	34.635
	(25.261)	(186.462)	(15.970)	(73.362)
Turnover	-2.128	-5.198	0.772	-1.447
	(3.417)	(5.913)	(3.116)	(5.015)
	Panel B: Bench Adj. (t+1), Net			
	(5)	(6)	(7)	(8)
Rebal Add	-75.175*	-35.495	-75.318*	-44.453
	(40.084)	(46.271)	(41.768)	(47.860)
Rebal Drop	-12.109	-35.988	-10.552	-27.912
	(42.314)	(46.173)	(43.322)	(46.638)
Active Share	-36.970**	-2.689	-32.904**	87.525
	(16.153)	(53.251)	(15.448)	(62.073)
Turnover	1.139	35.349**	-0.461	30.740**
	(6.844)	(16.623)	(7.020)	(15.427)
	Panel C: Bench Adj. (t+1), Gross			
	(9)	(10)	(11)	(12)
Rebal Add	-73.321*	-34.050	-73.942*	-43.693
	(40.047)	(46.259)	(41.743)	(47.855)
Rebal Drop	-11.550	-37.010	-9.817	-28.533
	(42.320)	(46.138)	(43.346)	(46.605)
Active Share	-32.785**	0.452	-26.385*	91.164
	(16.193)	(53.457)	(15.504)	(61.811)
Turnover	6.833	35.702**	5.076	30.961**
	(6.851)	(16.593)	(7.024)	(15.389)
Fixed Effects	None	Fund	YQ	YQ/Fund

Notes. We regress three measures of stock returns in quarter $t+1$ on four aggregated measures of active trading by fund j in quarter t over the set of passive index funds. The three measures of performance are (1) the value-weighted position-level alpha, (2) the net benchmark adjusted fund return, and (3) the gross benchmark adjusted fund return (all in basis points). The Rebal Add is the value-weighted average purchase size as a percent of AUM (both on the intensive- and extensive-margins). Similarly, Rebal Drop is the value-weighted average sell size. We also include two measures of fund activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). We include specifications with no fixed effects, with fund fixed effects, with quarter fixed effects, and fund and quarter fixed effects. Standard errors are double clustered at the fund and quarter level. See Section 5 for more details.

Next, we explore our results further through a series of sample splits. We find that the relation between positive rebalancing and future alpha is much stronger in quarters with inflows than in quarters with outflows. We argue this could be a symptom of portfolio adjustment constraints. Specifically, if a manager needs to sell stock to make room for a new addition, they may face both transaction costs and price impact. This is relaxed in the face of inflows, when the manager can add a stock without needed to do any other trading. We also show that our results are strongest among positions with embedded gains and funds that hold fewer stocks in their portfolio.

We also study the mechanism behind our results – *why* do fund managers do better in stocks they rebalance towards? We study the relation between positive rebalancing and future earnings surprises and find evidence that manager rebalancing is positively related to future good news for firm fundamentals.

Finally, we aim to understand whether our stock-fund-quarter-level results also apply when aggregated to the fund-quarter level. We find that larger positive active rebalancing is associated with higher future alphas, as well as higher future gross and net benchmark-adjusted fund returns when controlling for fund-level average activeness. These effects are economically larger than the stock-fund-quarter-level results, but they are less statistically significant.

Aggregating to the fund-quarter level speaks to whether total rebalancing in a given quarter predicts future returns. It does not, however, speak to the long-run returns to fund investors. Suppose, for example, that opportunities only arise infrequently. Then, investors may be paying fees for poor performance most quarters, and only receiving returns in excess of the benchmark and fees in the few quarters where opportunities arise. We believe that future work can explore whether our measures of active rebalancing can be useful in identifying managers capable of delivering long-run performance to investors in the active mutual fund space.

To this end, our paper provides two additional contributions that we hope will be useful for future research. First is our detailed data methodology for utilizing the entire position-level Thomson Reuters Mutual Fund (S12) dataset. We document and, to the best of our ability, address the numerous data issues that arise from using position-level data.

The second contribution is our empirical methodology for separating passive scaling from position-level holdings changes, leaving a position-level measure of active rebalancing by fund managers. This methodology is built on the empirical observation that passive equity index funds and ETFs typically respond to flows by perfectly and proportionally scaling their existing positions. In addition to our empirical methodology, this passive rebalancing benchmark itself may be useful in further studying position-level activity of mutual funds and ETFs.

References

Ahn, S. J., Arif, S., and Kaplan, Z. (2023). Neglected firms: Evidence from positions funds do not hold. *Available at SSRN*.

Akepanidtaworn, K., Di Mascio, R., Imas, A., and Schmidt, L. (2021). Selling fast and buying slow: Heuristics and trading performance of institutional investors. Technical report, National Bureau of Economic Research.

Alexander, G. J., Cici, G., and Gibson, S. (2007). Does motivation matter when assessing trade performance? an analysis of mutual funds. *The Review of Financial Studies*, 20(1):125–150.

An, L., Engelberg, J., Henriksson, M., Wang, B., and Williams, J. (2022). The portfolio-driven disposition effect. *PBCSF-NIFR Research Paper*.

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

Avramov, D. and Wermers, R. (2006). Investing in mutual funds when returns are predictable. *Journal of Financial Economics*, 81(2):339–377.

Baker, M., Litov, L., Wachter, J. A., and Wurgler, J. (2010). Can mutual fund managers pick stocks? evidence from their trades prior to earnings announcements. *Journal of Financial and Quantitative Analysis*, 45(5):1111–1131.

Baks, K. P., Metrick, A., and Wachter, J. (2001). Should investors avoid all actively managed mutual funds? a study in bayesian performance evaluation. *The Journal of Finance*, 56(1):45–85.

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.

Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In *Handbook of behavioral economics: applications and foundations 1*, volume 1, pages 79–175. Elsevier.

Barberis, N. and Huang, M. (2001). Mental accounting, loss aversion, and individual stock returns. *The Journal of Finance*, 56(4):1247–1292.

Barberis, N. and Xiong, W. (2012). Realization utility. *Journal of Financial Economics*, 104(2):251–271.

Barras, L., Scaillet, O., and Wermers, R. (2010). False discoveries in mutual fund performance: Measuring luck in estimated alphas. *The Journal of Finance*, 65(1):179–216.

Berk, J. B. and Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6):1269–1295.

Berk, J. B. and Van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of financial economics*, 118(1):1–20.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Chen, H.-L., Jegadeesh, N., and Wermers, R. (2000). The value of active mutual fund management: An examination of the stockholdings and trades of fund managers. *Journal of Financial and quantitative Analysis*, 35(3):343–368.

Cohen, R. B., Polk, C., and Silli, B. (2009). Best ideas. In *AFA 2011 Denver Meetings Paper*.

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

Cremers, K. M. and Petajisto, A. (2009). How active is your fund manager? a new measure that predicts performance. *The Review of Financial Studies*, 22(9):3329–3365.

Edelen, R. M. (1999). Investor flows and the assessed performance of open-end mutual funds. *Journal of Financial Economics*, 53(3):439–466.

Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance*, 65(4):1581–1611.

Fama, E. F. and French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance*, 65(5):1915–1947.

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, 61(4):2017–2046.

Giglio, S., Liao, Y., and Xiu, D. (2021). Thousands of alpha tests. *The Review of Financial Studies*, 34(7):3456–3496.

Glosten, L., Nallareddy, S., and Zou, Y. (2021). Etf activity and informational efficiency of underlying securities. *Management Science*, 67(1):22–47.

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

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

Grinblatt, M. and Titman, S. (1989). Mutual fund performance: An analysis of quarterly portfolio holdings. *Journal of Business*, pages 393–416.

Hartzmark, S. M. (2015). The worst, the best, ignoring all the rest: The rank effect and trading behavior. *The Review of Financial Studies*, 28(4):1024–1059.

Hartzmark, S. M. and Shue, K. (2018). A tough act to follow: Contrast effects in financial markets. *The Journal of Finance*, 73(4):1567–1613.

Hirshleifer, D. and Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1-3):337–386.

Holden, C. W. and Jacobsen, S. (2014). Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *The Journal of Finance*, 69(4):1747–1785.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, 23(2):389–416.

Jin, L. J. and Peng, C. (2023). The law of small numbers in financial markets: Theory and evidence. *Available at SSRN 3066369*.

Jorion, P. (2003). Portfolio optimization with tracking-error constraints. *Financial Analysts Journal*, 59(5):70–82.

Kaniel, R., Lin, Z., Pelger, M., and Van Nieuwerburgh, S. (2023). Machine-learning the skill of mutual fund managers. *Journal of Financial Economics*, 150(1):94–138.

Khan, M., Kogan, L., and Serafeim, G. (2012). Mutual fund trading pressure: Firm-level stock price impact and timing of seos. *The Journal of Finance*, 67(4):1371–1395.

Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006). Can mutual fund "stars" really pick stocks? new evidence from a bootstrap analysis. *The Journal of Finance*, 61(6):2551–2595.

Kothari, S. P. and Sloan, R. G. (1992). Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics*, 15(2-3):143–171.

Laarits, T. and Sammon, M. (2022). The retail habitat. *Available at SSRN 4262861*.

Lan, C., Moneta, F., and Wermers, R. (2018). Holding horizon: a new measure of active investment management. *Journal of Financial and Quantitative Analysis*, pages 1–80.

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

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

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance*, 53(5):1775–1798.

Pástor, L. and Stambaugh, R. F. (2002). Investing in equity mutual funds. *Journal of Financial Economics*, 63(3):351–380.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2017). Do funds make more when they trade more? *The Journal of Finance*, 72(4):1483–1528.

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

Petajisto, A. (2013). Active share and mutual fund performance. *Financial Analysts Journal*, 69(4):73–93.

Pomorski, L. (2009). Acting on the most valuable information:'best idea'trades of mutual fund managers. *Available at SSRN 1108186*.

Roll, R. (1992). A mean/variance analysis of tracking error. *Journal of Portfolio Management*, 18(4):13.

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

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *The Journal of Finance*, 54(2):581–622.

Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance*, 55(4):1655–1695.

Do Active Funds Do Better in What They Trade?

Marco Sammon and John J. Shim

INTERNET APPENDIX

A Robustness

A.1 Inactivity vs. Activeness

Our measure of active rebalancing presented in Section 2.2 uses deviations from passive scaling. For example, if there is an outflow, we would predict that the fund would scale down all existing positions in proportion to shares held if it behaved like a passively-managed index fund. If it were an active fund, it may decide to actively deviate by selling more than expected of some stocks and selling less or even buying some shares of other stocks. In this example, if a fund were to not trade a particular stock it held, it would effectively be an active buying decision.

We acknowledge that positive intensive-margin rebalancing that involves, say, buying more than or selling less than expected is a relatively straightforward signal that the fund views the stock positively (or at least more positively than other stocks it holds and could have held). When a fund does nothing, the signal is a bit more ambiguous. One alternative explanation is that the stock was one that was left out of consideration due to limited attention (Hirshleifer and Teoh, 2003).

To address this concern, we re-estimate our baseline regression specification described in Section 3.1 while including an indicator variable for inactivity. Specifically, we define the variable $noact_{i,j,t}$ which takes the value of 1 if fund j's holding in stock i was positive and unchanged from $t-1$ to t. We interact $noact_{i,j,t}$ with our two intensive-margin rebalancing measures as well to identify possible differences in the relation between active trading and future returns when the "activity" was deviating from expected by not trading. Table 16 shows that the results are qualitatively unchanged when including this measure of inactivity, as well as the interaction terms of inactivity and rebalancing. That is, the highlighted rows show that the interaction terms are insignificant, suggesting that the relation between active trading and future alpha does not come from what we deem are active decisions but materialize through no actual trading.

Table 16: Active Rebalancing and Inactivity

	Alpha (t+1)			Bench Adj Ret (t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	36.611***	26.798**	22.208*	32.37	19.09	9.34
	(13.96)	(10.73)	(11.91)	(44.94)	(44.24)	(43.98)
Rebal In (negative)	-34.997	-21.783	-17.663	-2.933	5.705	12.041
	(22.97)	(15.36)	(16.64)	(26.71)	(30.51)	(37.26)
Rebal In (positive) x No Change	20.71	6.46	27.95	-20.00	-22.88	-35.26
	(40.91)	(32.74)	(39.40)	(50.37)	(51.37)	(57.53)
Rebal In (negative) x No Change	-69.42	-25.08	-13.72	-56.01	-26.04	-20.42
	(49.47)	(33.11)	(35.16)	(37.07)	(37.61)	(32.13)
Rebal Out (adds)	21.938**	19.064***	18.113***	22.368	20.311	14.419
	(9.19)	(6.32)	(6.43)	(25.32)	(24.27)	(23.78)
Rebal Out (drops)	-16.394**	-17.006***	-17.449***	3.214	3.568	5.911
	(7.48)	(5.23)	(5.48)	(6.77)	(7.26)	(9.63)
Active Share	57.262	-27.716		-17.864	87.174	
	(41.92)	(59.25)		(34.59)	(68.97)	
Turnover	4.358	6.663		-2.465	-0.897	
	(5.47)	(4.07)		(6.88)	(4.67)	
Observations	10,073,284	10,073,284	10,073,284	10,073,284	10,073,284	10,073,284
R-squared	0	0.01	0.048	0	0.008	0.065
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$
\begin{aligned}
Return_{i,j,t+1} = {} & \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} \\
& + \beta_1^p ncrebalin_{i,j,t}^{positive} \times nochange_{i,j,t} + \beta_1^n ncrebalin_{i,j,t}^{negative} \times nochange_{i,j,t} \\
& + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},
\end{aligned}
$$

for our sample of active funds. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). We also interact *rebalin* with *nochange*, or an indicator variable that is 1 when the fund did not change shares held in a position. The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.1 for more details.

A.2 Individual Sample Splits

In Sections 4.2 and 4.3, we independently split on two variables which forms four samples from the main sample. Here, we conduct splits individually on each of the variables used in Sections 4.2 and 4.3. For more details on why splitting on these variables provide valuable insight, please see the main text.

Embedded Gains Here we estimate regressions split at the stock based on positive and negative embedded gains. See Section 4.2 for how we construct our measure of embedded gains. This splits stocks within a fund-quarter, so we must make adjustments to our baseline regressions. Just as in the main text, we make two adjustments: (1) we re-weight positions at the fund-quarter level such that each fund-quarter receives an equal weight within each subsample, and (2) we eliminate the rebalancing measure that captures extensive-margin additions because these trades by definition do not have embedded gains.

We report the findings on the relation between active trading and embedded gains in Table 17. First, we would like to highlight that the results for positive intensive-margin rebalancing are stronger in the subsample with embedded gains than embedded losses. One explanation for this difference is that managers have superior information about stocks with embedded gains. Therefore, when managers add further to these positions, it is more likely to be based on information and therefore predictive of future alpha.

Another difference between these two subsamples is that managers appear to have worse timing in selling stocks with embedded gains than embedded losses. Specifically, column 1 suggests that if a manager exits a position with embedded gains which was 1% of AUM, that stock is predicted to have 20 basis points higher alpha the following quarter. Compare this to column 4, which restricts the sample to positions with embedded losses and shows the relation is no longer statically significant and the economic magnitude is half as large. This is consistent with evidence on the disposition effect: managers may sell out of positions with embedded gains too quickly, and hold on to losing positions for too long.

Holding Period We again split our baseline sample into two slices based on holding period. If funds hold positions for longer periods of time, their active trading may be more incremental rather than a signal that the fund expects high returns in the following quarter. On the other hand, long holding periods could be indicative of continuous information gathering, and therefore one might expect trading in these positions to have more predictive power for future returns than trading in positions they have not held for a long time.

To this end, we split stocks within each fund's portfolio each quarter into those with above and below

Table 17: Sample Split: Embedded Gains

	Alpha (t+1)					
	Embedded Gains			Embedded Losses		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	42.364**	30.999**	27.781*	23.690*	9.534	10.276
	(18.24)	(14.99)	(16.61)	(12.78)	(11.30)	(13.63)
Rebal In (negative)	-38.923	-17.414	-13.968	-12.636	3.595	1.439
	(28.31)	(17.09)	(18.01)	(16.92)	(13.31)	(16.16)
Rebal Out (drops)	-20.638**	-19.418***	-19.101***	-10.138	-12.787	-11.81
	(8.86)	(6.12)	(6.81)	(8.88)	(8.18)	(8.78)
Active Share	127.004***	32.549		-106.074*	-162.017**	
	(41.83)	(65.95)		(54.90)	(72.11)	
Turnover	4.149	3.238		0.396	4.889	
	(6.39)	(5.60)		(6.34)	(5.70)	
Observations	5,596,295	5,596,295	5,596,294	3,124,482	3,124,482	3,124,482
R-squared	0	0.016	0.081	0	0.013	0.099
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for samples formed by splitting our baseline sample of active funds. We assign stocks into two subsamples if the embedded gain for that stock is above or below zero. The weights in the regression are equal to the size of the stock as a fraction of AUM of all stocks in fund j and quarter t for that subsample. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.2 for more details.

median holding periods. We make the same modifications for this setting: re-weighting and excluding additions (which have a holding period of 0). We report the results in Table 18, which shows that intensive-margin rebalancing trades are more predictive of future alphas in positions with long holding periods. This is consistent with the notion that managers may specialize in gathering information on a subset of stocks and hold those positions for relatively longer horizons.

On the other hand, managers seem to struggle more with timing the sales of stocks with long holding periods than short holding periods – similar to the difference between selling stocks with embedded gains and losses. Specifically, in the long holding period subsample, expected future alpha after intensive-margin selling is twice as large as in the sample with relatively short holding periods. This could be explained by several factors, including the fact that some stocks are held for a long time because they are difficult to exit, e.g., due to transaction costs or expected price impact.

Portfolio Count We consider how many stocks are held by each fund. As stated in the main text, holding many stocks may signal that is a fund is engaging in closet indexing or making broader style/sector bets. holding fewer stocks may be a sign of more targeted investing.

We split the sample based on how many stocks a fund holds in a given quarter – one sample where the number of stocks held is at or below the median number of stocks held across funds in that quarter, and the other for funds that held more stocks than the median. We report the estimates from regressing future alpha on our rebalancing measures for the two samples in Table 19.

The first striking finding is the difference in the predictive power of positive intensive-margin rebalancing for future returns. Columns 4, 5 and 6 show that positive intensive-margin rebalancing has strong predictive power for future alpha in funds with a small number of holdings, while columns 1, 2 and 3 show this relation is weak in funds with many holdings. One explanation for this is that in funds with many holdings, managers' attention may be limited, therefore making intensive-margin rebalancing more difficult.

On the other hand, extensive-margin rebalancing additions have much stronger predictive power for future alpha in funds with many holdings than funds with a small number of holdings. This is surprising, as one could imagine that funds with fewer holdings are more likely to be informed about stock-specific information.

AUM We split funds on above/below median assets under management each quarter. Table 20 shows that among small and large funds, the relation between active rebalancing and future returns are similar.

Table 18: Sample Split: Holding Period

	Alpha (t+1)					
	Long Holding Period			Short Holding Period		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	33.70**	26.106**	24.831*	24.07*	12.346	7.135
	(14.24)	(12.02)	(14.68)	(13.05)	(10.52)	(10.67)
Rebal In (negative)	-25.81	-11.196	-9.206	-50.00*	-29.815	-24.118
	(22.56)	(14.47)	(15.67)	(27.07)	(18.60)	(18.86)
Rebal Out (adds)	-15.20*	-16.322**	-15.075**	-14.24**	-12.602***	-11.875**
	(8.10)	(6.34)	(6.25)	(6.27)	(4.77)	(4.93)
Rebal Out (drops)	83.60**	9.275		54.11	-44.853	
	(41.05)	(59.05)		(42.53)	(65.27)	
Active Share	5.194	6.382		7.374	8.970**	
	(6.68)	(6.27)		(5.82)	(4.14)	
Turnover	-6.35	54.287	67.705***	19.58	99.552*	72.688***
	(31.96)	(44.25)	(3.64)	(34.16)	(51.71)	(2.89)
Observations	4,664,774	4,664,774	4,664,774	5,408,510	5,408,510	5,408,510
R-Squared	0	0.011	0.057	0	0.013	0.043
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for samples formed by splitting our baseline sample of active funds. We assign stocks into two subsamples if the holding period for that stock is or is not above the median holding period for fund j in quarter t. The weights in the regression are equal to the size of the stock as a fraction of AUM of all stocks in fund j and quarter t for that subsample. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t + 1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.2 for more details.

Table 19: Sample Split: Portfolio Count

	Alpha (t+1)					
	High Number of Holdings			Low Number of Holdings		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	27.851	23.859	24.041	42.80***	27.479***	20.994*
	(19.56)	(16.43)	(18.04)	(13.60)	(10.03)	(10.70)
Rebal In (negative)	-53.639	-29.332	-22.917	-33.92*	-17.867	-14.899
	(40.56)	(26.79)	(27.75)	(19.16)	(12.18)	(12.88)
Rebal Out (adds)	31.647**	24.863**	25.841**	18.95**	15.955***	14.591***
	(14.07)	(10.38)	(11.25)	(7.71)	(5.26)	(5.12)
Rebal Out (drops)	-23.319**	-22.053***	-24.581***	-14.19**	-14.487***	-13.903***
	(11.01)	(8.17)	(9.00)	(6.57)	(5.01)	(4.91)
Active Share	49.008*	-57.39		101.7	14.438	
	(27.56)	(45.88)		(81.75)	(112.60)	
Turnover	1.219	1.065		4.832	11.43	
	(5.95)	(4.22)		(5.67)	(7.17)	
Observations	7,744,547	7,744,547	7,744,547	2,328,737	2,328,737	2,328,737
R-Squared	0	0.01	0.031	0	0.013	0.07
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop}$$
$$+ \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for samples formed by splitting our baseline sample of active funds. We assign all stocks within a fund-quarter into two subsamples if the number of stocks held by fund j in quarter t is or is not above the median number of stocks held that quarter. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.2 for more details.

Table 20: Sample Split: Fund Size

	Alpha (t+1)					
	High AUM			Low AUM		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	33.29**	30.676**	26.060*	36.20**	23.269**	19.295
	(15.14)	(12.57)	(13.74)	(13.88)	(11.25)	(12.03)
Rebal In (negative)	-42.91	-31.531	-22.899	-32.42	-16.279	-13.667
	(30.18)	(19.75)	(19.40)	(20.73)	(14.72)	(16.10)
Rebal Out (adds)	19.81*	19.765**	18.643**	20.87**	18.062***	16.915***
	(10.10)	(8.32)	(8.26)	(8.70)	(6.11)	(5.93)
Rebal Out (drops)	-25.48**	-27.830***	-27.172***	-10.6	-11.931**	-12.248**
	(10.33)	(8.19)	(8.47)	(6.44)	(4.91)	(4.99)
Active Share	36.64	-20.597		78.20*	-63.168	
	(43.04)	(60.99)		(40.22)	(63.14)	
Turnover	15.72	6.654		0.877	4.589	
	(14.17)	(6.41)		(4.09)	(4.27)	
Observations	5,022,179	5,022,179	5,022,179	5,051,105	5,051,105	5,051,105
R-Squared	0	0.012	0.047	0	0.011	0.05
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} \\ + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for samples formed by splitting our baseline sample of active funds. We assign all stocks within a fund-quarter into two subsamples if the number of stocks held by fund j in quarter t is or is not above the median AUM that quarter. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.2 for more details.

The main difference between large and small funds arises for extensive-margin rebalancing drops, which are predictive of much higher returns for large funds than small funds. One explanation for this is that by nature of being large, such funds have a hard time choosing exactly which positions to sell in order to meet redemptions or make room for new ideas.

A.3 Alternative Split on Fund-Level Holding Period

In Table 18, we split observations within each fund quarter based on whether they had a holding period above or below the median for that portfolio at that point in time. In this subsection, we instead split the sample at the fund-quarter level based on average holding period. Specifically, we separate funds by first identifying each fund's average holding period over all stocks in their portfolio each quarter. This number can change quarter to quarter for a given fund, and will vary across funds even if they hold similar portfolios. That is, the average holding period in a fund-quarter will depend on when the fund purchased each position. We then take the median fund holding period each quarter and create two samples, one for funds above the median and one for funds at or below. We estimate the baseline regressions but separately for each sample, regressing future alpha on our four active rebalancing measures.

We report the coefficients for the two samples in Table 21. The results differ from those in Table 18 along several dimensions. First, when splitting at the fund level, positive intensive-margin rebalancing predicts future alpha for both the short holding period and long holding period sample. This suggests that different funds may specialize in investing at different horizons (Lan et al., 2018). Another difference is we can add back in the extensive-margin rebalancing measure, because even though they have a holding period of zero, we can still compute an average holding period for the fund. Broadly speaking, extensive margin additions for funds with short and long holding periods have similar future performance.

A.4 Missing Alphas

As discussed in the main body of the paper, we remove from our sample stocks for which we cannot compute quarter $t+1$ alpha. One might be concerned that these stocks have especially low returns, and this will bias our results towards finding more alpha. In Table 22, we add the stocks back to our sample, and set the alpha and benchmark-adjusted return equal to the delisting return, if that is available, and -100% if it is not. The results are broadly similar to those in Section 3.1, evidence that these missing alphas are not materially biasing our main results.

A.5 Fund-Quarter Aggregation

In Section 5, we discuss aggregating our rebalancing measures to the fund j quarter t level. As shown in Equation 11, these are constructed as weighted averages of positive and negative rebalancing across all positions in the fund, where the weights are proportional to the size of the position relative to the fund's AUM. In this subsection, we walk through a series of examples to illustrate the identifying variation in the regression of future fund-quarter level alpha on the measures of fund-quarter level rebalancing.

First, we would like to highlight that if the aggregation in Equation 11 did not weight individual stock-fund-quarter level rebalancing, $rebal_{j,t}^{add}$ and $rebal_{j,t}^{drop}$ would sum to zero at the fund-quarter level. This is

Table 21: Sample Split: Fund Average Holding Period

	Alpha (t+1)					
	Long Holding Period			Short Holding Period		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	37.57**	21.762**	14.352	36.09**	28.846**	28.532*
	(14.83)	(10.32)	(10.95)	(16.84)	(13.47)	(15.43)
Rebal In (negative)	-27.72	-12.711	-2.632	-46.64	-27.227	-30.08
	(20.45)	(14.63)	(14.56)	(29.45)	(18.62)	(20.90)
Rebal Out (adds)	21.33***	17.556***	17.352***	22.71**	19.067**	19.298**
	(8.07)	(6.05)	(6.41)	(10.28)	(7.49)	(7.92)
Rebal Out (drops)	-13.36	-10.108	-9.805	-18.02**	-19.744***	-21.045***
	(9.05)	(7.41)	(6.97)	(7.40)	(5.27)	(5.93)
Active Share	69.14	-4.833		48.55	-26.378	
	(44.78)	(79.81)		(42.67)	(47.65)	
Turnover	1.636	17.052**		7.447	2.36	
	(3.52)	(7.16)		(7.01)	(4.36)	
Observations	4,664,774	4,664,774	4,664,774	5,408,510	5,408,510	5,408,510
R-Squared	0	0.011	0.057	0	0.013	0.043
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for samples formed by splitting our baseline sample of active funds. We assign all stocks within a fund-quarter into two subsamples if fund j's average holding period in each quarter t is or is not above the median fund holding period that quarter. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.3 for more details.

Table 22: Missing Alphas

	Alpha (t+1)			Bench Adj Ret (t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebal In (positive)	36.039***	26.228**	21.813*	31.773	18.414	8.98
	(13.72)	(10.55)	(11.91)	(44.24)	(43.74)	(43.78)
Rebal In (negative)	-37.524	-22.727	-17.633	-5.435	4.428	10.721
	(23.87)	(15.96)	(16.61)	(27.55)	(31.07)	(36.65)
Rebal Out (adds)	21.638**	18.893***	17.743***	22.379	20.377	14.621
	(9.04)	(6.28)	(6.30)	(25.49)	(24.38)	(23.61)
Rebal Out (drops)	-14.924**	-15.563***	-15.828***	4.769	5.228	7.191
	(7.38)	(5.31)	(5.52)	(7.15)	(7.40)	(9.53)
Active Share	56.855	-28.598		-20.583	85.523	
	(40.97)	(59.22)		(34.25)	(68.70)	
Turnover	4.141	6.622		-2.583	-0.892	
	(5.43)	(4.09)		(6.93)	(4.72)	
Observations	10,105,252	10,105,252	10,105,252	10,105,252	10,105,252	10,105,252
R-squared	0	0.01	0.048	0	0.008	0.065
Fixed Effects	None	Fund YQ	Fund x YQ	None	Fund YQ	Fund x YQ

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+1} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop}$$
$$+ \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for samples formed by splitting our baseline sample of active funds. We assign all stocks within a fund-quarter into two subsamples if the number of stocks held by fund j in quarter t is or is not above the median number of stocks held that quarter. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. If a return is missing, we replace it with its delisting return; if the delisting return is missing, we assume the return was -100%. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). There are no fixed effects included, and standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 and Section A.4 for more details.

because $rebalscale$ accounts for changes in AUM due to flows, and therefore the money to buy any new stock must come from (negative) deviations from perfect scaling of existing positions or portfolio deletions. When including the weights, $rebal_{j,t}^{add}$ and $rebal_{j,t}^{drop}$ do not have to add up to zero, as we illustrate below.

To make this more concrete, in Panel A of Table 23, we present an example fund with \$6000 of AUM in quarter $t-1$. Suppose for simplicity that between quarter $t-1$ and t the fund receives no flows, and the prices of the stocks it holds do not change. Therefore, its AUM in quarter t must also be \$6000.

Between quarter $t-1$ and t, suppose this fund actively rebalances all its existing positions, but does not add any new stocks or drop any stocks entirely. Specifically, the fund scales down its holdings of AAPL, NFLX and GOOG by 3.45% so that it can increase its positions in FB and AMZN by 5%. FB is initially a relatively large position, which is why the fund has net positive rebalancing of 26.1 basis point of AUM (adding together the weighted positive rebalancing for FB and AMZN yields 74.4 basis points, while adding together the weighted negative rebalancing for AAPL, NFLX and GOOG yields -48.3 basis points).

In Panel B of Table 23, consider the same fund, but suppose instead it scales down its holdings of FB, AAPL and GOOG by 3.6% to increase its holdings of AMZN and NFLX by 20%. The difference from Panel A is that unlike FB, AMZN and NFLX are relatively small positions in quarter $t-1$. This is why the fund has net negative rebalancing of 57.5 basis points of AUM (37.7 basis points - 95.2 basis points).

Taking a step back, we believe Table 23 illustrates the intuition behind $rebal_{j,t}^{add}$ and $rebal_{j,t}^{drop}$. These measures capture rebalancing *in the fund's largest positions*. So, a fund will have positive net rebalancing if it is adding to large positions, and negative net rebalancing if it is adding to small positions. Our economic intuition behind constructing $rebal_{j,t}^{add}$ and $rebal_{j,t}^{drop}$ this way is that adding to large positions is likely more information-motivated than adding to small positions. The logic is that (1) adjusting a large position will have a larger effect on the fund's overall return and, as a consequence, (2) will entail larger tracking error and portfolio adjustment costs.

A.6 Stock-Quarter Aggregation

As discussed in the main body of the paper, a possible concern with our baseline results in Table 6 is that they are due to price pressure. Specifically, there are several papers documenting price pressure effects of mechanical buying/selling by mutual funds (Coval and Stafford (2007), Lou (2012), Khan et al. (2012)), and therefore it could be that, e.g., the high returns after positive active rebalancing are due to temporary price

Table 23: Example Fund-Quarter Aggregation

Stock			Before Rebalancing					After Rebalancing			
Panel A: Positive net rebalancing											
Stock	$shares_{t-1}$	p_{t-1}	$dheld_{t-1}$	wt_{t-1}	$shares_t$	p_t	$dheld_t$	$rebalin_t$ (shares)	$rebalin_t$ (%AUM)	wt_t	$rebalin_t \times wt_t$
FB	100	$ 22.50	$ 2,250.00	38%	105.00	$ 22.50	$ 2,362.50	5.00	1.88%	39.38%	0.738%
AAPL	35	$ 57.35	$ 2,007.25	33%	33.79	$ 57.35	$ 1,937.99	(1.21)	-1.15%	32.30%	-0.373%
AMZN	10	$ 19.99	$ 199.90	3%	10.50	$ 19.99	$ 209.90	0.50	0.17%	3.50%	0.006%
NFLX	20	$ 36.23	$ 724.60	12%	19.31	$ 36.23	$ 699.60	(0.69)	-0.42%	11.66%	-0.049%
GOOG	20	$ 40.91	$ 818.25	14%	19.31	$ 40.91	$ 790.02	(0.69)	-0.47%	13.17%	-0.062%
Total			$ 6,000.00				$ 6,000.00		Fund-Quarter Level	$rebal^{add}_{j,t}$	0.744%
										$rebal^{drop}_{j,t}$	-0.483%
Panel B: Negative net rebalancing											
Stock	$shares_{t-1}$	p_{t-1}	$dheld_{t-1}$	wt_{t-1}	$shares_t$	p_t	$dheld_t$	$rebalin_t$ (shares)	$rebalin_t$ (%AUM)	wt_t	$rebalin_t \times wt_t$
FB	100	$ 22.50	$ 2,250.00	38%	96.36	$ 22.50	$ 2,168.03	(3.64)	-1.37%	36.13%	-0.494%
AAPL	35	$ 57.35	$ 2,007.25	33%	33.72	$ 57.35	$ 1,934.13	(1.28)	-1.22%	32.24%	-0.393%
AMZN	10	$ 19.99	$ 199.90	3%	12.00	$ 19.99	$ 239.88	2.00	0.67%	4.00%	0.027%
NFLX	20	$ 36.23	$ 724.60	12%	24.00	$ 36.23	$ 869.52	4.00	2.42%	14.49%	0.350%
GOOG	20	$ 40.91	$ 818.25	14%	19.27	$ 40.91	$ 788.44	(0.73)	-0.50%	13.14%	-0.065%
Total			$ 6,000.00				$ 6,000.00		Fund-Quarter Level	$rebal^{add}_{j,t}$	0.377%
										$rebal^{drop}_{j,t}$	-0.952%

Notes. Example calculations for $rebal^{add}_{j,t}$ and $rebal^{drop}_{j,t}$. In Panel A, the fund has net positive rebalancing of 26 basis point of AUM, while in Panel B, the fund has net negative rebalancing of 58 basis points of AUM.

pressure.

Recall, however, that in Table 6 we are looking at the *active* rebalancing decisions of *individual funds*. While we believe it is unlikely that active trading by an individual fund is going to have a large impact on prices, if buying is correlated across funds (and persistent, because we are focused on active rebalancing in quarter t and returns in quarter $t+1$), it could be that their collective trading has an impact on prices. To test this hypothesis, we aggregate our active rebalancing measures to the *stock-quarter* level:

$$\text{Stk. Rebal Add}_{i,t} = \left(\sum_j rebalinshrs^{pos}_{i,j,t} + \sum_j rebaloutshrs^{add}_{i,j,t} \right) / shrout_{i,t} \tag{12}$$

$$\text{Stk. Rebal Drop}_{i,t} = \left(\sum_j rebalinshrs^{neg}_{i,j,t} + \sum_j rebaloutshrs^{drop}_{i,j,t} \right) / shrout_{i,t} \tag{13}$$

In words, Stk. Rebal Add is total positive rebalancing (both from positive intensive-margin rebalancing and additions) as a fraction of stock i's shares outstanding in quarter t. Stk. Rebal Drop is total negative rebalancing (both from negative intensive-margin rebalancing and deletions) as a fraction of stock i's shares outstanding in quarter t. These measures are designed to quantify aggregate active rebalancing relative to each firm's market capitalization.

Note that Stk. Rebal Add and Stk. Rebal Drop seem similar to the measure of aggregate active mutual fund trading in Chen et al. (2000). The key difference is that our measures take out expected scaling in response to flows, which as we show in Section 3.2 does not have predictive power for future returns.

We then run the following regression to estimate the relation between aggregate active rebalancing and future stock-level performance:

$$Performance_{i,t+1} = \alpha + \beta_1 \text{Stk. Rebal Add}_{i,t} + \beta_2 \text{Stk. Rebal Drop}_{i,t} + \gamma_i + \gamma_t + \varepsilon_{i,t} \tag{14}$$

where $Performance$ is either 4-factor alpha or market-adjusted return, γ_i are stock fixed effects and γ_t are year-quarter fixed effects.

The results are in Table 24. The first column suggests that if, in aggregate, active managers have positive rebalancing equal to 1% of a stock's shares outstanding (i.e., Stk. Rebal Add = 0.01), the stock will have -1.66 basis points of alpha the following quarter. Further, if active managers have negative rebalancing equal to 1% of a stock's shares outstanding, the stock will have -0.60 basis points of alpha the following quarter. Collectively, these results imply that there is reversal after positive and negative active rebalancing.

The second column shows that the general pattern is different by including stock and year-quarter fixed effects – positive rebalancing leads to reversal but negative rebalancing leads to continuation. Further, with the firm and time fixed effects, the magnitudes of the estimated coefficients are significantly larger. This could be due to the fact that the residual variation in future alphas is significantly smaller after including these fixed effects. The third column shows the results are very small and insignificant for market-adjusted returns without fixed effects. The last column shows that, with fixed effects, the results are also broadly similar when using market-adjusted returns instead of alphas.

It's possible that the results in Table 24 don't tell the whole story, because it is focused only on the active part of rebalancing. And as shown in previous literature (Coval and Stafford (2007), Lou (2012), Khan et al. (2012)), the mechanical part of active manager rebalancing may also have an effect on stock prices.

To quantify this, we run the same regression in Equation 14, except we swap the variables "Stk. Rebal Add" and "Stk. Rebal Drop" for "Pos. Δ Shr. Mkt. Cap." and "Neg. Δ Shr. Mkt. Cap." Rather than calculate the positive/negative active rebalancing as a share of each stock's market capitalization, these alternative measures quantify the *total* rebalancing as a share of each stock's market capitalization, i.e., they include the mechanical part of rebalancing due to scaling positions in response to flows.

The results are in Table 25. The results for Pos. and Neg. Δ Shr. Mkt. Cap. show reversal for both total buying and selling, respectively, though with mixed statistical significance. As an example to interpret the

Table 24: Aggregate Active Stock-Level Rebalancing and Returns

	Alpha (t+1)		Mkt. Adj. (t+1)	
	(1)	(2)	(3)	(4)
Stk. Rebal Add	-166.751*	-600.200**	0.002	-0.091**
	(96.84)	(285.55)	(0.02)	(0.04)
Stk. Rebal Drop	-60.886	446.976	-0.024	0.142***
	(199.80)	(281.44)	(0.03)	(0.05)
Observations	524,893	524,893	524,893	524,893
R-squared	0	0.066	0	0.094
Fixed Effects	None	Stock // YQ	None	Stock // YQ

Notes. This table shows coefficients from estimating

$$Performance_{i,t+1} = \alpha + \beta_1 \text{Stk. Rebal Add}_{i,t} + \beta_2 \text{Stk. Rebal Drop}_{i,t} + \gamma_i + \gamma_t + \varepsilon_{i,t}$$

$Performance$ is either 4-factor alpha or market-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Market-adjusted return is the return on the stock minus the return on the CRSP value-weighted market index. We include specifications with no fixed effects and with stock and year-quarter fixed effects. Standard errors are double clustered at the stock and quarter level.

magnitudes, after a quarter where active mutual funds have aggregate positive rebalancing equal to 1% of a firm's shares outstanding, alpha is 1.06 basis points lower over the following quarter. The market-adjusted returns coefficients are relatively small are mostly insignificant.

Collectively, the results in Tables 24 and 25 show a statistically weak and economically small relation between aggregate rebalancing and future returns, allaying concerns of mechanical price pressure driving our main results.

A.7 Horizons

To this point, we have only examined the returns one quarter after the managers' active rebalancing decisions. In this subsection, we examine returns up to four quarters after the original rebalancing decision. Looking at longer horizons is useful along two different dimensions. First, it speaks to concerns about price pressure driving our main results, as if that were the case, we would expect reversion in quarters after $t+1$ (Greenwood, 2005). Second, it may speak to the nature of the information managers are utilizing to inform their active rebalancing decisions.

Specifically, we run the same regression as in Equation 10, except the left-hand-side variable is 4-factor alpha in each of the 4 quarters after the active rebalancing occurred. The results are in Table 26. The first column

Table 25: Aggregate Total Stock-Level Rebalancing and Returns

	Alpha (t+1)		Mkt. Adk. (t+1)	
	(1)	(2)	(3)	(4)
Pos. Δ Shr. Mkt. Cap.	-106.858*	-83.318	-0.004	-0.022*
	(58.28)	(113.05)	(0.01)	(0.01)
Neg. Δ Shr. Mkt. Cap.	-216.343	-444.138*	-0.001	0.026
	(224.85)	(268.01)	(0.03)	(0.04)
Observations	524,893	524,893	524,893	524,893
R-squared	0	0.066	0	0.094
Fixed Effects	None	Stock // YQ	None	Stock // YQ

Notes. This table shows coefficients from estimating

$$Performance_{i,t+1} = \alpha + \beta_1 \text{Pos. } \Delta \text{ Shr. Mkt. Cap.}_{i,t} + \beta_2 \text{Neg. } \Delta \text{ Shr. Mkt. Cap}_{i,t} + \gamma_i + \gamma_t + \varepsilon_{i,t}$$

$Performance$ is either 4-factor alpha or market-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Market-adjusted return is the return on the stock minus the return on the CRSP value-weighted market index. We include specifications with no fixed effects and with stock and year-quarter fixed effects. Standard errors are double clustered at the stock and quarter level.

is a replication of the results in Table 6, restricting to the subset of stock-quarter observations for which we can compute alpha over all of the four quarters after the active rebalancing occurred. The results are very similar, with positive alpha after positive intensive- and extensive-margin rebalancing.

Columns 2 to 4 show that, for the most part, there is no relation between active rebalancing in quarter t and alphas in quarters $t+2$, $t+3$ and $t+4$. This suggests active rebalancing decisions forecast near-term future returns, rather than long-run future returns, consistent with managers acting on relatively short-lived information. Importantly, even though these point estimates are mostly insignificant, they are not the opposite sign as the estimates in column 1. This is further evidence against the price pressure story, as if this were purely due to price pressure, we would expect reversion over the following few quarters (Greenwood, 2005).

As a final check, in Table 27, we replicate the results from Table 26 for our set of passive funds. Again, the first column replicates Table 7 for the subset of observations with non-missing alphas over all 4 quarters following the rebalancing decision. Again, for the active rebalancing decisions of passive funds, there is almost no relation with future alpha, and there is no evidence of reversion over the following quarters either. The anomalous exception is for negative intensive-margin rebalancing, which shows continued poor performance over subsequent quarters.

Table 26: Active Rebalancing and Performance: Horizons Analysis, Active Funds

	Alpha			
	t+1	t+2	t+3	t+4
	(1)	(2)	(3)	(4)
Rebal In (positive)	43.269***	20.382	-1.183	1.967
	(13.80)	(15.22)	(9.56)	(13.07)
Rebal In (negative)	-41.616*	-36.134*	-40.948	-1.938
	(24.05)	(18.98)	(27.90)	(22.72)
Rebal Out (adds)	22.991**	4.46	16.808***	9.505
	(9.22)	(7.29)	(6.15)	(8.11)
Rebal Out (drops)	-11.946	-6.639	-1.592	-3.728
	(7.41)	(7.01)	(7.96)	(10.24)
Active Share	22.48	17.75	28.906	58.660*
	(41.42)	(34.89)	(36.55)	(34.32)
Turnover	5.044	8.209*	6.097	4.83
	(5.49)	(4.40)	(4.65)	(4.49)
Observations	9,770,329	9,770,329	9,770,329	9,770,329
R-Squared	0	0	0	0
Fixed Effects	None	None	None	None

Notes. This table shows coefficients from estimating a weighted regression:

$$Return_{i,j,t+h} = \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t},$$

for our sample of active funds, and horizons of $h = 1$, 2, 3 and 4 quarters. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t+1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 for more details.

Table 27: Active Rebalancing and Performance: Horizons Analysis, Passive Funds

	Alpha			
	t+1	t+2	t+3	t+4
	(1)	(2)	(3)	(4)
Rebal In (positive)	-36.61	-48.521	-70.926	-61.785
	(37.82)	(40.09)	(44.57)	(41.45)
Rebal In (negative)	30.831	87.163**	81.727*	92.551**
	(44.50)	(42.99)	(42.92)	(43.52)
Rebal Out (adds)	-18.056**	-12.937	-14.357	-4.063
	(8.07)	(9.06)	(10.59)	(8.16)
Rebal Out (drops)	-2.414	13.458	5.746	0.263
	(7.90)	(8.20)	(10.12)	(8.55)
Active Share	5.956	7.877	11.948	4.004
	(25.57)	(20.83)	(19.53)	(19.60)
Turnover	-1.96	0.783	1.648	2.178
	(3.62)	(3.83)	(3.87)	(3.98)
Observations	8,604,960	8,604,960	8,604,960	8,604,960
R-Squared	0	0	0	0
Fixed Effects	None	None	None	None

Notes. This table shows coefficients from estimating a weighted regression:

$$\begin{aligned} Return_{i,j,t+1} = {} & \beta_1^p rebalin_{i,j,t}^{positive} + \beta_1^n rebalin_{i,j,t}^{negative} + \beta_2^a rebalout_{i,j,t}^{add} + \beta_2^d rebalout_{i,j,t}^{drop} \\ & + \gamma_1 actshare_{j,t} + \gamma_2 turnover_{j,t} + \varepsilon_{i,j,t}, \end{aligned}$$

for our sample of passive funds, and horizons of $h = 1$, 2, 3 and 4 quarters. The weights in the regression are equal to the size of the stock as a fraction of AUM in quarter t. *Return* is either quarterly alpha or benchmark-adjusted return of stock i. Alpha is estimated against the market, SMB, HML, and UMD using daily data in quarter $t + 1$. Benchmark-adjusted return is the return of stock i minus the total return of the fund's benchmark index. Both return measures are in quarterly basis points. We regress the return measures on our four measures of active rebalancing: positive and negative rebalancing for existing positions (*rebalin*), and adding new or completely dropping past positions (*rebalout*). The active rebalancing measures are signed, and in units of a percent of AUM. We also include two measures of fund-level activeness: Active Share (Cremers and Petajisto, 2009) and Turnover (Pástor et al., 2017). The fixed effects specifications are as follows: Fund YQ corresponds to fund and quarter fixed effects, and Fund x YQ corresponds to fund-by-quarter fixed effects. Standard errors are triple clustered at the fund, stock, and quarter level. See Section 3.1 for more details.

B Thomson Reuters Mutual Fund (S12) Data Methodology

To our knowledge, we are the first paper to study the full set of changes in holdings at the stock-fund-quarter level. The replication code can be found on the authors' websites.

Given that we are focused on changes, data problems can be magnified because changes are typically much smaller than the associated levels. As a specific example of this, in Table 28, we show data for T. Rowe Price Personal Strategy Income Fund's (PRSIX, fundno 48) holdings of Seneca Foods Corporation (permno 62958). What stands out about the data is that the fund appears to drop the stock in 2008Q2, and the re-buys the same number of shares just a quarter later in 2008Q3. Then, the fund again appears to drop the stock for 3 quarters (2009Q2-2009Q4) before again re-buying the exact same number of shares. Eventually, in 2010Q4, the fund trims its position in the stock, and holds it steady at 500 shares for several quarters.

We believe this is a data error, because it seems unlikely to us that a fund would drop a stock, and then re-buy the exact same number of shares within a quarter. These data errors compound, in the sense that the missing holding in 2008Q2 for Seneca Foods creates a mistaken drop from 2008Q1 and a mistaken add in 2008Q3.[22]

In this Appendix, we describe our approach to cleaning the S12 data. As we describe below, there are some issues we believe we can fix and some issues we cannot – and as a result, we must drop the associated observations.

Fund-Quarter Data Filters Our first set of filters cleans data at the fund-quarter level. The number of observations that remain after applying these filters (in this order) is presented in Table 29.

1. We start with the universe of fund-quarters in the MF links database, which connects CRSP mutual fund data with the Thomson Reuters S12 data.

2. We start by removing observations with missing S12 fundnos, missing WFICNs, and duplicate WFICNs within a given year-quarter. We remove these observations because they cannot be matched to the CRSP mutual fund data.

3. We then match to the CRSP mutual fund data, and keep the subset of fund-quarter observations with

[22] We would also like to note that in the raw S12 data, there is no observation for T. Rowe Price Personal Strategy Income Fund's (PRSIX, fundno 48) holding Seneca Foods Corporation (permno 62958) in e.g., 2008Q2 (i.e., there is not an observation which states that the fund holds zero shares). When adjusting the holdings data to account for these types of issues, we create a (new) observation for each quarter we believe the fund held the stock, but did not report holding the stock in the S12 holdings data. As we describe below, this applies for gaps of up to three quarters, e.g., the gap in Table 28 between 2009Q2 and 2009Q4.

Table 28: Missing data and erroneous additions and deletions

S12 Fundno	Permno	Year-Quarter	Shares Held
48	62958	2006Q4	700
48	62958	2007Q1	700
48	62958	2007Q2	700
48	62958	2007Q3	700
48	62958	2007Q4	700
48	62958	2008Q1	700
48	62958	2008Q2	
48	62958	2008Q3	700
48	62958	2008Q4	700
48	62958	2009Q1	700
48	62958	2009Q2	
48	62958	2009Q3	
48	62958	2009Q4	
48	62958	2010Q1	700
48	62958	2010Q2	700
48	62958	2010Q3	700
48	62958	2010Q4	500
48	62958	2011Q1	500
48	62958	2011Q2	500
48	62958	2011Q3	500
48	62958	2011Q4	500

Notes. Example of possibly erroneous additions and deletions from a fund's portfolio

non-missing returns, flows and lagged total net assets (TNA). These are all key variables for computing our $rebalscale$ measure.

4. Following Cremers and Petajisto (2009), we compute the total equity holdings of each fund each quarter using the S12 holdings data. This is computed by first merging the holdings data to CRSP stock-level data, multiplying shares held by the stock's price, and adding up across all stocks in each managers' portfolio each quarter. We then drop funds for which total equity holdings are less than 67% of reported TNA in the CRSP mutual fund database. This is designed to remove funds which are not focused on investing in US equities (e.g., equity-bond blended funds).

5. Similarly, following Cremers and Petajisto (2009), we compute the fraction of holdings which can be matched to CRSP stock-level data. We then drop funds for which we cannot match at least 90% of their individual holdings to CRSP stock-level data (based on the count of matches, rather than dollars matched, as outlined in the previous filter). This again is designed to remove funds which are not focused on US equities.

6. Following Pástor et al. (2017), we remove small funds – specifically those funds which, at the S12 fundno level, have less than 15 million in TNA (where all TNAs have been deflated to 2015 dollars

using the CPI).

7. Following Evans (2010), we remove funds during their incubation period – which we identify as dates before the fund was first offered and dates where the fund has a missing name in the CRSP mutual fund database.

8. Given that we control for lagged fund-level turnover in our main regressions, we drop observations with missing values for this variable. Fund-level turnover comes from the CRSP mutual fund database.

9. We also remove fund-quarter observations with missing expense ratios, as we use these to compare gross and net returns in our fund-quarter level analysis.

10. Given that we are focused on changes in holdings, we aim to remove fund-quarters associated with stale data. The specific issue is that sometimes, a fund's last filing will be carried forward if they do not submit a new filing, i.e., in the S12 fund summary data there will be a single report date (RDATE) associated with multiple filing dates (FDATES).

    We label every FDATE that uses an RDATE after the first instance as stale. Because we are looking at changes in holdings from quarter $t-1$ to t, we exclude not only all quarters where the data is stale, but also the first quarter after the data stops being stale. Excluding the stale quarters avoids putting in zeros for changes in positions, which will be attributed to intensive margin rebalancing if the fund has any flows. Excluding the first quarter after the data stops being stale avoids mistakenly attributing the entire portfolio change that occurred over the stale quarters to the last quarter.

    - For example, suppose that a fund has a filing in 2019Q1, then has stale data in 2019Q2 (i.e., the same RDATE appears again), and reports fresh data in 2019Q3. Suppose, that a new stock appears in the 2019Q3 filing, but the fund actually added the stock to their portfolio in 2019Q2. If this fund has skill, the addition in 2019Q2 might lead to alpha in 2019Q3 (but, say, none in 2019Q4). Without removing the stale data, we would (mistakenly) attribute the purchase to 2019Q3 and compare it to alpha in 2019Q4, where we would erroneously find that the manager did not have skill in adding this stock.

    - This filter also removes all observations associated with the *re-use* of S12 fundnos. According to Thompson's documentation, if the same fundno has more than a year gap between RDATEs, one should assume this is actually a new fund with a re-used fundno. To avoid using the first observation associated with a new fund (where every stock has been added) and the last observation associated with a defunct fund (where every stock has been deleted), we drop the last observation associated with the old fund, and the first observation associated with the new fund. This is taken care of by the stale data filter described above.

11. We require each fund have at least 10 non-zero holdings changes (in terms of adjusted shares held) from quarter $t-1$ to t among stocks that are not added, dropped or involved in corporate actions like mergers or acquisitions. This is an additional safeguard against stale data, where all the changes in holdings would be zero.

12. We remove fund-quarter observations with missing next-quarter fund-level returns or alpha. This is because these are key outcome variables in our fund-quarter level analysis. We repeat our tests using delisting returns if available; if not, we assume a return of -100%. See Appendix A.4 for details.

13. We restrict to observations with non-missing active share. We obtain active share data from Martijn Cremers' website. This limits our sample to the 1990-2019 period where active share data is available.

14. We remove observations with missing total returns for the associated benchmark. Following Cremers and Petajisto (2009), we identify the benchmark based on the index with which the manager's portfolio has the smallest total difference in deviations from portfolio weights. This drops about 30,000 fund-quarter observations because we were unable to obtain total return data for several of the benchmark indices in the early part of our sample.

15. We remove passive funds, which we identify using the procedure in Appel et al. (2016). We classify a fund as passive if the procedure in Appel et al. (2016) identifies any fund-quarter in the fund's history as passive. We classify passive funds this way because the passive fund flag in the CRSP database is not well populated before 1999.

16. We remove funds with only one observation remaining after all the previous filters have been applied, because for such funds, it would be impossible to estimate a fund fixed-effect.

After applying all these filters, we end up with 72,907 fund-quarter observations from 3,190 unique funds. This is similar to the number of funds in Cremers and Petajisto (2009) (2,647), Lou (2012) (2,989), and Pástor et al. (2017) (3,126). We would also like to highlight that although the sample of Pástor et al. (2017) runs from 1979-2011 (8 years shorter than ours) we have fewer unique funds. The main differences come from our restriction on non-stale data, which is crucial when working with changes in individual holdings.

Stock-Fund-Quarter Data Filters We also apply a set of filters at the stock-fund-quarter level (i.e., individual holdings), which we summarize in Table 30. Note that the counts of affected observations only

Table 29: Fund-quarter filters

Fund Quarter Obs.	Filter
2,077,342	All fund-quarters in MF Links database
552,210	Non-missing S12 fundno, non-duplicate WFICN
511,565	Non-missing CRSP MF data for returns, flows and lagged TNA
280,258	Equities holdings at least 67% of reported TNA
270,834	Matched 90% of individual positions to CRSP
246,015	Fund with $15M real TNA (2015 dollars)
244,802	Remove incubation bias
208,594	Non-missing turnover and lagged turnover
208,527	Non-missing expense ratios
153,798	Non-stale holdings, not associated with re-used S12 fundnos
142,343	At least 10 non-zero holdings changes from $t-1$ to t
139,742	Non-missing next quarter fund-level returns and alpha
109,969	Dropping observations with missing active share
109,890	Remove missing t or $t+1$ benchmark returns
86,588	Remove passive funds
86,298	Remove singeltons
3,320	Unique Funds

Notes. This table provides filters, described in detail in Appendix B, and the number of observations in the sample after each filter is applied.

applies to the set of fund-quarter observations which pass the filters described in Table 29:

1. In the left column of Table 30, we consider all fund-quarters associated with funds that ever appear in our sample. In the 2nd column, we start with the universe of 72,907 fund-quarters described above (baseline FQ sample). In both cases, we restrict to the set of individual holdings that can be matched to CRSP.[23]

2. We label an observation as an addition if the fund did not hold the stock the previous quarter, but holds it at the end of the current quarter. We label an observation as a deletion if the fund held the stock at the end of the previous quarter, but not the end of the current quarter.

   - Because we attribute the deletion to the end of the quarter following the last quarter the stock appeared, we need to add a synthetic observation for each time a stock is deleted (e.g., if a fund adds and drops a stock 3 times we will need to add 3 observations – one at the end of each run of the fund holding the stock) where we fill in zero shares held in that quarter for that fund.

3. Something we noticed in working with the S12 data is that there are instances where an individual

[23] Note that the 50,200,000 observations indicated in column 1 row 1, and the 9,142,902 observations indicated in column 2 row 1 are larger than in the raw data after matching to CRSP. This is because they reflect the additional observations created to fill in the types of missing observations documented in Table 28 and the additional observations created to indicate deletions.

position will disappear and reappear several quarters later with the exact same number of shares. For example, suppose a manager has 100 shares of AAPL in 2019Q1, then reports having no shares of AAPL in 2019Q3, and again reports having 100 shares of AAPL in 2019Q3. We provide a specific example of this in our data in Table 28. While of course managers might add and drop the same stock, we find it unlikely that they would add and drop entirely *exactly* the same number of shares. So, any time a stock appears to disappear, and reappears within 4 quarters, we classify the associated observations as data errors. We make two adjustments to the data to account for these data errors. First, we argue that in such instances, the fund is not actually adding or dropping the stock (even though it appears to exit and re-enter the funds' portfolio). Second, we exclude the observations where the stock disappears, and the first observation where it reappears. The logic is that we don't know exactly what the fund did with the shares in the interim – so to be conservative, we just drop these cases.

- Note that in the 1st column, this changes additions and deletions by the same number of observations (as expected – because for every erroneous deletion, there should be a matching erroneous addition). In the 2nd column, the change is not symmetric because we have filtered for the subset of quarters described in Table 29.
- Also note that in the 1st column, there are many more cases where we fill in gaps in missing holdings than the second column. This implies that our fund-quarter filters remove a significant number of fund-quarter data with erroneously missing holdings data.

4. As an alternative, rather than requiring the exact same number of shares when the stock reappears, we also allow for 10% absolute difference between the number of shares held before and after the gap.

5. We remove observations where, in the original data, shares held is missing.

6. We remove all observations where the firm is an acquiring permno or has an acquiring permno in quarter $t-1$ or quarter t. This is because these cases can involve dramatic changes in shares outstanding (e.g., due to issuing shares to buy another company) which are not accounted for by CRSP's CFACSHR variable.

7. We remove observations for which we cannot compute alpha or benchmark-adjusted returns over the following quarter. This is because future alpha and returns are our key outcome variables of interest.

Finally, although this is not a filter, we would like to highlight that Thompson did a major update to the data from 2017 onward.

Table 30: Stock-fund-quarter changes and filters

Stock Fund Quarter Observations		
Our Funds, All Quarters	Our Funds, Filtered Quarters	Data Subset
27,300,000	10,600,000	All S12 matched to CRSP
3,831,121	1,467,582	Additions
3,831,234	1,452,826	Deletions
3,808,818	1,459,659	Additions (after fix for exact same # shares reappearing)
3,808,931	1,446,503	Deletions (after fix for exact same # shares reappearing)
3,761,869	1,437,668	Additions (after fix for $<10\%$ diff. in shares reappearing)
3,761,982	1,425,093	Deletions (after fix for $<10\%$ diff. in shares reappearing)
14,276	4,035	# of 1 period gaps filled (exact same # shares before/after)
11,548	2,757	# of 2 period gaps filled (exact same # shares before/after)
6,825	2,155	# of 3 period gaps filled (exact same # shares before/after)
37,050	14,112	# of 1 period gaps filled ($<10\%$ diff. in shares before to after)
41,578	15,575	# of 2 period gaps filled ($<10\%$ diff. in shares before to after)
34,239	14,831	# of 3 period gaps filled ($<10\%$ diff. in shares before to after)
1,094,738	372,559	Is acquirer or acquired in quarter t or $t-1$
26,451,060	10,433,424	Non-missing $t+1$ alpha and return, and t price

Notes. Stock-fund-quarter-level data filters and associated change in sample. In the left-most column, we consider all fund-quarters associated with funds that ever appear in our sample. In the 2nd column, we start with the universe of 72,907 fund-quarters described in Table 29. In both cases, we restrict to the set of individual holdings that can be matched to CRSP.