Internet Appendix for
"Index Rebalancing and Stock Market Composition: Do Indexes Time the Market?"

Marco Sammon and John J. Shim

A S12 Data Cleaning

In this section, we provide additional details on how we processed the S12 data, and how we identify index funds tracking major indexes.

A.1 Standard Data Quality Filters

We apply several filters to the fund-quarter data which are standard in the literature (see e.g., Pástor et al. (2017)). 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 S12 holdings to CRSP stock data and for which the value of their total S12 stock holdings represents at least 66% of their total AUM 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 drop all fund-quarters where a fund was involved with a merger in that quarter or the previous quarter, as this can lead to large changes in shares held which are not well explained by our ownership ratio framework.

A.2 Timing Issues

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 the *SPDR S&P 500 ETF Trust* (Ticker: 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 only has one "new" report each year in the Thomson S12 data, and Thomson 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.

Next, we must account for reporting delays, which are identified by fund-quarters where the RDATE lags significantly behind the FDATE. Specifically, we identify holdings data as "delayed" if the RDATE is not in the same quarter as the FDATE (results are similar if we instead require that the RDATE and FDATE be within one week of each other). We then drop observations which meet this criteria in the current quarter or previous quarter. It is necessary to exclude these observations, as they could have a misalignment between measures of fund flows and returns from the CRSP mutual fund data and the changes in holdings.

A.3 Re-Use of S12 Fund Numbers (fundnos)

Another issue related to gaps between RDATEs is Thomson'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 remove the first observation for a new fund (because all holdings would be classified as additions) and the last observation for the old fund (because all holdings would be classified as deletions). For the same reason, we also drop the overall first and last observation for every $fundno$.

A.4 Identifying Value-Weighted Index Funds Tracking Major Indexes from Holdings Data

For most of our analysis, we restrict to the S&P family of indexes (500, 400, and 600), the Russell family of indexes (1000, Mid Cap, 2000, and 3000) and the CRSP market-capitalization-based indexes and the CRSP total market index. For indexes in the S&P and Russell families, we identify these funds as passive funds with a median active share (Cremers and Petajisto, 2009) of less than 5% relative to their respective benchmarks. Initially, a 5% threshold appears large, especially given that many index funds exhibit tracking errors close to zero. However, there are substantial data issues in the S12 dataset – particularly prior to the 2017 update – which result in missing holdings and consequently a non-zero active share, as discussed in Appendix B.3. For the CRSP-index-tracking funds, we manually identify them, as active share is not computed for these indexes. To our knowledge, Vanguard is the only large asset manager with index funds tracking CRSP indexes.

Filtering for funds which meet these criteria ultimately restricts our sample to data between 1996 and 2023. This seems surprising, as e.g., SPY was launched in 1993. However, as discussed in Appendix B.3, several data issues in SPY, including stale holdings and missing holdings, lead it to be removed from our final sample.

A.5 Stock-Quarter level Data Filters

In Section 4.2, when constructing index funds' rebalancing portfolios, we also apply several data-cleaning filters at the stock-quarter level. First, following Greenwood and Sammon (2025), we exclude all stocks that are involved in a merger in quarter t or quarter $t-1$ as this can lead to dramatic changes in shares held by the index fund for reasons outside our ownership ratio framework. Further, we winsorize stock returns at the 1% and 99% level to reduce the effect of outliers on the extensive margin rebalancing portfolios. This is because an index fund might only add/drop one or two stocks in a regular quarterly rebalance, so even if the portfolio is "value weighted," all the weight would be placed in a small number of names.

B Value-Weighted Stock Index Framework

B.1 Ownership Ratio Perspective

In this appendix, we provide supporting materials for Section 3.1, and show how the ownership ratio framework is equivalent to that of a value-weighted index.

We describe a hypothetical index that tracks a value-weighted portfolio of the universe of U.S. stocks. The framework can easily be extended to subsets of this universe (for the purpose of this framework, our index does not apply any float adjustments). This means that each stock's weight in the index is equal to each stock's market capitalization (i.e., price times shares outstanding) divided by total stock market capitalization, or

$$w_{i,t} = \frac{me_{i,t}}{totalme_t}, \tag{A.1}$$

where $me_{i,t}$ (market equity) is equal to $p_{i,t} \cdot shrout_{i,t}$ (price times shares outstanding). $totalme_t$ is the sum of market equity for all stocks.

We now introduce an index *fund* that tracks our hypothetical stock index. The total dollars in the fund is given by aum_t (assets under management). It is useful to measure the fund as owning a fraction of the overall stock market. To this end, define the *fund ownership ratio* Ω_t as

$$\Omega_t = \frac{aum_t}{totalme_t}. \tag{A.2}$$

We can also compute the ownership ratio at the stock level, i.e., the fraction of shares outstanding of each stock i owned by this fund. We first compute the number of shares held in stock i by the fund, which is

$$sharesheld_{i,t} = \frac{(totalme_t \cdot \Omega_t) \cdot w_{i,t}}{p_{i,t}} = \Omega_t \cdot shrout_{i,t}, \tag{A.3}$$

where the term in parenthesis is the AUM of the fund. This means that the numerator is the total dollar value of the fund's position in stock i. Using the shares held in stock i, we can compute the *stock-specific ownership ratio* $\Omega_{i,t}$, as

$$\Omega_{i,t} = \frac{sharesheld_{i,t}}{shrout_{i,t}}. \tag{A.4}$$

Substituting Equation A.3, we see that, for a value-weighted index fund, the ownership ratio is invariant across stocks, or

$$\Omega_{i,t} = \Omega_t. \tag{A.5}$$

While the algebra is straightforward, the resulting Equation A.5 clarifies an important insight into how value-weighted indexes and index funds work. It says that owning a value-weighted total stock market portfolio is identical to owning the same fraction of shares outstanding for every stock, and that fraction is given by what fraction of the overall market the fund owns. That is, if a fund tracking our example value-weighted index has AUM equal to 1% of the market's total capitalization, it achieves this by owning 1% of every stock's shares outstanding. This means that the ownership ratio can be measured in a variety of ways. If there are N stocks in the market, then the ownership ratio is

$$\Omega_t = \frac{aum_t}{totalme_t} = \frac{sharesheld_{1,t}}{shrout_{1,t}} = \frac{sharesheld_{2,t}}{shrout_{2,t}} = \cdots = \frac{sharesheld_{N,t}}{shrout_{N,t}}. \tag{A.6}$$

Note that Ω_t is not indexed by i, i.e., it is completely stock invariant. This means that if a stock i's shares outstanding changes – something that is outside of the fund's control – the fund must adjust its holdings in that stock (as well as many others) to keep Ω_t constant across all stocks.

B.2 Rebalancing

In this subsection, we provide complete details on the assumptions made to derive the rebalancing trade described in Section 3.4 and presented in Equation 4.

We denote the set of stocks that exist at time t as $\mathcal{S}_t$. The rebalancing we provide can also be applied to a segment of stocks (e.g., a 500 stock large-cap index or a mid-cap value index), where the stocks that fall in the segment, and are thus included in the index, constitute the set $\mathcal{S}_t$. It is also simple to adapt our methodology to include float adjustments by replacing every instance of $shrout_{i,t}$ with $shrout_{i,t} \times IWF_{i,t}$, i.e., multiply each stock's shares outstanding (or index-eligible shares outstanding) with a float adjustment or investable weight factor (IWF).

We consider index rebalancing periods that are far enough apart such that many events can occur between index reconstitutions. We use the index-eligible shares outstanding approach, outlined in Section 3.3, to update shares outstanding for all ownership ratios at each rebalancing period. As discussed in the main text, there are several types of events that cause the index fund to trade in order to reestablish a constant ownership ratio, e.g., IPOs, SEOs, and buybacks. We assume a particular timing of these events to simplify the analysis. At $t-1$, the index fund holds a portfolio of stocks in the set $\mathcal{S}_{t-1}$. The index fund sets its portfolio holdings and rebalances at t after the following events occur in this particular order:

1. **Price Changes.** Each stock $i \in \mathcal{S}_{t-1}$ has some change in prices. The return of stock i is given by $r_{i,t} = p_{i,t}/p_{i,t-1} - 1$. The fund return is simply the return of its constituent stocks, and is given by

$$r_{fund,t} = \sum_i w_{i,t} \cdot r_{i,t}. \tag{A.7}$$

   The overall market return is equal to the fund return because, in our stylized fund, there is no *index* tracking error. That is, $r_{fund,t} = r_{idx,t}$.

2. **Flows.** The fund has some amount of flows. Flows, as a percentage of fund size, are denoted $flow_t$, and defined as

$$flow_{j,t} = \frac{aum_{j,t}}{aum_{j,t-1}} - (1 + r_{j,t}) \, . \tag{A.8}$$

   With a flow, the fund will trade by perfectly scaling up or down each position proportionally, which changes each stock-level ownership ratio to be equal to the new fund-level ownership ratio and reestablishes a constant ownership ratio.

3. **Issuance/Buybacks.** Each stock $i \in \mathcal{S}_{t-1}$ can engage in buybacks or issue shares. The net issuance as a percentage of shares outstanding is given by

$$issue_{i,t} = \frac{shrout_{i,t}}{shrout_{i,t-1}} - 1. \tag{A.9}$$

Issuance changes the stock-specific ownership ratio through shares outstanding. In isolation, a single stock's issuance causes that stock's ownership ratio to decline, which would require the fund to buy that stock's shares and fund it by proportionally scaling down all other stocks until all stock-specific ownership ratios are equal to each other. We call this type of adjustment intensive margin rebalancing because it adjusts positions among existing stocks. When many firms adjust shares outstanding (which is likely to occur between, say, quarter rebalancing periods), intensive margin rebalancing will typically involve buying and selling stocks *disproportionately*. In addition, all stocks' net issuance collectively will have an impact on total market equity, which also changes the *fund* ownership ratio. The change in total market equity from all stocks' issuance and buybacks is represented as

$$issue_{mkt,t} = \frac{\sum_i issue_{i,t} \cdot shrout_{i,t-1} \cdot p_{i,t-1}}{totalme_{t-1}}. \tag{A.10}$$

4. **Additions/Deletions.** Stocks regularly enter and exit the public market. For indexes that track market segments, stocks can enter or exit from changing characteristics as well. Denote the set of new stocks as $\mathcal{A}_t$. The set of stocks that was included in $\mathcal{S}_{t-1}$ but are dropped in this addition/deletion period is $\mathcal{D}_t$. Therefore, the set of stocks for period t is $\mathcal{S}_t = \mathcal{S}_{t-1} + \mathcal{A}_t - \mathcal{D}_t$. The fund will buy the added stocks and sell all shares of the dropped stocks. These entries/exits also change the composition of the overall market – e.g., if there is a difference in total value between the added and dropped stocks, total market equity will change. We can think of entry and exit in the market similar to the flows for the fund – they change the fund ownership ratio, not from dollars coming in and out of the fund (numerator), but from dollars coming in and out of the overall market (denominator). We measure the addition/deletion flow as a fraction of total market equity, and it can be represented as

$$adddrop_{mkt,t} + issue_{mkt,t} = \frac{totalme_t}{totalme_{t-1}} - (1 + r_{mkt,t}). \tag{A.11}$$

Changes in total market equity, once adjusting for returns, must come from changes on the extensive margin (entry/exit from the market or the tracked market segment) and changes on the intensive margin (issuance/buybacks from existing stocks). An increase in $adddrop_{mkt,t}$, all else equal, leads to an increase in total market equity that the fund was not previously exposed to. We refer to this trading from adds/drops as extensive margin rebalancing since it involves heterogeneous buying and selling but only for stocks that enter or exit the market – all existing positions are scaled up and down based on the new fund ownership ratio.

5. **Index Rebalancing.** Many of the prior events change the ownership ratio, as was outlined for each event in isolation. Index rebalancing takes *all* of these factors into account and adjusts the ownership ratio, which requires changes in holdings. That is, the index fund must trade to match the new fund ownership ratio. The new constant ownership ratio for t is

$$\Omega_t = \frac{aum_t}{totalme_t}. \tag{A.12}$$

Among the stocks that were neither added or dropped, the new position that the fund targets in each stock i relative to the old position is

$$\frac{shares_{i,t}}{shares_{i,t-1}} = \frac{\Omega_t \cdot shrout_{i,t}}{\Omega_{t-1} \cdot shrout_{i,t-1}}, \tag{A.13}$$

which ultimately yields the expression

$$\frac{shares_{i,t}}{shares_{i,t-1}} = \frac{(1 + r_{fund,t} + flow_t)\,(1 + issue_{i,t})}{(1 + r_{mkt,t} + issue_{mkt,t} + adddrop_{mkt,t})}. \tag{A.14}$$

The expression shows that the fund will increase its position in a stock if the numerator, which contains fund-specific and stock-specific changes, exceeds market-level composition changes. That is, if a stock's net percentage issuance is large relative to overall market issuance and new additions, the fund will be a net buyer in that stock.

B.3 Empirical Validation

Equation 4 provides a guide for how our stylized index fund rebalances, which we can take to the data. The purpose of taking our methodology to the data is to provide a check that, despite the simplifying assumptions we have made, the ownership ratio framework is supported by the way real-world index funds operate.

Because we make some assumptions to simplify our framework, we expect that our predicted rebalancing will not be a perfect fit empirically. This is because the data we use is delayed or is estimated with error.[20] In addition, the mutual fund holdings data quality is low, with many holdings that are simply missing. As a specific example of this, the SPDR S&P 500 ETF Trust (Ticker: SPY) is a large S&P 500 tracking ETF. Despite the fact that it should hold at least 500 stocks, from 2011 to 2018, our examination of the Thomson Reuters S12 data shows that it erroneously reports only 420-460 holdings. We have contacted Thomson Reuters and have resolved a small fraction of these errors, but the majority remain, especially for data before 2017.

Importantly, however, many of the variables in Equation 4 are fund or overall market aggregates, and shares held can be properly measured for the holdings that are non-missing in the data. This implies that, despite the many data issues described above, our framework will still retain significant predictive power. This is in contrast to computing and working with index fund *weights*, which rely on having all holdings for accuracy.

With these caveats in mind, we provide a regression that can be used for estimation. Taking logs of Equation 4 and allowing for fund-by-fund estimation yields the following regression:

$$\log \frac{shares_{j,i,t}}{shares_{j,i,t-1}} = \alpha + \beta_j^{flow} \cdot \log\left(1 + r_{j,t} + flow_{j,t}\right) + \beta_j^{issue} \cdot \log\left(1 + issue_{j,i,t}\right) + \beta_j^{mkt} \cdot \log\left(1 + r_{j,t} + issue_{j,t} + adddrop_{j,t}\right) + \varepsilon_{i,j,t}. \tag{A.15}$$

Here, the j subscript identifies each fund in our data and dictates the universe of stocks to compute aggregate issuance and add/drop imbalances. Our rebalancing framework states that, for a value-weighted index fund, $\alpha = 0$, $\beta^{flow} = \beta^{issue} = 1$, $\beta^{mkt} = -1$. When bringing Equation 4 to the data, we use quarterly holdings and quarterly stock data to match the quarterly rebalancing used by many indexes (again acknowledging that there are many exceptions).

[20] For example, shares outstanding in CRSP can be somewhat stale, and firm buyback activity may not be publicly reported until months after the actual buying took place. This comes from examining the CRSP data and comparing with primary sources like 10-K filings and other sources like Bloomberg. Other variables, like fund flows, are *estimated*. For example, flows are estimated using AUM and fund returns following the methodology of Barber et al. (2016).

For this empirical exercise, we focus on stocks that were held for at least two consecutive quarters. We do not examine stocks added to the index or dropped from the index because 100% of dropped stocks are sold, and added stocks' trading is determined completely by the new fund-level ownership ratio.

We construct the dependent and independent variables directly from the mutual fund holdings data. $shares_{j,i,t}$ is the split-adjusted shares for stock i held by fund j in quarter t. We estimate $flow_{j,t}$ following Barber et al. (2016) as

$$flow_{j,t} = \frac{aum_{j,t}}{aum_{j,t-1}} - (1 + r_{j,t})\,. \tag{A.16}$$

where AUM is fund j's assets under management and $r_{j,t}$ is the fund's total return. Total AUM data comes from CRSP and is aggregated across all share classes to the S12 $fundno$ level.[21]

The stock-level issuance or buyback, $issue_{j,i,t}$ is the quarterly change in stock i's float-adjusted shares outstanding, scaled by last quarter's float-adjusted shares outstanding.[22] For fund-level aggregate issuance and add/drop percentages (i.e., $issue_{j,t}$ and $adddrop_{j,t}$), we take the signed total dollar value of all stock-level issuance (and buybacks) and the signed total dollar value of adds and drops, both scaled by the total market capitalization of all stocks in the index at the end of the current quarter.

When estimating Equation A.15, we value-weight observations within each fund-quarter (but each fund-quarter is weighted equally). The logic of using these weights is that some funds may replicate an index through selective sampling, but the deviations from expected behavior are likely to be smaller for stocks that comprise a relatively larger share of the index (in dollar terms). By weighting toward larger holdings, we focus on the positions most likely to be rebalanced, in-line with the methodology we outline in Section 3.4. In addition, before computing these weights and running

[21] It is important to use total AUM from CRSP as opposed to adding up the total value of reported holdings, as that would omit any portfolio cash, and any holdings which cannot be matched to CRSP. $r_{j,t}$ is the weighted average return across all $crsp_fundno$s associated with a given S12 fundno, where the weights are proportional to lagged AUM in each share class of the fund. For the effects of dividends and dividend reinvestment on the estimated $flow_{j,t}$ and the estimation of Equation A.15, see Appendix B.4.

[22] For each major index family we consider, we obtain index membership data directly from the index provider. We obtain quarterly data on S&P Total Market and S&P 500 index membership directly from S&P. We obtain daily data on Russell 1000 and 2000 index membership directly from FTSE Russell. Finally, we obtain daily data on CRSP Total Market index membership directly from CRSP.

regressions, we winsorize all of the right-hand-side and left-hand-side variables at the 1% and 99% level by fund-quarter to reduce the influence of data errors.

We examine four index funds that follow the methodology we outline in Sections 3.1 and 3.4. We also use these four index funds because they track the most popular stock market indexes, and we have data on float adjustments used for each of the indexes. The four funds are SPDR S&P 500 ETF Trust (Ticker: SPY), iShares Russell 3000 ETF (Ticker: IWV), iShares Core S&P Total US Stock Market ETF (Ticker: ITOT) and Vanguard Total Stock Market Index Fund ETF (Ticker: VTI), which tracks the CRSP Total Market index.[23] We use these index funds/ETFs because they nearly perfectly track the underlying index. For this analysis, we do not need to use data on all index funds that track these indexes because they all rebalance in exactly the same way that the index does. We provide plots of the three beta estimates using rolling 4-quarter regressions for each fund in Figure A.1.

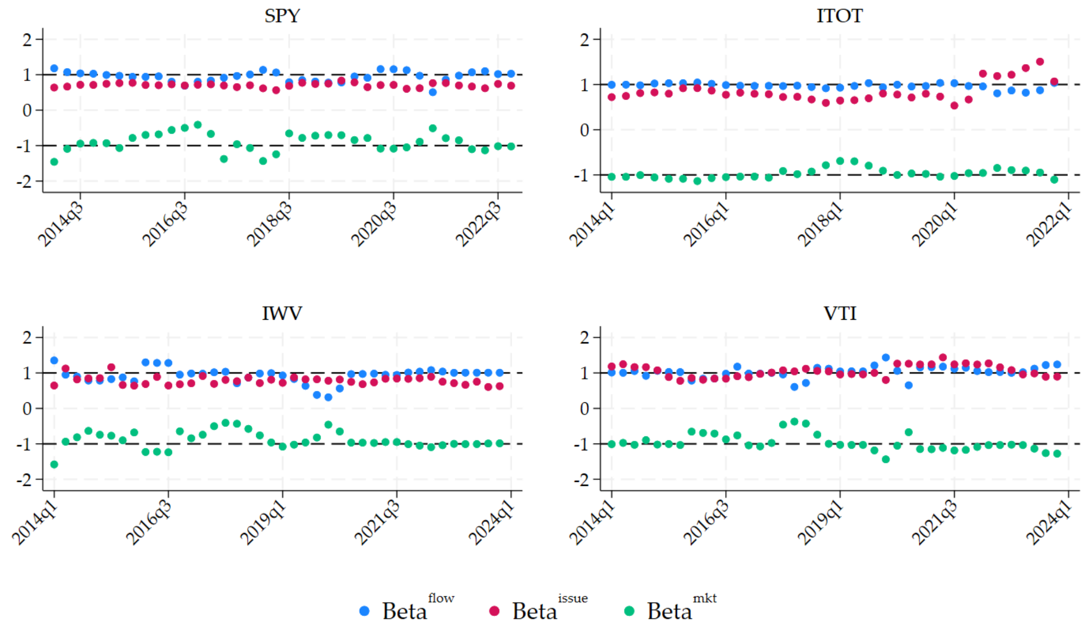
Most observations for SPY, ITOT, IWV and VTI have betas of 1, 1, and -1 as predicted by the empirical methodology. Therefore, despite some variation around the predicted values, we believe that Figure A.1 strongly supports our ownership-ratio framework. To further validate our methodology, we bring the rebalancing predictions of Equation A.15 to our full sample of index funds in Appendix D.1. There, we find that our decomposition is an accurate description of how most index funds work, further bolstering the idea that ownership ratio logic underlies the design of almost all real-world index funds, even though many focus on particular subsets of the stock market.

B.4 Real-World Clarifications and Refinements

We now discuss minor adjustments to the above framework to incorporate many real-world considerations.

[23]VTI previously tracked the MSCI total market index. We only plot data for VTI starting in 2014Q1 when VTI completed the transition to tracking a CRSP-based index. For SPY, there is only one filing per year before 2008, so we plot data starting in 2008 when we have non-stale filings every quarter.

Figure A.1: Betas for Large Value-Weighted Index Funds



Notes. This figure plots the betas from the regression in Equation A.15. The dashed horizontal lines represent the predicted values of 1 and -1.

Flows While most indexes rebalance quarterly, index funds must still trade daily due to flows. Thus, index funds will often use index-eligible shares between rebalances (fixing the denominator of stock-level ownership ratios and the non-price component of total market equity), and adjust the numerator because flows increase or decrease the fund's assets under management. As a result, index funds simply scale their positions up or down to reestablish an invariant ownership ratio.

Cash Dividends Most real-world index funds hold stocks that pay dividends, and therefore the funds themselves pay dividends to investors. In practice, there are two common ways that index funds handle dividends. The first, used by unit investment trusts (e.g., SPY), is to accumulate the cash dividends throughout the quarter and distribute them in a lump sum at the end of the quarter. The second, used by index mutual funds and ETFs (e.g., VTI), minimizes cash drag by reinvesting dividends as they are received. Specifically, the dividends are put to work by proportionally scaling up all existing positions. To fund the dividend payment to shareholders, these funds later raise cash by proportionally scaling down all existing holdings.

In both approaches, the payment of dividends does not affect the ownership ratio of the fund. This can be seen in two ways. First, consider the stock ownership ratio in Equation 2: when a company pays a cash dividend, neither the number of shares outstanding nor the number of shares held by the fund changes, so the ownership ratio remains constant. Second, the fund ownership ratio in Equation 1 is unaffected because both the numerator (the value of the fund's holdings) and the denominator (the total market value) decrease by the same amount in percentage terms (i.e., by the dividend yield) once the dividend is paid. This holds regardless of whether the dividends are temporarily reinvested before the fund itself pays the dividend, because any reinvestment is subsequently reversed by a proportional scaling down all positions.

A situation in which dividends could affect the ownership ratio is if dividends are reinvested. When dividends are reinvested, the fund will use the cash received to scale up its existing positions, so the number of shares held by the fund increases while the number of shares outstanding remains unchanged. As a result, the fund's ownership ratio *increases* after the dividend payment and subsequent reinvestment. This also makes it clear that dividend reinvestment should be treated as an inflow in Equation 4.

Empirically, neither actual flows nor dividend reinvestment rates are directly observed. Instead, these quantities must be estimated. Following Barber et al. (2016), we estimate flows using Equation A.16. Essentially, this equation estimates flows as the difference between the fund's observed AUM and the "expected" AUM given the fund's returns. Importantly, the fund return $r_{j,t}$ in Equation A.16 is a *total* return i.e., it includes the dividend yield. So, Equation A.16 only identifies flows above and beyond the expected AUM of the fund if all investors fully reinvested their dividends. As we explain in more detail below, the ultimate effect of this assumption is to systematically understate true expected scaling by the dividend yield.

To clarify this, consider two extreme examples. First, suppose that 100% of dividends are reinvested. Assume the fund begins with an AUM of \$100, experiences total capital gains of \$1, and has a dividend yield of 1% (i.e., \$1), for a total return of *approximately* 2% (we say approximately here because we are ignoring the precise timing of dividend payments and capital gains). Now suppose the fund also receives an inflow of \$1. At the end of the period, the fund's AUM is \$103: the initial \$100 grows by \$2 (from the reinvested dividend and capital gain), and the last \$1 comes from the inflow. By contrast, the expected AUM based on the total return is \$102. Thus, using Equation A.16, the estimated flow is \$1. However, because dividends were reinvested, the fund must scale up all existing positions by 2%: 1% for the reinvested dividend and 1% for the inflow. This will affect the estimation of Equation A.15, as the estimated 1% flow implies that positions should be scaled up by 1%, but in fact they need to be scaled up by 2%. Finally, we believe this example clarifies the claim that Equation A.16 identifies flows above and beyond what would be expected if 100% of dividends were reinvested.

Alternatively, consider the other extreme in which no dividends are reinvested. Suppose again that the fund begins with an AUM of \$100, earns a \$1 capital gain, and has a 1% dividend yield (i.e., \$1), while also receiving an inflow of \$1. At the end of the quarter, the fund's AUM is \$102: the initial \$100 grows by \$1 from the capital gain and \$1 from the inflow, but the \$1 from the dividend is not reinvested. In this case, the expected AUM from Equation A.16 is also \$102, so the estimated flow is zero. The fund, however, must still scale up its holdings by 1% to invest the \$1 inflow. Thus, the estimate of flows from Equation A.16 again understates the actual expected scaling of the index fund's positions, and again the gap is exactly equal to the dividend yield of 1%.

To summarize, whether dividends are fully reinvested or not reinvested at all does not *differentially* affect our calculations. In either case, Equation A.16 underestimates the true expected scaling by exactly the dividend yield. So, this methodology for estimating flows will lead us to systematically understate expected scaling. The reason the dividend reinvestment rate doesn't differentially affect this understatement is because Equation A.16 identifies flows based on AUM growth relative to expected AUM growth given total returns (and thus is agnostic to AUM growth from dividend reinvestment or other sources). As mentioned above, however, dividend reinvestment should be treated as an inflow from the perspective of Equation 4. And, because the expected scaling given estimated flows is systematically too small relative to true scaling, β_j^{flow} in Equation A.15 will be biased upward.

So, the natural final question is whether the assumptions embedded in our methodology for estimating flows has a large effect on the estimated β_j^{flow}s. Ultimately, we find that abstracting from dividend payments and possible reinvestment does not have a large economic effect on estimating Equation A.15, as Figure A.5 shows that the β_j^{flow}s are centered at the expected value of 1. One reason for this is that index funds' quarterly dividend yields are small (in magnitude) compared to quarterly flows. This gives us confidence that our method for estimating flows is not systematically biasing our results.

Universe of Stocks Most indexes track a specific subset of stocks based on a variety of criteria. The most common is market capitalization (e.g., the S&P 500 index tracks 500 of the largest U.S. stocks with some additional requirements), but many indexes focus on value/growth, industry, and other well-known factors and characteristics. This change has little effect on our example index and fund. We can fit stock segments into our framework by treating stocks that are newly included or excluded from the index as if they were an IPO or a delisting. In this case, total market equity would describe the sum of market capitalization within the targeted market segment.[24]

Float Adjustments Most modern stock indexes apply float adjustments. This is done by using an investable weight factor (IWF), which typically ranges from 0 to 1. If some of a stock's shares outstanding are privately held by, say, company founders and considered to be unavailable to the public, then the shares outstanding are adjusted by multiplying the IWF with shares outstanding to compute *index-eligible shares*. The reason for this adjustment is that acquiring a significant fraction of shares outstanding might be impossible in some cases because those shares are effectively unavailable to the public (ignoring the possibility of short-selling), and, even for smaller fractions, the fund may encounter illiquidity due to fewer publicly available shares.[25] Float adjustments can easily be incorporated into our framework – simply replace all occurrences of shares outstanding with IWF multiplied by shares outstanding, and all corresponding market capitalization calculations

[24]Many indexes hold a fixed number of stocks (e.g., S&P *500* and Russell *2000*), meaning that an addition is almost always coupled with a deletion. In this case, the fund would scale up and down its existing positions based on the difference between the size of the additions and deletions. For example, if the total size (in dollar terms) of the additions is larger than the deletions, then the fund must generate cash to invest in the additions, which is typically achieved by proportionally scaling down all other positions.

[25]In general, the same index provider will use the same float adjustment for the same stock included in its own stock indexes (to the extent that there is overlap). For example, the S&P 500 index uses the same IWFs as the S&P 500 Growth index and the S&P 1500 Composite index. However, it is important to note that there are significant differences in IWF for the same stock across index companies. This suggests that float-adjustments are at least somewhat subjective.

with those using float-adjusted shares outstanding. This means that an index fund could rebalance due to a change in shares outstanding itself or a change in the IWF that effectively changes shares outstanding.

Large Single Changes in Market Composition Many real-world stock indexes will make an exception to the quarterly rebalancing frequency or make other interim adjustments when there are large-enough changes to the composition of the market. For example, many index families (e.g., CRSP, Russell and S&P) will incorporate a large SEO into its index in the days following the SEO, which results in updating index-eligible shares based on the new shares outstanding, and triggers a need to trade based on a shock to at least two ownership ratios. This leads index funds to trade all of the stocks in their portfolios – as in this specific example, the fund would need to scale down all existing positions to purchase part of the SEO – outside of regularly scheduled rebalancing dates. Similarly, a delisting due to acquisition may be reflected in the index in the days that follow, sometimes by adding another stock (because many indexes keep the number of stocks constant). Just like for an SEO, some index families (e.g., CRSP) have a process for quickly adding large IPOs to the index outside of regular rebalancing dates (Sammon and Murray, 2024).

Cash Buffers We do not consider cash in our analysis. While there is no cash in many indexes, most index funds and ETFs hold at least some cash. For the ETFs we consider in Section 4, we find that the median cash holding as reported in ETF Global is about 25 bps of AUM, which lines up with anecdotal evidence from speaking with index fund managers and their internal targets. For a more comprehensive look, we examine cash holdings for all index funds as reported in the CRSP mutual fund data. There, we find a median of about 2 bps of AUM. This is compared to active funds, which have a median of around 1 to 2 percentage points over the last decade. Moreover, our understanding from speaking with index fund managers is that these cash positions do not have much "cash drag" because managers use futures and swaps to maintain full exposure.

Conceptually, our framework is unchanged with a constant cash buffer. This is because we can treat cash as an asset just like any other. The modification comes from taking the original index, scaling it down slightly across all positions, and adding a cash position to come up with a modified or effective index. When there are inflows or outflows, the fund will still scale up and down all holdings (including cash) to maintain proportionality across all positions.

B.5 Discussion and Dispelling Myths

We provide discussion on additional considerations regarding value-weighted index funds, as well as dispel some common myths.

Changing Prices If stock prices change, our example index fund does not need to trade any stock and will still maintain a value-weighted portfolio. This is because, even if prices change, the fund still holds a constant fraction of each stock's shares outstanding, and thus price changes do not affect the stock-specific ownership ratios ($\Omega_{i,t}$). In addition, the fund ownership ratio's numerator and denominator change by the same factor, leaving the ratio unchanged. An increase in prices will change a stock's weight $w_{i,t}$, but holding a constant fraction of each stock's shares outstanding *automatically* adjusts weights from price changes without any need for a change in holdings.

Cash and Proportional Scaling Index funds will often receive cash (through dividends or inflows) or need to generate cash (to satisfy outflows). This means that the fund must put cash to work by investing in stocks, or sell stocks to raise cash. Unlike with changes in prices, dollars moving in to and out of the fund change AUM but not the total stock market capitalization, changing the fund ownership ratio. And, because the ownership ratio determines the fraction of shares outstanding held in every stock, the fund will exactly scale up or down its existing positions to achieve the new fund ownership ratio.[26]

Index Fund Demand Shocks The literature sometimes uses index fund demand as a heterogeneous demand shock for different stocks. Usually, the argument is based on fund *weights*: because most index funds are capitalization-weighted, relatively large-cap stocks have larger weights in the

[26]ETF managers do not deal with cash because authorized participants (APs) deliver or receive a basket in creation/redemption that is representative of what an index fund manager would do if it received cash or needed to generate cash. Often, ETFs use creation/redemption to implement the rebalancing described below.

index fund than relatively small-cap stocks. This is true, and can be observed in Equation A.1. However, when an index fund receives an inflow or outflow, the demand shock is the same for each stock as a percentage of each stock's shares outstanding. That is, when an index fund needs to allocate dollars, it perfectly scales up its share holdings in each stock by the same *percentage*. This preserves value weighting, but does not mean that the index fund buys more shares in large capitalization stocks than small capitalization stocks. The dollar value of purchases will vary – buying a percentage of ownership for large capitalization stocks will have a greater notional value than small capitalization stocks. However, the size of the demand shock, expressed as a percentage of each stock's shares outstanding, is identical across stocks, and this is how demand shocks are typically expressed in demand-based asset pricing models (see, e.g., Haddad et al. (2022)).

The reason why some demand shocks can be properly used to exploit heterogeneity across stocks is because they lead to rebalancing (rather than scaling). This can only occur from a change in the composition of the index via the extensive margin (additions/deletions) and/or the intensive margin (shares outstanding). The commonly used Russell inclusion/exclusion shock at the Russell 1000/2000 breakpoint is correctly deemed a positive demand shock for stocks switching between the Russell 1000 to the 2000 (and negative for those going from 2000 to 1000). This is *not* because the stock moves from the "bottom" of the 1000 to the "top" of the 2000, leading to a larger weight. The reason is that the funds collectively tracking the 2000 have a much larger AUM than those tracking the 1000, relative to each index's total capitalization. That is, the Russell 2000 funds have a much larger fund *ownership ratio* – and therefore a stock moving from the Russell 1000 to 2000 must be purchased to equalize the ownership ratio of the new stock with all of the other stocks in the index.

Index Fund Trading One misconception is that index funds do not need to trade. The reality is that index funds trade quite a bit – by our estimates, a minimum of about 15% of their AUM. One reason is because of regular cash-based events. If cash needs to be put to work (e.g., a stock dividend) or cash needs to be raised (e.g., an outflow), the fund will scale its existing portfolio up or down.

Index funds also engage in a more complex form of trading associated with trying to hold a value-weighted portfolio of the market (or a particular segment of the market). One might sensibly expect that if a fund were to hypothetically own 100% of the market, and as a result shut down flows into the fund, the fund would not need to trade. For example, if some stocks' prices increase and others do not, the index fund still holds "the market" and it engages in no trading. However, this misses intensive and extensive margin rebalancing motives for trade. Our hypothetical fund must still respond to changes in the composition of the market. This means that the fund will purchase a portion of each IPO and issuance, and sell a portion of each buyback. And further, when, e.g., buying a portion of a new issue, it must proportionally sell all other existing stocks to raise enough cash to participate in the issuance – leading to further trading.

What Determines the Stock-Level Passive Ownership Share Many papers studying the effects of passive ownership on financial markets use the *passive ownership share* as the key independent variable of interest. This is typically defined as the fraction of the firm's market capitalization owned by index mutual funds and ETFs. To see how this is determined, consider a stock i which at time t belongs to J indexes. Further, suppose each index has its own fund ownership ratio $\Omega_{j,t}$ and applies its own investable weight factor $IWF_{i,j,t}$. Then, under our framework, the passive ownership share can be expressed as:

$$passive\_ownership\_share_{i,t} = \sum_{j \in J} \Omega_{j,t} \cdot IWF_{i,j,t} = \sum_{j \in J} \frac{aum_{j,t}}{totalme_{j,t}} \cdot IWF_{i,j,t} \tag{A.17}$$

In words, Equation A.17 shows that what determines the passive ownership share is the size of the index funds tracking each individual index, $aum_{j,t}$, relative to the total market capitalization of all stocks in the index $totalme_{j,t}$. This is why, e.g., switching from the Russell 1000 to the Russell 2000 increases passive ownership share$_{i,t}$: because the aum in Russell 2000 tracking funds is larger *relative to total index capitalization* than aum in Russell 1000 funds.

Equation A.17 also highlights why passive ownership share$_{i,t}$ will vary in a given stock over time. This can be because: (1) the stock is added to or dropped from a given index, thus changing the set J, (2) money flows into or out of an index it belongs to, changing $aum_{j,t}$ and thus $\frac{aum_{j,t}}{totalme_{j,t}}$, (3) the index adjusts the investable weight factor for stock i $IWF_{i,j,t}$, or (4) there are compositional changes in the index (e.g., issuance, buybacks, additions and/or deletions) which affect $\Theta_{j,t}$. As an example of this, consider what happened to all S&P 500 index funds in 2020 Q4. This is when the S&P 500 added Tesla, which was roughly 2% of $totalme_{j,t}$, and much larger than the associated dropped stock, Apartment Investment & Management. All other stocks in the S&P 500 index had to be scaled down to make room for Tesla, effectively lowering $\Omega_{j,t}$ and therefore $passive\_ownership\_share_{i,t}$.

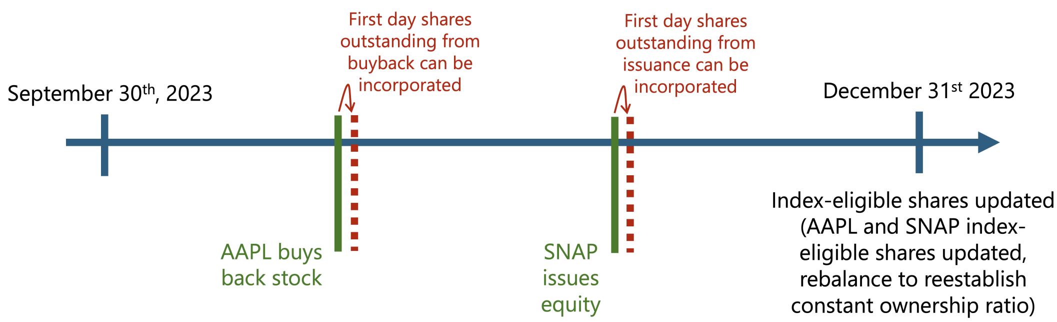
C Hypothetical Index Examples: Rebalancing Frequency and Shares Outstanding Delay

In this appendix, we present examples of hypothetical indexes that vary on the two key choices that define any value-weighted index: rebalancing frequency and shares outstanding delay. The purpose of these examples is to highlight that these are two distinct dimensions of index design.

Figure A.2 presents an example index design that rebalances quarterly and uses a one-day delay to update shares outstanding. In this example, Apple (AAPL) conducted a stock buyback and Snap (SNAP) issued new equity during the quarter. The first day that the changes in shares outstanding can be incorporated is one day after each firm's respective event (red dotted lines). However, the index does not immediately adjust holdings when these events occur or when they are eligible to be incorporated after the delay has been applied. Instead, it waits until the next rebalance period at the end of the quarter – December 31, 2023 – to update index-eligible shares based on the delayed data and to rebalance the portfolio to reestablish a constant ownership ratio with the updated index-eligible shares outstanding.

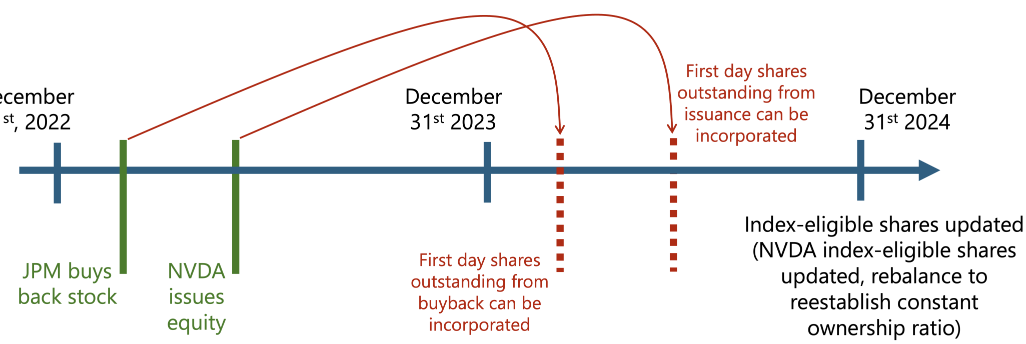
Figure A.3 shows another possible rebalancing policy: an index that rebalances annually and uses a one-year delay to update shares outstanding. In this example, JPMorgan (JPM) bought back stock and Nvidia (NVDA) issued new equity during 2023. Because the delay is one year, the changes in shares outstanding resulting from these events cannot be incorporated until a minimum of a year after the change date. After that point, the shares become eligible to be updated, and are updated

Figure A.2: Example 1: Quarterly Rebalancing, One-Day Delay



Notes. The index updates shares outstanding one day after issuance or buyback events, but rebalances only at the end of the quarter.

Figure A.3: Example 2: Annual Rebalancing, One-Year Delay

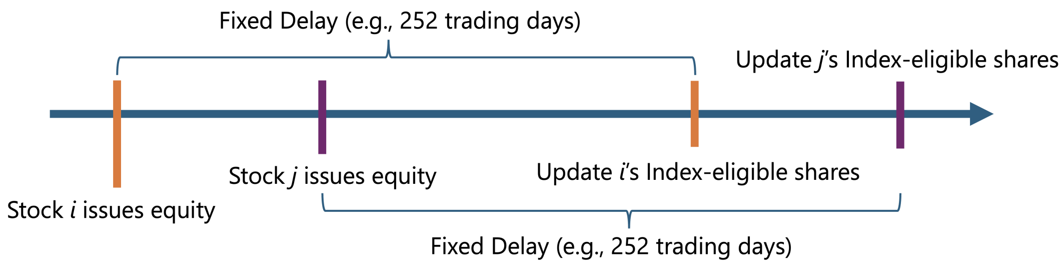


Notes. The index waits a full calendar year after a buyback or issuance before updating index-eligible shares, but the change does not lead to rebalancing until the end of the year.

at the next rebalance date. Since this example index uses an annual rebalancing frequency, it waits until the end of the year – December 31, 2024 – and, at that point, updates index-eligible shares based on the JPM buyback and the NVDA issuance. At that point, the index updates index-eligible shares and rebalances to reestablish constant ownership ratios, even though the events occurred more than a year earlier. Note that these changes are not incorporated into the December 31, *2023* rebalance because the shares outstanding are not sufficiently stale given the index's rebalancing rule.

As a final example, Figure A.4 presents a possible policy of daily rebalancing with a 1-year shares outstanding delay. As discussed in Section 5.2.2, this is designed to isolate the market timing benefit of delayed responses to changes in shares outstanding.

Figure A.4: Daily Rebalancing, One-Year Delay



Notes. The index updates shares outstanding a fixed delay after issuance or buyback events, but rebalances daily. In this example, we use a 252 trading day delay to shares outstanding. So, if there is some issuance and buyback activity on a given day, the index will incorporate these changes into index-eligible shares outstanding exactly a year later, and rebalance to reestablish constant ownership ratios.

Together, these examples illustrate how rebalancing frequency and shares outstanding delay are distinct dimensions of index design. Rebalancing frequency governs when rebalancing occurs, while shares outstanding delay determines what information is used at each rebalance.

D Additional Empirical Results

D.1 Rebalancing Decomposition: All Funds

As highlighted in Section B.3, for a value-weighted index fund, we expect the following point estimates: $\beta^{flow} = 1$, $\beta^{issue} = 1$ and $\beta^{mkt}_j = -1$. Given data issues, we expect that there may be significant variation around these predictions.[27]

Outside of data issues, there is also a simple reason why any one fund may (correctly) have beta estimates very different from our priors: some index funds use weighting schemes that are different from value-weights. Value-weights are critical for the methodology we outline in Section 3, and

[27] In addition to the data issues and timing assumptions mentioned and highlighted above, the beta coefficients may differ from the decomposition-predicted values for several reasons. First, $flow_{j,t}$ is based on an approximation and therefore is measured with error. Similarly, $adddrop_{j,t}$ is calculated assuming all adds/drops happen at the end of the quarter. Although many large and well-known index funds have a "regular" rebalance at or near the end of each quarter (e.g., those tracking S&P, CRSP or Russell indexes), this does not hold universally. Further, some drops are due to delistings, which may trigger index fund rebalancing away from "regular" rebalancing dates. Therefore, the end of quarter value of added or dropped stocks may not reflect the true funding need created by additions/deletions. Finally, given stale data issues in CRSP, $issue_{j,i,t}$ may be measured with error, and for the case of buybacks, many stock indexes are not aware of a buyback until sometimes months after the buyback has occurred, and thus update index-eligible shares only with a delay.

are (up to float-adjustments) used by the largest broad-based index funds in the world (e.g., the four funds in Figure A.1). However, there are many index funds that tend to be thematic (e.g., tracking cash flow yield) and these funds may use different weighting schemes as a selling point.[28] Even for funds that appear to be value weighted, many index funds apply weight caps (e.g., the Nasdaq 100) or float adjustments to firms' shares outstanding, which can be an additional source of rebalancing activity. In fact, changes in the investable weight factor (IWF) mimic the effects of a genuine buyback or equity issuance.[29] As a result, without access to data on float adjustments for every single index, rebalancing driven by IWF changes may not align perfectly with observable net issuance, creating an additional source of index fund trading unexplained by the factors in Equation 4.

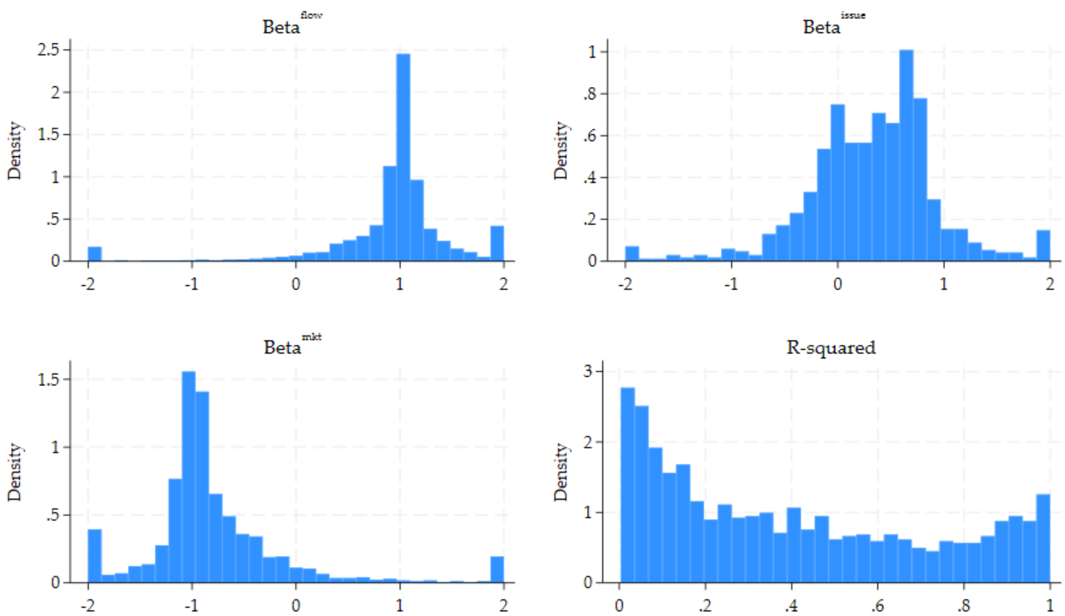
With the caveat that there are potentially a non-trivial number of index funds that do not adhere to our outlined value-weighted index fund methodology, we present histograms of fund-level beta and R-squared estimates in Figure A.5. To minimize the influence of the small number of quarters with serious data errors that lead to extreme beta estimates, we run the regression in Equation A.15 in 4-quarter rolling windows for each fund, and take the median of these estimates at the fund level. The first panel presents the histogram of fund-level estimates for β_j^{flow}. Here, the median value is 0.91, close to our expected value of 1. Further, most of the mass of the distribution is centered around 1, further evidence that flow-based scaling affects funds' holdings in the way our methodology predicts. There are, however, many β_j^{flow} estimates which are significantly less than one but are still positive. These data points are likely from funds that use alternative weighting schemes (including equal weighting, weight caps, weighting by characteristics, etc.), and thus scale up/down positions in response to flows, just not perfectly proportionally to lagged holdings.

The second panel presents the distribution of fund-level estimates for β_j^{issue}. Here, our expected value is again 1, but the median is lower, at 0.32. That being said, the distribution still has

[28]For example, the Invesco S&P 500 Equal Weight ETF (Ticker: RSP) uses equal weights rather than value weights. Weighting schemes can also implement ceilings/floors on individual stock weights, and weighting or re-weighting based on a variable other than market capitalization (e.g., there are some dividend-focus funds that weight based on dividend yield like the SPDR S&P Dividend ETF (Ticker: SDY)).

[29]For example, if an index provider increases a firm's IWF, the float-adjusted shares outstanding rise, requiring the fund to purchase more shares of that firm to maintain a constant ownership ratio. Conversely, a reduction in the IWF would trigger index fund selling, much like a buyback would. This is further complicated by the fact that each index provider applies its own methodology for float adjustments, resulting in variation across indexes in whether and how these changes are implemented.

Figure A.5: Index Fund Trading Betas



Notes. This figure plots histograms of the fund-level betas in Equation A.15. Specifically, we run the regression in equation A.15 in 4-quarter rolling windows for each fund, and take the median of these estimates at the fund level to construct a fund-level estimate. All variables, except for R-squared, have been truncated at ±2.

most of the mass in the positive range between zero and 1. One explanation for the gap between the distribution's median and the expected value of β_j^{issue} is that β_j^{issue} may be subject to more measurement error concerns than β_j^{flow}. This is for at least three reasons. First, as discussed above, shares outstanding data in CRSP may be stale, so our measure of issuance may not align with measures of issuance used by index providers. Further, buybacks may not be recorded and made public for many months after they occur. Second, indexes that use alternative weighting schemes will not necessarily buy after issuance and sell after buybacks. As a specific example of this, the Invesco S&P 500 Equal Weight ETF (Ticker: RSP) has a negative value of β_j^{issue} (despite having a β_j^{flow} close to 1). Finally, as mentioned above, changes in IWF effectively act as net issuance, but we do not have data on IWFs for each fund in our sample.

The third panel presents the estimates for β_j^{mkt}. Here, the median is -0.79, relative to the expected coefficient of -1. While the measurement error concerns are not as extreme as they are for β_j^{issue}, they are still potentially important and may push β_j^{mkt} toward zero. The most likely discrepancy is our timing assumption of additions and deletions occurring at the end of the quarter. While this is close to true on average, there are many indexes with ad-hoc additions or deletions, or alternative rebalancing calendars. For example, the Invesco S&P 500 Low Volatility ETF (ticker: SPLV) is rebalanced after market close on the third Friday in February, May, August, and November. These differences will likely lead us to misestimate the true funding need in $adddrop_{j,t}$, attenuating our estimate of β_j^{mkt}.

Finally, the fourth panel presents the distribution for the R-squared. Here the mean is 0.39, relative to the expected R-squared of 1. This is perhaps unsurprising given all the alternative index methodologies, measurement error issues and data errors in the mutual fund holdings discussed so far. Overall, however, we believe that Figure A.5 provides broad evidence that our methodology accurately describes how index funds work, despite the many exceptions we know exist in the data. Given the wide distribution of point estimates in Figure A.5, we clearly have some passive funds in our sample which are not value-weighted index funds (which matches our survey of many individual index funds and the indexes they track). Nevertheless, Figure A.5 highlights that flow-based proportional scaling, responding to stock-level issuance and buybacks, and adjusting to changes in market composition are integral features of almost every index fund.

Table A.1: Rebalancing Portfolio Returns (All Passive Funds)

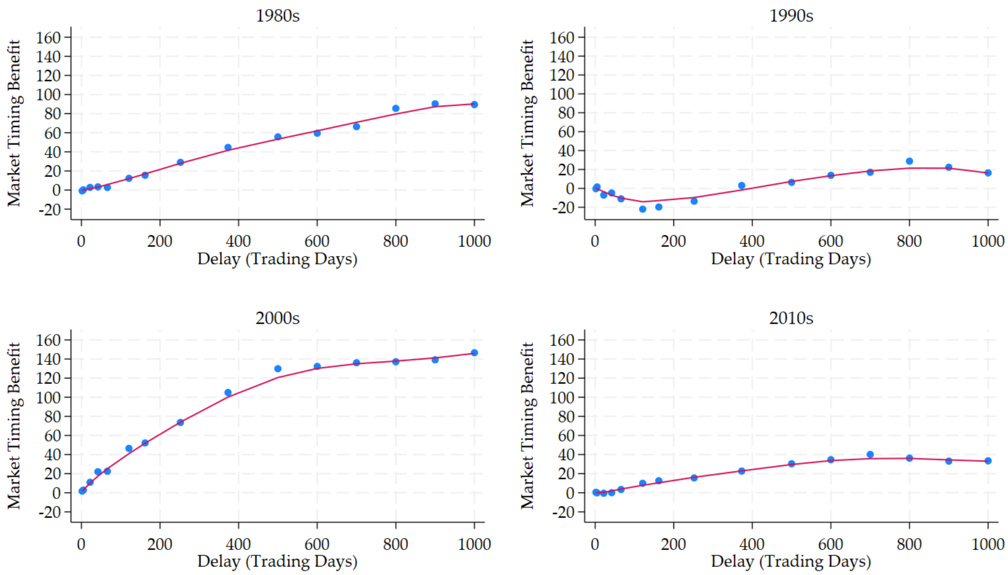
	Intensive Margin Rebalancing			Scaling			Extensive Margin Rebalancing			Total Rebalancing		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Market		0.103***	0.0303		-0.0134	-0.0153		0.0253	0.171***		0.0932***	0.0527
		(0.035)	(0.042)		(0.017)	(0.016)		(0.054)	(0.064)		(0.031)	(0.034)
Size			-0.0111			-0.0189			0.243**			0.0594
			(0.081)			(0.047)			(0.096)			(0.062)
Value			-0.0438			-0.00876			-0.0236			-0.0475
			(0.062)			(0.021)			(0.087)			(0.048)
Profitability			-0.360***			0.0129			-0.315**			-0.345***
			(0.081)			(0.033)			(0.132)			(0.076)
Investment			-0.0238			0.029			-0.734***			-0.230***
			(0.076)			(0.026)			(0.146)			(0.075)
Momentum			0.0757			-0.0256**			0.609***			0.241***
			(0.051)			(0.011)			(0.068)			(0.059)
Reversal			-0.0308			0.0277			-0.11			-0.0229
			(0.051)			(0.019)			(0.081)			(0.059)
Issuance			-0.0247			0.0128			0.423***			0.0189
			(0.065)			(0.032)			(0.111)			(0.064)
Alpha	-4.056***	-4.951***	-2.488**	0.182	0.298	0.147	-4.911***	-5.130***	-7.726***	-3.848***	-4.656***	-2.994***
	(0.920)	(0.989)	(1.225)	(0.480)	(0.573)	(0.521)	(1.526)	(1.454)	(1.584)	(0.871)	(0.837)	(0.950)
Observations	41,080	41,080	41,080	41,080	41,080	41,080	35,698	35,698	35,698	41,153	41,153	41,153
R-squared	0	0.013	0.077	0	0.001	0.008	0	0	0.128	0	0.009	0.133

Notes. This table presents regressions of our rebalancing and scaling portfolios' returns on asset pricing factors for all index funds. The intensive margin captures rebalancing activity for existing positions. The scaling portfolio is based on how the index fund is predicted to scale existing holdings up and down based on flows. The extensive margin portfolio is long/short stocks added/dropped by the index fund. The total rebalancing portfolio is composed of all intensive and extensive margin rebalancing trades. Portfolios are constructed in quarter t to examine returns in $t+1$. We use quarterly index fund holdings data from 1996 Q1 to 2023 Q3 to construct portfolios. Within each quarter, fund-quarter observations are weighted in proportion to their share of total AUM. The unit of observation is at the fund-quarter level. All returns are in percent and annualized. Robust standard errors are in parenthesis.

D.2 All Funds Rebalancing Portfolios

Table A.1 replicates the results in Table 2, but for all index funds, defined as funds in the CRSP mutual fund database with a non-missing index fund flag, or a name that identifies them as an index fund (Appel et al., 2016). To align with the time-series sample in Table 2, we use data from 1996 Q1 to 2023 Q3. If a fund-quarter has no stocks with $\Delta shares_{i,j,t} > 0$ or $\Delta shares_{i,j,t} < 0$, that leg of the portfolio's return is assumed to be the value-weighted market return. This allows us to still assess the future returns to intensive margin rebalancing as a long-short stock portfolio instead of mixing a long-short portfolio with a long-only or short-only portfolio. The unit of observation is at the fund-quarter level, and observations are weighted by the fund's share of total index fund AUM that quarter. The number of observations is slightly smaller in columns 7-9, as there are some index funds that add and drop zero stocks in a given quarter.

Figure A.6: Market Timing by Decade



Notes. This figure plots the market timing benefit by decade. The market timing benefit is the hypothetical index return minus the daily total market return. The hypothetical indexes considered here rebalance daily but vary the delay in responding to changes in shares outstanding.

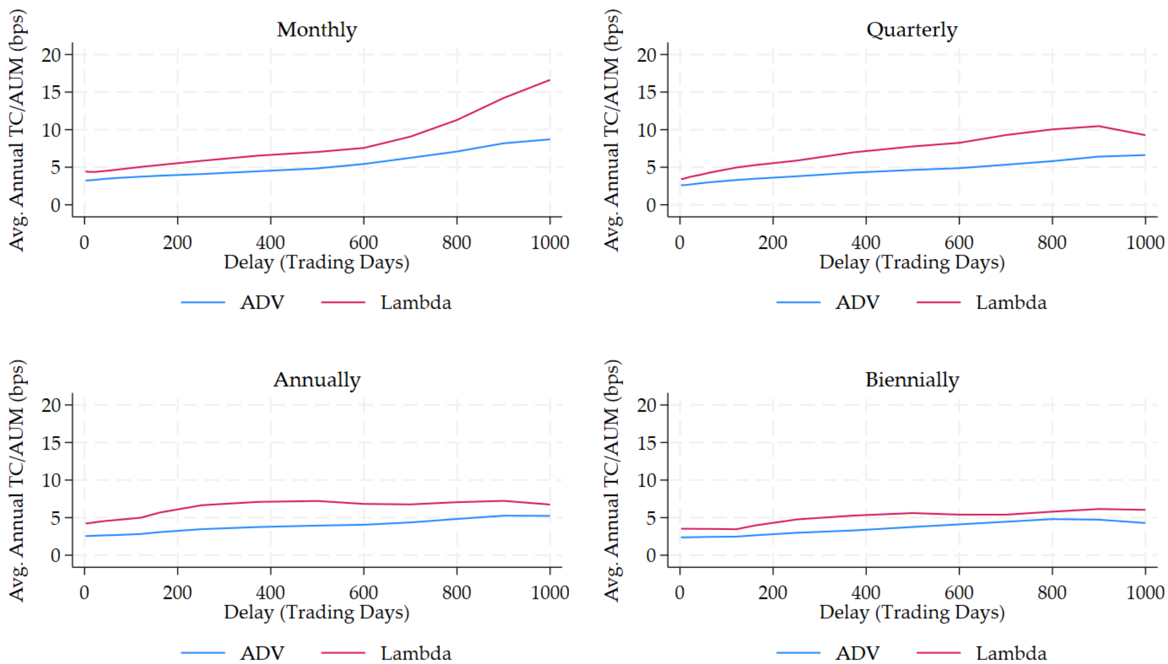
D.3 Market Timing by Decade

Figure 5 shows that delaying responses to compositional changes increases returns on average. We find, however, that sleepy rebalancing does not universally yield higher returns across all subsamples. In Figure A.6, we plot the market timing benefit by decade. In the 1990s, the delayed rebalancing index underperformed in the 6 months to 1 year range because it missed the boom after tech-related IPOs and SEOs. That is, the 1990s happened to be a very good short-term strategy to buy immediately after an IPO or equity issuance, as it effectively allowed an index to ride the tech boom. But, delayed rebalancing dramatically outperforms in the 2000s, when many of those same tech firms that issued shares underperform following the tech crash.

D.4 Trading Cost Robustness

In this subsection, we compare trading cost estimates for our hypothetical index funds under our two primary models: the ADV-based model (Equation 8) and the Lambda-based square-root model

Figure A.7: Estimated Trading Costs by Rebalancing Frequency and Shares Outstanding Delay



Notes. This figure reports the average annual trading costs as a fraction of AUM for hypothetical index funds under different rebalancing frequencies and shares outstanding staleness. Trading costs are estimated using the ADV-based model in Equation 8 and the estimated price impact model in Equation 9, over the 2004-2023 sample.

(Equation 9). Because estimating $\lambda_{i,t}$ requires high-frequency TAQ data, the Lambda-based model is only available beginning in 2004 when millisecond TAQ coverage begins. Figure A.7 shows average annual trading costs as a fraction of AUM across rebalancing frequencies and shares outstanding delays for both models.

Two main patterns emerge. First, the two models yield consistent qualitative results across all strategies: trading costs rise with shorter rebalancing horizons and increase modestly with shares outstanding delay. Second, although the ADV model typically yields lower cost estimates overall, the shape of the cost profiles is similar across both specifications.

D.5 Factor Exposure of Delayed Rebalancing Indexes

In this subsection, we examine the factor exposures of the delayed rebalancing indexes. For simplicity, we examine each monthly rebalanced hypothetical index return varying the delay between

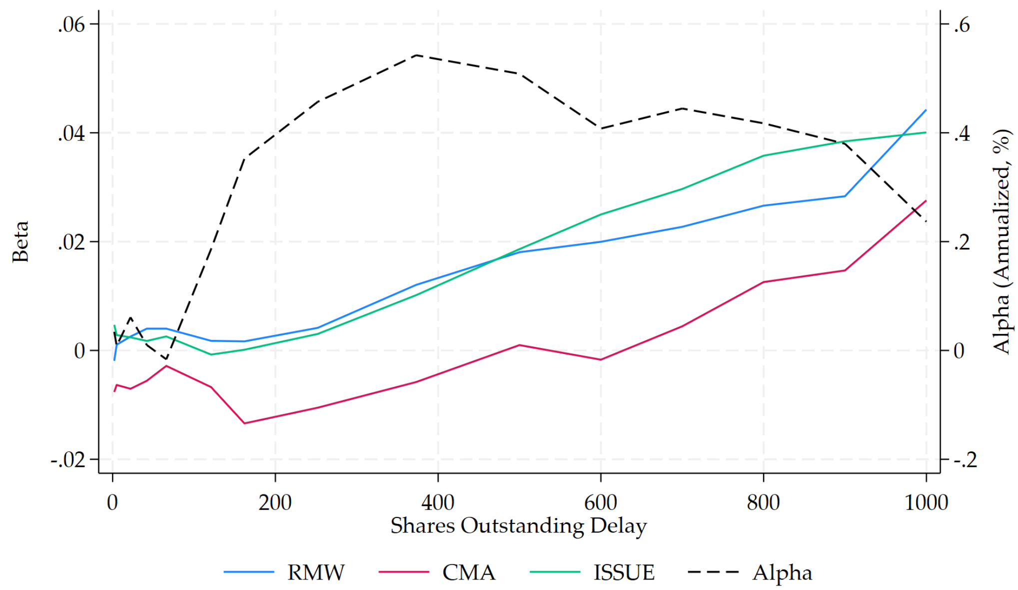
0 and 1000 trading days. We regress the return of each hypothetical index against the market, size, value, investment, profitability, momentum, reversal, and issuance factors. While there are several possible issuance factors, we use the 1-year issuance factor formed using quintiles from Chen and Zimmermann (2022) (based on Pontiff and Woodgate (2008)) for this exercise, which has the most explanatory power among issuance factors for the intensive margin rebalancing portfolio in Table 2. Figure A.8 plots the loadings on the investment (CMA), profitability (RMW), and issuance (ISSUE) factors for each delay on the left y-axis. We omit the other factors, even though they are included in the estimation, because their loadings are close to zero (and the market loading is close to one) for every delay. The figure also presents the alpha for each delay on the right y-axis. The figure shows that longer delays lead to roughly monotonically increasing loadings on the three aforementioned factors. This helps explain where the delayed rebalancing indexes come from – longer delays effectively lead the indexes to load on the profitability and issuance factors. And, because both have positive average returns over our sample, the loadings contribute to beating the market portfolio.

Importantly, however, not all of the alpha is eliminated by controlling for these factors' returns. Figure A.8 shows that alphas increase and peak at a delay of around 1 to 2 years at around 50 bps per year, but are still positive even at a 4-year delay. So, while exposure to well-known risk factors explains some of the returns to sleepy rebalancing, there is still residual variation driven by a delayed response to issuance and repurchase activity that shows up as alpha. We view this as evidence that existing issuance factors (as well as other factors known to predict returns) do not fully capture the relevant issuance effect for a value-weighted total market investor.

D.6 Effect of Including Alternative Issuance Factors on Rebalancing Portfolios' Alphas

In the main body of the paper, Table 2 conditions on an issuance factor from Chen and Zimmermann (2022) forming quintiles based on net issuance over the past year following Pontiff and Woodgate (2008). There are, however, many ways to construct such a factor. In this subsection, we show that using different methods for constructing the issuance factor has minimal effects on our findings.

Figure A.8: Factor Loadings and Alphas of Hypothetical Index Funds



Notes. For each shares outstanding delay, we run a regression of the hypothetical index fund's return on the following asset pricing factors: The market, size, value, investment, profitability, momentum, reversal and issuance. We plot the loadings on the investment (CMA), profitability (RMW) and issuance (ISSUE) factors, as well as the alpha. For each shares outstanding delay, the hypothetical index fund is rebalanced monthly. All quantities are annualized.

For the factors constructed following Daniel and Titman (2006) – which we refer to as 5-Year DT factors – the portfolios are formed based on net issuance over the last 5 years. For the factors constructed following Pontiff and Woodgate (2008) – which we refer to as 1-Year PW factors – portfolios are formed based on net issuance over the last year (all split adjusted). For the data from Ken French's website – which we refer to as Lag 1-Year FF factors because they follow Fama and French (2008a) – the portfolios are formed "*in June of year t based on the change in shares outstanding from the fiscal year end in $t-2$ to the fiscal year end in $t-1$ (all split adjusted)*". As additional robustness, we consider long-short versions of these issuance factors based on quintiles (5-1) and deciles (10-1). In each case, the issuance factor is constructed to have positive returns, i.e., it is long firms that were small net issuers/did buybacks, and short firms with large net issuance.

We then re-estimate the regressions in Table 2, where we regress the rebalancing and scaling portfolios on market, size, value, profitability, investment, momentum, and reversal factors, as well as an issuance factor. We iterate through all of the issuance factors to understand how loadings and alphas vary with the different factors. For brevity, Table A.2 below just reports the loading on the issuance factor and the alpha. To be clear, the full set of factors is included as independent variables in each regression, but we only report the loading on the issuance factor and the alpha.

Consistent with Table 2, the intensive margin rebalancing portfolio typically loads negatively on the issuance factor (note that the second panel corresponds exactly to the results in Table 2). Importantly, the alphas remain negative and statistically significant with every issuance factor. As discussed in the main body of the paper, we believe that these issuance factors do not fully eliminate the alpha because each issuance factor focuses on issuance at a specific horizon or in certain parts of the cross-section (e.g., top versus bottom quintile). In other words, these issuance factors do not necessarily explain the extent of issuance exposure borne by value-weighted market investors.

As additional robustness, Table A.3 below shows the results for all index funds (where again, the second panel of the table corresponds exactly to Table A.1). Broadly, the results are similar to this report's Table A.2, again suggesting that our results are not sensitive to the choice of issuance factor.

Table A.2: Loadings on Issuance Factors and Alphas for Rebalancing Portfolios (Value-Weighted Index Funds)

Portfolio	Overall	Intensive	Scaling	Extensive
		1-Year PW Deciles		
Issuance Factor	0.00127	-0.0806**	0.0061	-0.0591
	(0.046)	(0.036)	(0.018)	(0.072)
Alpha	-3.408***	-1.844**	-0.281	-4.456***
	(0.694)	(0.753)	(0.306)	(1.401)
		1-Year PW Quintiles		
Issuance Factor	0.0453	-0.108**	-0.0096	0.0173
	(0.057)	(0.048)	(0.022)	(0.094)
Alpha	-3.569***	-2.287***	-0.179	-5.158***
	(0.643)	(0.671)	(0.299)	(1.286)
		5-Year DT Deciles		
Issuance Factor	0.128***	0.00543	0.0251	0.142**
	(0.039)	(0.033)	(0.016)	(0.060)
Alpha	-4.059***	-2.730***	-0.347	-5.839***
	(0.664)	(0.661)	(0.324)	(1.240)
		5-Year DT Quintiles		
Issuance Factor	0.0826	-0.0744	0.0318	0.0094
	(0.071)	(0.057)	(0.025)	(0.110)
Alpha	-3.907***	-2.240***	-0.413	-5.150***
	(0.701)	(0.750)	(0.320)	(1.358)
		Lag 1-Year FF Deciles		
Issuance Factor	0.052	-0.0886*	(0.025)	0.028
	(0.056)	(0.046)	(0.022)	(0.089)
Alpha	-3.551***	-2.435***	-0.142	-5.176***
	(0.645)	(0.660)	(0.306)	(1.250)
		Lag 1-Year FF Quintiles		
Issuance Factor	0.017	-0.0607*	(0.009)	(0.111)
	(0.045)	(0.036)	(0.016)	(0.070)
Alpha	-3.452***	-2.497***	-0.185	-4.710***
	(0.640)	(0.652)	(0.297)	(1.242)

Notes. This table which presents regressions of our rebalancing and scaling portfolios' returns on asset pricing factors for value-weighted index funds tracking the S&P 500, 400 and 600, the Russell 1000, 2000 and 3000, the CRSP market-capitalization based indexes. The first column uses the returns to the total rebalancing portfolio, composed of all intensive and extensive margin rebalancing trades. The second column uses the returns to the intensive margin portfolio, composed of rebalancing for stocks held both in quarter t and quarter $t-1$. The third column uses returns to the scaling portfolio based on what the index fund held at the end of the previous quarter. The fourth column uses the returns to the extensive margin rebalancing portfolio, composed of stocks added and dropped by the index fund. Each regression contains the following factors, in addition to the issuance factor denoted: market, size, value, profitability, investment, momentum and reversal. The unit of observation is at the fund-quarter level. All returns are in percent and annualized. Robust standard errors are in parenthesis.

Table A.3: Loadings on Issuance Factors and Alphas for Rebalancing Portfolios (All Index Funds)

Portfolio	Overall	Intensive	Scaling	Extensive
		1-Year PW Deciles		
Issuance Factor	-0.037	-0.0676	0.0101	0.322***
	(0.039)	(0.058)	(0.027)	(0.066)
Alpha	-2.583**	-1.977	0.108	-8.849***
	(1.055)	(1.392)	(0.792)	(1.602)
		1-Year PW Quintiles		
Issuance Factor	0.0189	-0.0247	0.0128	0.423***
	(0.064)	(0.065)	(0.032)	(0.111)
Alpha	-2.994***	-2.488**	0.147	-7.726***
	(0.950)	(1.225)	(0.521)	(1.584)
		5-Year DT Deciles		
Issuance Factor	0.0476	0.0806**	0.0165	0.103*
	(0.035)	(0.037)	(0.023)	(0.061)
Alpha	-3.146***	-2.976**	0.12	-6.440***
	(0.976)	(1.159)	(0.508)	(1.633)
		5-Year DT Quintiles		
Issuance Factor	-0.0361	0.0611	0.023	0.173*
	(0.071)	(0.086)	(0.040)	(0.104)
Alpha	-2.771***	-2.836**	0.106	-6.657***
	(1.012)	(1.301)	(0.482)	(1.621)
		Lag 1-Year FF Deciles		
Issuance Factor	(0.059)	-0.160***	(0.016)	0.209**
	(0.064)	(0.062)	(0.020)	(0.106)
Alpha	-2.717***	-2.047*	0.253	-6.673***
	(0.970)	(1.165)	(0.580)	(1.558)
		Lag 1-Year FF Quintiles		
Issuance Factor	(0.061)	-0.117**	-0.0308**	0.128
	(0.049)	(0.046)	(0.015)	(0.083)
Alpha	-2.637***	-2.053*	0.341	-6.544***
	(0.965)	(1.145)	(0.617)	(1.608)

Notes. This table which presents regressions of our rebalancing and scaling portfolios' returns on asset pricing factors for all index funds. The first column uses the returns to the total rebalancing portfolio, composed of all intensive and extensive margin rebalancing trades. The second column uses the returns to the intensive margin portfolio, composed of rebalancing for stocks held both in quarter t and quarter $t-1$. The third column uses returns to the scaling portfolio based on what the index fund held at the end of the previous quarter. The fourth column uses the returns to the extensive margin rebalancing portfolio, composed of stocks added and dropped by the index fund. Each regression contains the following factors, in addition to the issuance factor denoted: market, size, value, profitability, investment, momentum and reversal. The unit of observation is at the fund-quarter level. All returns are in percent and annualized. Robust standard errors are in parenthesis.