Benchmarking Intensity and Long-Run Returns

Anna Pavlova* and Taisiya Sikorskaya†

March 31, 2020

Abstract

In a model with fund managers evaluated relative to benchmark indices, stocks belonging to these benchmarks have lower expected returns. Motivated by theory, we construct a stock-level measure of benchmarking intensity (BMI) from a unique dataset of 33 U.S. equity indices. The BMI of a stock is computed as the cumulative weight of the stock in all benchmarks, weighted by assets under management following each benchmark. Exploiting a discontinuity in the BMI of stocks at the bottom of the Russell 1000 and the top of the Russell 2000 index, we show that stocks with higher BMIs indeed have lower long-run returns. A growing literature similarly uses the Russell cutoff for identification. We document emerging challenges to this approach: (1) the change in incentives to hold stocks around the cutoff of funds benchmarked to the Russell MidCap/1000/2000 and (2) the growing importance of other benchmarks, including recently introduced CRSP indexes.

JEL Classification: G11, G12, G23

Keywords: Benchmark, index, preferred habitat, benchmarking intensity, index effect, mutual funds, Russell cutoff, optimized sampling, CRSP indexes

*London Business School, apavlova@london.edu

†London Business School, tsikorskaya@london.edu

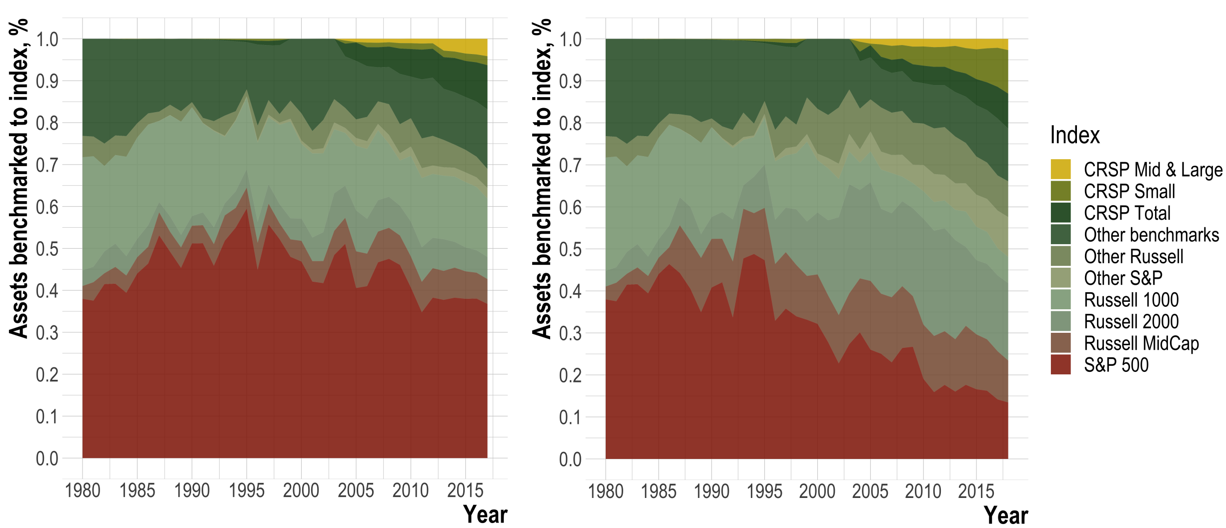
1 Introduction

The asset management industry has been growing in size and importance over time. As of 2018, it has amassed more than \$74 trillion[1] in assets under management. A large fraction of these funds are managed against benchmarks (e.g., the S&P 500, FTSE-Russell indexes, etc.). Benchmarks convey to end investors information about the types of stocks a fund invests in and act as a useful tool for performance evaluation of fund managers. With growing investor appetite for different investment styles, benchmarks are becoming increasingly heterogeneous. Figure 1 plots assets under management of US domestic equity mutual funds, by benchmark. The heterogeneity of benchmarks is apparent from the figure, especially for mid-cap and small stocks. Our objective is to link benchmark membership to stock prices and expected returns, as well as to stock ownership by asset managers.

Figure 1: Assets benchmarked to indices

(a) All stocks

(b) Mid- and small-cap stocks



This figure shows the evolution of the share of benchmark groups in the total assets under management of US domestic equity mutual funds. Mid- and small-cap stocks are in $75^{th}-95^{th}$ *percentile of market capitalization. All reported indices include blend, value and growth types, e.g. Russell 1000 above represents the sum of Russell 1000, Russell 1000 Value and Russell 1000 Growth.*

In this paper, we argue that stocks included in a benchmark form a preferred habitat for fund managers evaluated against that benchmark. That is, benchmarked fund managers have an additional reason to hold stocks in their benchmarks, which results in an inelastic demand for these stocks. We derive a measure, which we call the benchmarking intensity,

[1]Based on BCG report, https://www.bcg.com/en-gb/publications/2019/global-asset-management-will-these-20s-roar.aspx.

that reflects the size of the preferred habitat investors. We define the benchmarking intensity of a stock as the cumulative weight of the stock in all benchmarks, weighted by assets under management following each benchmark. Exploiting a discontinuity in the benchmarking intensity of stocks at the bottom of the Russell 1000 and the top of the Russell 2000 index, we test the prediction of the theory that stocks with higher benchmarking intensities have higher prices and lower expected returns. The data supports these predictions. We also document that, consistent with the preferred habitat view, both active and passive managers buy additions to their benchmarks and sell deletions. Interestingly, these trends are harder to spot with coarse data on benchmarks but they become apparent if we use more granular data.[2]

We start with a simple model that highlights the channel through which a stock's benchmarking intensity affects its price and expected return. The model features fund managers alongside standard direct investors. All investors are risk-averse. A fund manager's compensation depends on performance relative to her benchmark. The model predicts that such performance evaluation makes benchmark stocks the preferred habitat of managers evaluated against that benchmark. The fund manager's higher demand for her benchmark stocks makes prices of these stocks higher in equilibrium and their expected returns lower. This effect is permanent, persisting for as long as the stocks remain in the benchmark. In an equilibrium with heterogeneous benchmarks, the variable that captures the additional (inelastic) demand of benchmarked managers – beyond what the standard risk-return trade-off would predict – is exactly the benchmarking intensity.

In our empirical analysis, we test whether stocks with higher benchmarking intensities have lower long-run returns. Identifying this effect is challenging. Stocks with higher benchmarking intensities are included in more benchmarks and have larger weights in them. Since most of the benchmarks are value-weighted, our measure is closely tied to firm size. As the firm grows, its benchmarking intensity increases but its expected return goes down (due to the size effect). Moreover, as the firm grows, its stock's liquidity improves, which in turn leads to lower expected returns. There may be other confounding effects if one simply compares stocks in major benchmarks to those that are not. Our solution to these challenges is to exploit the cutoff between the Russell 1000 and 2000 indexes, which separates stocks that are very similar in size and other characteristics but differ significantly in terms of their benchmarking intensities. The close-to-random index assignment into the Russell 1000 and 2000 indexes serves as a source of (conditionally) exogenous variation in the benchmarking intensity. In other words, we use index membership as an instrument for the benchmarking

[2]We scrape historical mutual fund benchmarks directly from the website of the US Securities and Exchange Commission.

intensity. We perform our analysis on stocks added to or deleted from the Russell 2000, using stocks close to the cutoff that do not switch indices as the control group.

We find that stocks with higher benchmarking intensities have lower long-run returns. An increased inelastic demand of benchmarked fund managers does indeed lead to significantly lower expected returns of these stocks for horizons of up to 5 years. This result is robust to alternative specifications and we point out that its size depends on whether stocks switching indices multiple times are excluded from the sample, as our theory would suggest. We can interpret our finding as a negative long-run return of a long-short portfolio that buys stocks with high BMI and sells stocks with low BMI. Corroborating the results of Chang, Hong, and Liskovich (2014), we also document price pressure upon index reconstitution (index effect). We link price pressure experienced by a stock to the change in its benchmarking intensity.

We show that both active and passive investors have a considerable fraction of holdings concentrated in their benchmarks and that they rebalance around the Russell cutoffs according to their benchmarks. The majority of recent studies attributed the discontinuities in ownership around the cutoff to passive investors, i.e., index and exchange-traded funds. Consistent with the literature, we find highly significant rebalancing of index additions and deletions for passive funds in line with the direction imposed by their benchmarks. For example, passive funds benchmarked to the Russell 2000 increase their ownership in stocks added to the Russell 2000 by 103bps. These funds also sell deleted stocks in similar proportions. The granularity in our data allows us to see the same pattern in active funds. We find that active funds benchmarked to the Russell 2000 also sell deletions, decreasing their ownership share by 100bps. Active funds benchmarked to the Russell 1000 and Russell MidCap increase their ownership shares in stocks added to the Russell 1000 and MidCap by 23bps and 68bps, respectively.

We highlight several considerations that may affect research design based on the Russell cutoff. First, our model abstracts from transaction costs but, in practice, they are important. Fund managers often deal with them using a portfolio construction approach known as optimized sampling. This approach implies a trade-off between the fund's tracking error and transaction costs and often leads to leaving out the smallest stocks in the benchmark. In our data, it mostly affects stocks added to the Russell 1000 after 2007, when a change in Russell's reconstitution methodology (the introduction of 'banding') increased the potential contribution of these stocks to the tracking error. Second, we document considerable growth of the CRSP indexes after 2013 and explain how index construction poses a confounding problem for using the Russell cutoff. This may inform the growing literature that exploits the Russell cutoff for identification.

Related research. This paper is related to several strands of literature, including equilibrium asset pricing with benchmarked fund managers, index effect, and empirical research on the effects of institutional ownership.

Among theoretical contributions, the first paper to study benchmarking is Brennan (1993). Brennan derives a two-factor asset pricing model in a two-period economy with a benchmarked fund manager. Cuoco and Kaniel (2011), Basak and Pavlova (2013) and Buffa, Vayanos, and Woolley (2014) investigate equilibrium asset pricing effects of delegated portfolio management in dynamic economies. The closest paper to ours in this strand of literature is Kashyap, Kovrijnykh, Li, and Pavlova (2018). None of these works, however, considers heterogeneous benchmarks. The only papers that do are Barberis and Shleifer (2003) and Buffa and Hodor (2018), but they focus primarily on asset return comovement. Barberis and Shleifer consider investors with different styles, which one can interpret as different benchmarks. Our results also highlight heterogeneous habitats of equity fund managers, which relates our work to preferred habitat models of the term structure of interest rates (e.g., Vayanos and Vila (2009)).

Both our theoretical and empirical results are related to the index effect literature. The index effect was first documented by Shleifer (1986) and Harris and Gurel (1986) for additions to the S&P 500 index and subsequently found in many other markets and asset classes.[3] The existence of the index effect challenges the standard theories, which predict that demand curves for each stock are very elastic and therefore index inclusion should have no effect on asset prices and expected returns. The literature offers four competing hypotheses that explain the index effect: the *price pressure hypothesis*, *imperfect substitutes hypothesis*, *information hypothesis*, and *information costs and liquidity hypothesis*. The latter two explain the index effect by reduced information asymmetry and improvements in investor recognition and liquidity upon index inclusion, respectively. Recent literature has contested these views, showing a large and symmetric index effect for stocks with similar liquidity and investor recognition moving between the Russell 1000 and 2000 indexes (Chang, Hong, and Liskovich (2014)).

The price pressure hypothesis posits that demand curves for stocks are downward sloping but only in the short term, as dealers accommodate excess demand for newly included stocks (Scholes (1972)). As dealers demand compensation for providing liquidity around index inclusions, prices of included stocks should rise but then revert quickly back to the

[3] Most of the studies focus on S&P 500 and Russell composition changes, though others also cover such index families as MSCI, DJIA, Nikkei, FTSE, CAC, Toronto Stock Exchange Index, etc. For example, Chen, Noronha, and Singal (2005) documents a long-lasting price increase of the S&P 500 additions, which increases in magnitude through time, but not deletions. Hacibedel and van Bommel (2007) also find permanent price increase for emerging markets indices within the MSCI family.

pre-inclusion levels. Evidence in support of the price pressure hypothesis is mixed; however, index effects lasting for over a month have been documented even for mechanical index reconstitutions[4]. Finally, imperfect substitutes hypothesis posits that long-term demand for securities is not perfectly elastic since securities are not perfect substitutes and therefore the price pressure from index additions is permanent. Our preferred habitat view provides a microfoundation for the imperfect substitutes hypothesis. In our model, fund managers affect stock prices and expected returns for as long as the stocks remain in the benchmark.

The closest empirical work to ours is Chang, Hong, and Liskovich (2014). It is the first paper to build a regression discontinuity design (RDD) on the cutoff between the Russell 1000 and 2000 indexes in order to quantify the price pressure stemming from institutional demand. The paper finds a 5% index effect in the month of addition to and deletion from the Russell 2000. It also documents a decreasing trend in this index effect and attributes it to alleviation of limits to arbitrage. Even though we use the same cutoff for identification, we focus on the long-run returns (twelve months to five years). Chang, Hong, and Liskovich assume that assets under management benchmarked to the Russell 1000 and Russell 2000 are of similar magnitude and focus only on discontinuities in stocks' index weights upon reconstitutions. In constrast, our measure, the benchmarking intensity, takes into account both assets benchmarked to each index and index weights. It also picks up cutoffs of other indices around the Russell 1000/2000 cutoff, which confound the results in the latter part of the sample, and provides an alternative explanation to their findings.

There is a growing body of literature studying corporate outcomes of institutional ownership using the Russell cutoff.[5] In an attempt to reconcile differing findings, there has been debate on the implementation of the identification strategy that exploits the Russell cutoff. Some early papers used June (post-announcement) index weights in their empirical approach. It was later criticized by Appel, Gormley, and Keim (2019a) and Wei and Young (2017), among others, who highlighted a mechanical relationship between institutional ownership and June weights. Importantly, most researchers cannot observe the true running variable used to assign stocks into indices. Hence, identification requires either predicting index membership using public data as of May or assuming that controlling for researcher-constructed running variable achieves conditional exogeneity of the index dummy. We address all aforementioned points by (1) using Russell's proprietary running variable, which delivers 99% accuracy of assignment and (2) orthogonalizing any remaining error by instrumenting index membership. Furthermore, we discuss the implications of industry-wide

[4]See, for example, Kaul, Mehrotra, and Morck (2000).

[5]The list of papers includes but is not limited to: Appel, Gormley, and Keim (2019b), Glossner (2018), Heath, Macciocchi, Michaely, and Ringgenberg (2018), Schmidt and Fahlenbrach (2017), Appel, Gormley, and Keim (2016), Crane, Michenaud, and Weston (2016).

practice of optimized sampling, which affects funds' rebalacing around the cutoff, and shed light on the identification threat from the entry of the CRSP indexes with an overlapping cutoff.

The paper proceeds as follows. Section 2 explains the implications of heterogeneous benchmarks for stock returns. In Section 3, we construct the measure of benchmarking intensity, describe our identification strategy and present estimation results. We discuss ownership trends and relate them to funds' preferred habitats in Section 4. Section 5 concludes. Omitted details are relegated to the Appendix.

2 Model of Delegated Asset Management with Heterogeneous Benchmarks

To illustrate the main mechanism, we first develop a simple model of asset prices in the presence of benchmarking. It builds upon Brennan (1993) and Kashyap, Kovrijnykh, Li, and Pavlova (2018) and introduces heterogeneous fund managers whose performance is evaluated relative to a variety of benchmarks. The goal of the model is to derive a relationship between benchmarking intensity, our measure of capital that is inelastically supplied by fund managers, and stock returns.

2.1 Model

Except for the presence of fund managers, our environment is standard. There are two periods, $t = 0, 1$. The financial market consists of a riskless asset with an exogenous interest rate normalized to zero (e.g., a storage technology) and N risky assets paying cash flows $D_i,\ i = 1, \dots, N$ in period 1. The cash flows of the risky assets are given by

$$D_i = \overline{D}_i + \beta_i Z + \epsilon_i, \quad c_i > 0,\ i = 1, \dots, N,$$

where $Z \sim N(0, \sigma_z^2)$ is a common shock and $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ is an idiosyncratic one. The vectors $D \equiv (D_1, \dots, D_N)^\top$ and $S \equiv (S_1, \dots, S_N)^\top$ denote vectors of period-1 cash flows and period-0 risky asset prices, respectively. Period-1 risky asset prices equal to D. The risky assets are in fixed supply of $\overline{\theta} \equiv (\overline{\theta}_1, \dots, \overline{\theta}_N)$ shares. It is convenient to introduce the notation $\Sigma \equiv \Sigma_z + I_N \sigma_\epsilon^2$ for the variance-covariance matrix of cash flows D, where Σ_z is a $N \times N$ matrix with a typical element $\beta_i \beta_j \sigma_z^2$ and I_N is an $N \times N$ identity matrix. We also set $\overline{D} \equiv (\overline{D}_1, \dots, \overline{D}_N)^\top$ and $\beta \equiv (\beta_1, \dots, \beta_N)^\top$.

There are J benchmark portfolios that are used for performance evaluation. Each benchmark j is a portfolio of $\omega_j \equiv (\omega_{1j}, \ldots, \omega_{Nj})^\top$ shares of the assets described above. Some components of ω_j can be zero.

There are two types of investors: direct investors and fund managers. Direct investors, whose mass in the population is λ_D, manage their own portfolios. Fund managers manage portfolios on behalf of fund investors. Fund investors can buy the riskless asset directly, but cannot trade stocks; they delegate the selection of their portfolios to portfolio managers. The managers receive compensation from fund investors. Each manager is evaluated relative to a benchmark. We denote the mass of managers evaluated relative to benchmark j by λ_j.[6] All investors have a constant absolute risk aversion utility function over terminal wealth (or compensation), $U(W) = -\exp^{-\gamma W}$, where γ is the coefficient of absolute risk aversion.

The terminal wealth of a direct investor is given by $W = W_0 + \theta_D^\top(D - S)$, where the $N \times 1$ vector θ_D denotes the number of shares held by the direct investor, and W_0 is the investor's initial wealth. The direct investor chooses a portfolio θ_D to maximize his utility $U(W)$. A fund manager's j compensation w_j consists of three parts: one is a linear payout based on absolute performance of the fund, the second piece depends on the performance of the fund relative to the benchmark portfolio j, and the third is independent of performance (c). Specifically,

$$w_j = aR_j + b(R_j - B_j) + c, \quad a \geq 0,\, b > 0$$

where $R_j \equiv \theta_j^\top(D - S)$ is the performance of the fund's portfolio and $B_j \equiv \omega_j^\top(D - S)$ is the performance of benchmark j.[7] The parameters a and b are the contract's sensitivities to absolute and relative performance, respectively. The fund manager chooses a portfolio of θ_j shares to maximize his utility $U(w_j)$.

2.2 Portfolio Choice and Asset Prices

The optimal portfolio of the direct investors is the standard mean-variance portfolio:[8]

$$\theta_D = \frac{1}{\gamma}\Sigma^{-1}\left(\overline{D} - S\right). \tag{1}$$

[6] For simplicity, we assume that each fund investor employs one fund manager, but this can easily be relaxed.

[7] Ma, Tang, and Gómez (2019) analyze compensation of fund managers in the US mutual fund industry and provide evidence supporting our specification here. Endogenizing this compensation structure is beyond the scope of this paper; see Kashyap, Kovrijnykh, Li, and Pavlova (2020) who derive it as part of an optimal contract.

[8] We omit proofs in the main text and relegate them to Appendix B, available upon request.

In contrast, the fund managers do not have the same risk-return trade-off as direct investors, because of their compensation contracts. The optimal portoflio of manager j is given by

$$\theta_j = \frac{1}{\gamma(a+b)}\Sigma^{-1}\left(\overline{D} - S\right) + \frac{b}{a+b}\omega_j. \tag{2}$$

The fund manager splits his risky asset holdings across two portfolios: the mean-variance portoflio (the first term in (2)) and the benchmark portfolio (the second term). The latter portfolio arises because the manager hedges against underperforming the benchmark. Consistent with the preferred habitat view, the manager thus has a higher demand for stocks in her benchmark. Notice that the demand for the benchmark portfolio ω_j is inelastic. It does not depend on the riskiness of the assets and depends only on the parameters of the compensation contract. It follows that, *ceteris paribus,* stocks with a higher benchmark weight have a higher weight in the fund manager's portfolio.

By clearing markets for the risky assets, $\lambda_D\theta_D + \sum_{j=1}^{J}\lambda_j\theta_j = \overline{\theta}$, we compute equilibrium asset prices.

$$S = \overline{D} - \gamma A\Sigma\left(\overline{\theta} - \frac{b}{a+b}\sum_{j=1}^{J}\lambda_j\omega_j\right), \tag{3}$$

where $A \equiv \left[\lambda_D + \frac{\sum_j \lambda_j}{a+b}\right]^{-1}$ modifies the market's effective risk aversion.[9]

Equation (3) elucidates the determinants of the index effect in our model. The index effect manifests itself through the benchmarking-induced price pressure term $\sum_{j=1}^{J}\lambda_j\omega_j$. This term reflects the cumulative inelastic demand of fund managers and motivates our benchmarking intensity measure used in the empirical part of the paper. Equation (3) implies that if a stock gets added to a benchmark or if its weight in a benchmark increases, its price goes up. Another implication is that the larger the mass of fund managers (λ_j's) following a benchmark, the higher the benchmarking-induced price pressure and hence the bigger the index inclusion effect. The more benchmarks a stock belongs to and the bigger its weight in the benchmarks, the more demand from fund managers it attracts and therefore the higher the stock's price. Another implication is that the larger the mass of managers following a benchmark (λ_j), the higher the price pressure.

[9]Our model can be extended to incorporate passive managers, who simply hold the benchmark portfolio. Suppose the total mass of fund managers benchmarked to index j, λ_j, consists of a mass λ_j^P of passive managers and a mass λ_j^A of active. Then the expression for stock prices is:

$$S = \overline{D} - \gamma A\Sigma\left(\overline{\theta} - \sum_{j=1}^{J}\left[\frac{b}{a+b}\lambda_j^A\omega_j + \lambda_j^P\omega_j\right]\right)$$

Our next set of predictions is about the expected stock returns (or the cost of equity). The expected return of stock i, expressed as a per-dollar return $E[r_i] \equiv (\overline{D}_i - S_i)/S_i$, is given by

$$E[r_i] = \frac{\gamma A}{S_i}\beta_i \sigma_z^2 \beta^\top \left(\overline{\theta} - \frac{b}{a+b}\sum_{j=1}^{J}\lambda_j \omega_j\right) + \frac{\gamma A}{S_i}\sigma_\epsilon^2 \left(\overline{\theta}_i - \frac{b}{a+b}\sum_{j=1}^{J}\lambda_j \omega_{ij}\right) \tag{4}$$

Equation (4) implies that the price pressure we discussed above is permanent, and it lasts for as long as a stock remains in the fund managers' benchmarks. Therefore, *ceteris paribus*, stocks with higher benchmarking intensities have lower expected returns. Furthermore, if a stock's benchmarking intensity goes up (e.g., due to an index inclusion), its price should rise upon announcement and the expected return after the announcement should be lower.

In summary, our model produces the following predictions:

Prediction 1: Stocks with higher benchmark intensities have lower expected returns.

Prediction 2: If a stock's benchmarking intensity goes up (e.g., due to an index inclusion), its price should rise.

Prediction 3: If a stock's benchmarking intensity goes up, the funds' ownership of the stock ($\sum_j \theta_{ij}$) should rise.

3 Benchmarking Intensity and Long-Run Returns

3.1 Empirical Measure of Benchmarking Intensity

Guided by the model, we calculate the benchmarking intensity (BMI) for stock i in month t as

$$BMI_{it} = \sum_{j=1}^{J}\omega_{ijt}\lambda_{jt}$$

where $\lambda_{jt} = \frac{\Lambda_{jt}}{\sum_{j=1}^{J}\Lambda_{jt}}$ is the assets under management (AUM) share of mutual funds benchmarked to index j in month t, with Λ_{jt} being their dollar AUM[10], and ω_{ijt} is the weight of stock i in index j in month t. In our annual panel, we use BMI_{it} calculated as of September.[11]

Because the largest market indices are value-weighted, BMI is closely tied to the market cap of the company. In the literature, firm size was related to stock returns, which poses a challenge for our empirical analysis.

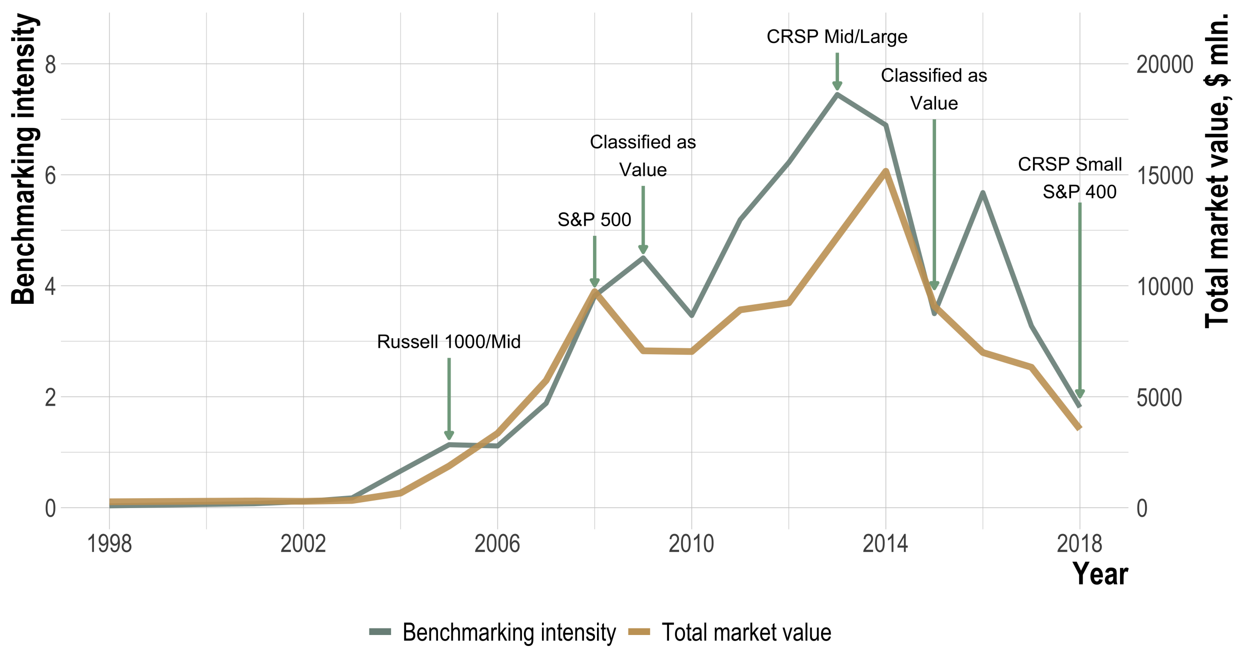
Even though benchmarking intensity is typically slow-moving, considerable variation comes from index membership. If a stock switches indices, for example, moves out of the S&P

[10] We also experimented with scaled AUM shares. For example, we used shares scaled to the industry AUM in 2014 as it is the time when the CRSP indexes were introduced, completing our universe of benchmarks.

[11] The reason is that we want to avoid sorting on the initial price pressure that occurs mostly in June as discussed in later sections.

500 index, its BMI changes. A useful illustration is a natural gas company Range Resources Corp. (ticker *RRC*). Figure 2 depicts a year-on-year evolution of its benchmarking intensity. Despite the evident comovement between size and benchmarking intensity, the latter has slightly more variation due to the changing index membership and index asset flows: in 2005 *RRC* joins the Russell 1000 and Russell MidCap, in 2008 – S&P 500, in 2012 *RRC* gets into the CRSP Mid Cap, in 2018 it exits the S&P 500 and gets added to S&P 400 and the CRSP Small Cap.

Figure 2: Benchmarking Intensity of *RRC*

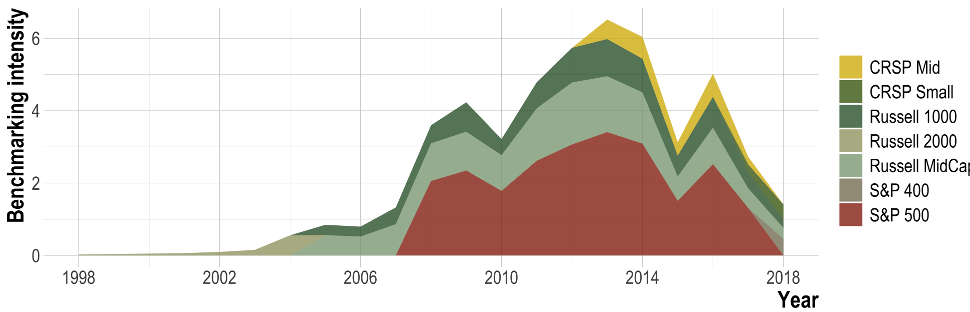


This figure plots the benchmarking intensity (left axis) and the total market value (right axis) of RRC stock over time. Arrows point to the years when the stock was added to the benchmarks. Total market value is scaled by 2,500.

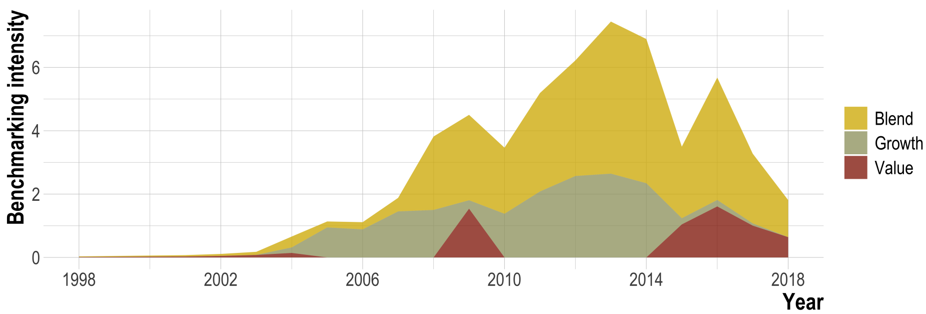
Figure 3 illustrates the contribution of membership in each index into the benchmarking intensity of *RRC* (Panel (a)). Even though most of the time the stock's S&P 500 membership contributes over 50%, size and variation of other components are significant. Panel (b) of the same Figure shows how much different benchmark styles (i.e., value, growth, and blend) contribute to *RRC*'s BMI. In our data, we only have style indices for the Russell and CRSP families, so the rest is attributed to blend. Even with this limitation, it is apparent that style benchmarks occupy a considerable fraction of BMI. These two illustrations highlight one of the key contributions of our measure – it takes into account the heterogeneity of benchmarks and overlaps between them.

Figure 3: Decomposition of the Benchmarking Intensity of *RRC*

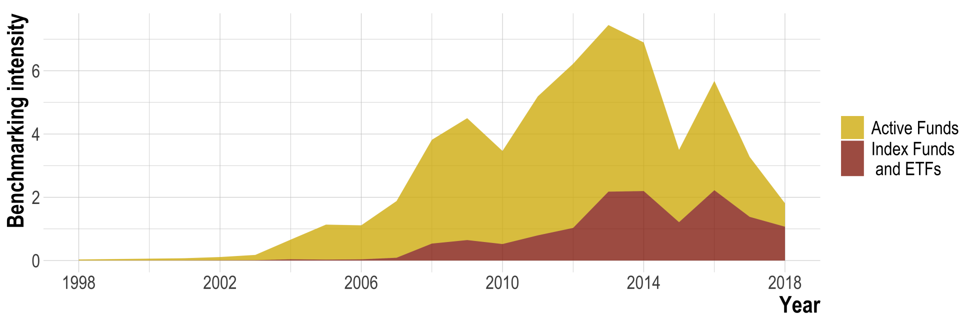
(a) Index Group



(b) Index Style



(c) Fund Type



These figures plot the evolution of each component of the benchmarking intensity of RRC stock over time. Figure (a) plots index groups, each including Blend, Value and Growth indices. Figure (b) plots Russell and CRSP style components. Figure (c) plots the contribution of active and passive funds.

Since the benchmarking intensity measure is built using the AUM of both active and passive funds, there is a variation coming from the relative importance of these two fund types as depicted in Panel (c) of Figure 3. The BMI of *RRC* is dominated by the benchmarking demand from active funds, which changes in 2018 as the stock exits the S&P 500 universe. This illustrates another important contribution of BMI – unlike passive ownership, a measure of institutional demand used in the extant literature – the BMI accounts for the inelastic demand of active funds as well.

Table 1 documents descriptive statistics for BMI in our sample. S&P 500 stocks have the highest average BMI, while the BMI of almost all Russell 2000 stocks is below the sample average. The reported statistics also highlight the increasing heterogeneity of benchmarks for the U.S. equities: the average number of benchmarks increased from 7 to 10 and the benchmark Herfindahl-Hirschman index went down from 1100 to 740. Together, value and growth indices are at least as important as blend indices, contributing on average almost 60% to BMI. Furthermore, active funds contribute 83% to the BMI over the full sample period, even though their share declined to an average of 65% in the recent 5 years.

Table 1: Properties of benchmarking intensity

	By time period					**By benchmark**					
	Full sample	1998-2000	2001-2006	2007-2012	2013-2018	S&P 500	Russell 1000	Russell 2000	Russell MidCap	Russell Value Indexes	Russell Growth Indexes
Average BMI, bps	3.2	3.3	3.3	3.4	2.9	16.2	9.0	0.5	3.4	3.2	3.6
St. dev. of BMI, bps	12.4	15.5	13.8	12.0	10.3	27.7	20.9	0.5	3.1	11.6	13.9
Minimum BMI, bps	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
Maximum BMI, bps	439.4	439.4	439.4	356.3	308.3	439.4	439.4	4.8	37.3	439.4	439.4
Average no. of benchmarks	9.5	7.5	9.0	9.9	10.4	9.6	11.0	9.5	11.5	10.5	10.6
Average benchmark HHI	969	1103	1077	1064	736	5224	2749	138	2184	1080	989
Average contribution of indices, %:											
- S&P 500	55.5	59.3	58.9	57.6	48.3	55.5	55.7	43.3	49.6	55.7	54.8
- S&P 400	20.1		21.5	20.2	18.7	21.5	23.0	14.0	23.0	20.5	19.5
- Russell 1000	24.1	36.4	26.6	22.0	18.7	19.3	24.1	16.9	23.8	25.1	22.8
- Russell 2000	80.6	76.2	81.9	85.8	75.2	49.9		80.6		81.7	79.4
- Russell MidCap	28.8	24.5	29.8	33.3	24.9	20.8	28.8	27.9	28.8	26.3	32.1
- CRSP Large and Mid	8.2				8.2	6.5	8.2	6.6	10.7	8.2	8.1
- CRSP Small	11.5				11.5	9.8	15.0	10.2	15.0	11.8	11.2
Average contribution of styles, %:											
- blend	42.9	31.5	37.6	43.3	51.6	56.8	38.8	43.4	33.7	47.2	45.3
- value	26.4	25.6	26.3	29.1	24.2	21.5	28.9	25.8	31.1	30.0	18.5
- growth	30.7	42.9	36.1	27.6	24.3	21.7	32.3	30.7	35.2	22.8	36.2
Average contribution of fund types, %:											
- active	83.0	96.6	93.7	89.1	65.4	79.3	81.9	88.8	82.7	86.1	87.1
- passive (index and ETFs)	17.0	3.4	6.3	10.9	34.6	20.7	18.1	11.2	17.3	13.9	12.9

This table reports the descriptive statistics for benchmarking intensity. Columns 'By time period' show statistics for the respective period. Columns 'By benchmark' show statistics for stocks that belong to the respective benchmark. BMI statistics (average, standard deviation, minimum, and maximum) are in basis points. Contribution is in percentage points. Contribution of indices is average of the ratios of BMI coming from the AUM benchmarked to an index to total BMI of the stock. Contribution of indices is across index styles, e.g., line for the Russell 1000 includes blend, value, and growth. Average number of benchmarks is for a stock. Benchmark HHI is a Herfindahl-Hirschman index computed using benchmark AUM shares (scaled by 10000, so that index below 1500 indicates an unconcentrated industry). Averages are simple arithmetic means across stock-years.

3.2 Dataset

The main sample is a yearly panel of stocks which were the Russell 3000 constituents in 1998-2018. The main three pillars of data are historical benchmark weights, mutual fund holdings, and stock characteristics.

In contrast to the previous studies, the dataset is granular with respect to benchmark information. It includes primary prospectus benchmarks scraped directly from historical fund prospectuses available on the website of the U.S. Securities and Exchange Commission[12] and augmented with a Morningstar snapshot. The scraping procedure and its validation are described in detail in Section A.1 in the Appendix. We obtain benchmark constituent data from the following sources. All the constituent weights for 22 Russell benchmark indices are from FTSE Russell (London Stock Exchange Group). The Russell indexes include (all total return in USD): Russell 1000/2000/2500/3000/3000E/Top 200/Mid Cap/Small Cap Completeness as well as their Growth and Value counterparts. Constituent weights for the S&P 500 TR USD and S&P 400 MidCap TR USD are from Morningstar and available from September 1989 and September 2001, respectively, to October 2015. We construct constituent weights for S&P 500 before September 1989 and after October 2015 manually from constituent lists and prices available through CRSP. We generate the S&P 400 weights from holdings of index funds (Dreyfus and iShares).[13] Weights for the CRSP US indexes are accessed via Morningstar and are available from 2012. These indices include (all total return in USD): Total Market, Large Cap, Mid Cap, Small Cap as well as their Growth and Value counterparts.

Our benchmark data has two advantages to prior research. First, the benchmark information is a dynamic panel encompassing benchmark changes.[14] Therefore, it accurately reflects the benchmark used by funds at any point in time. Secondly, we obtain Russell index data from FTSE Russell directly: our dataset includes proprietary total market values (capitalization) as of the rank day in May and provisional constituent lists available before the reconstitution day in June.

In fund rebalancing analyses, we use holdings available in the CRSP Mutual Fund Database (CRSP, June 2010 - December 2018) and Thomson Reuters S12 (TRS12, March 1980 - December 2018). Our main source is TRS12 and we use CRSP to add funds for which data is not available in TRS12. Moreover, CRSP is used to validate the net assets of the funds and pull various fund-level characteristics, such as returns, expense ratios, equity percentage, and others. We merge the databases using MFLINKS following steps described in Section A.2 in the Appendix. We follow several data validation procedures and impose

[12] Follow https://www.sec.gov/edgar/searchedgar/mutualsearch.html

[13] Since the S&P 400 index is relatively small, these weights do not contribute much to the analysis. We do not include the S&P 600 index because its share is even smaller and the holdings-based weights are not of sufficient quality.

[14] See Appendix, in which we show that our scraping procedure picked up such important benchmark changes as Vanguard's move from the MSCI to CRSP indexes in 2013.

typical mutual fund filters, which are outlined in the Appendix as well (Section A.4).

We use several databases of ownership data. Mutual-fund ownership share for any stock is computed as the percentage of shares held by funds of a certain type in the total number of shares outstanding for the stock (using TRS12 and CRSP as above).

We classify funds into active and passive based on the *index_fund_flag* in CRSP and by screening fund names. All ETFs in our sample are classified as passive. A fund is classified as an ETF if its *et_flag* in CRSP is non-empty or it has *exchange-traded* or *etf* in its name. We manually resolve and exclude exceptions when the same portfolio has share classes of both active and passive funds. Detailed steps as well as the textual rules we deploy for the screening are listed in Section A.6 of the Appendix.

We use daily fund returns from CRSP and benchmark returns from Morningstar in order to compute tracking errors (net).

Stock data comes from standard sources. We take daily returns, prices, adjustment factors, and bid and ask prices from CRSP.[15] Market, risk-free rate, and factor returns are from Ken French's Database. All fundamental accounting data comes from Compustat. We use CRSP-Compustat linking table and take into account release dates to make sure that the variables are available to the public by the rank date in May.

We report the descriptive statistics of the main calculated variables used in analysis in Table 10 in the Appendix.

3.3 Evidence From Russell Indices Reconstitution

Our goal is to test the relationship between two stock-level variables: long-run returns and the benchmarking intensity. We exploit the cutoff between the Russell 1000 and 2000 indexes, which separates stocks that are very similar in size and other characteristics but differ significantly in terms of their benchmarking intensities. The close-to-random index assignment into the Russell 1000 and 2000 indexes serves as a source of (conditionally) exogenous variation in the benchmarking intensity. In other words, we use predicted index membership as an instrument for the benchmarking intensity.

3.3.1 The Russell Index Cutoff

Russell indexes undergo a yearly reconstitution at the end of June. The reconstitution is a two-step process: assigning a stock to an index and determining the weight of the stock in that index.

[15] Returns are adjusted for delisting in the standard way.

The first step is solely based on the ranking of all eligible securities by their total market capitalization on the rank day in May. For most of the years in our sample, the rank day falls on the last trading day in May[16]. Russell uses its broadest Russell 3000E index as the universe of eligible securities together with newly admitted stocks[17]. Ranks are computed based on the proprietary measure of the total market capitalization of eligible securities. This proprietary measure has been made available to us by Russell[18][19] and hence we are able to replicate the assignment rule very closely.[20]

In the second step, each stock in the index is assigned a weight based on its float-adjusted market capitalization. To define the adjustment, Russell uses proprietary float factors, which we can infer from total and float-adjusted market capitalization. These factors do not affect index assignment but they explain some variation in the benchmarking intensity due to their direct relationship with index weights: all else equal, stocks will have lower index weight if the float adjustment is larger, and hence lower BMI.

Before 2007, a firm would be assigned to the Russell 2000 index if and only if its total market value rank falls between 1000 and 3000. Since the assignment is based on ranks, firms cannot manipulate it.[21] Moreover, an idiosyncratic shock to the market value on the rank date can bring the stock to the other side of the cutoff. Hence, the assignment is as good as random. Panels (a) and (b) of Figure 4 plot index weights and benchmarking intensities of stocks on the rank day (May 31^{st}) in 2006. All stocks to the right of 1000^{th} rank cutoff in May are assigned to the Russell 2000 in June. Because of the discontinuity in index weights at the cutoff, our benchmarking intensity measure also has a discontinuity at the cutoff.

In order to reduce the turnover between indices, FTSE Russell introduced a 'banding' policy in 2007. According to the new rule, a stock is assigned to the Russell 2000 index if and only if:

- it was in the Russell 2000 in the previous year and its total market value rank in May falls between the left cutoff $(1000 - c_1)$ and 3000[22]

- it was in the Russell 1000 and its total market value rank in May falls between the

[16]Exceptions are recent years, when the rank days were: 05/27/2016, 05/12/2017, and 05/11/2018.

[17]See the details on the methodology in the official and publicly available guide.

[18]We match this measure to the May Russell 3000E constituent lists as well as the preliminary constituent lists from June in order to arrive at the universe of eligible securities. The preliminary lists have also been provided by Russell.

[19]We performed our analysis with the market value measure constructed from CRSP and Compustat as in Chang, Hong, and Liskovich (2014) as well. This measure delivers qualitatively identical main results.

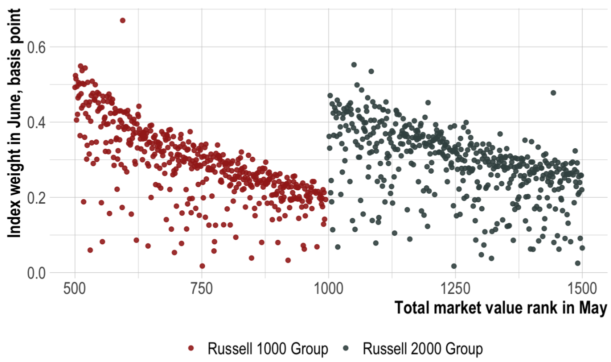
[20]For comparability with other papers, we include a Table 8 in the Appendix.

[21]Typically, bunching is formally tested for with McCrary (2008) test but since the assignment variable is a rank, which is relative to other stocks, bunching is not possible.

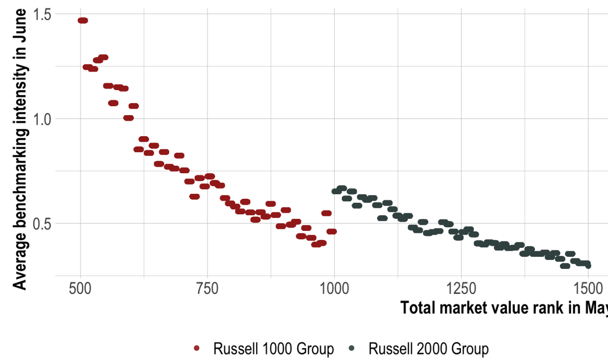
[22]The rule is similar for stocks moving to the Russell 2000 from below, i.e., around 3000 rank. We are omitting it here for brevity.

Figure 4: Discontinuities in Index Weights and BMI before and after 2006

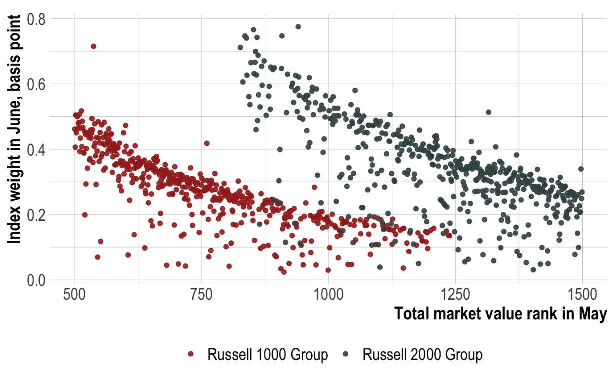
(a) Index weights upon reconstitution in 2006



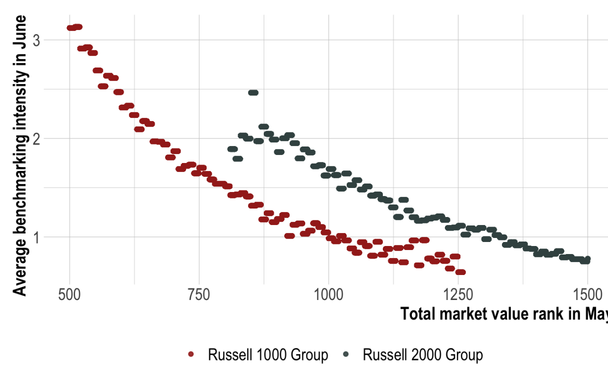
(b) Benchmarking intensity in 1998-2006



(c) Index weights upon reconstitution in 2012



(d) Benchmarking intensity in 2007-2018



This figure plots index weights and benchmarking intensity against the total market value rank on the rank day in May. Index weights are a snapshot on the reconstitution date in 2006 (June 30^{th}*) and 2012 (June* 29^{th}*). Benchmarking intensity is averaged for constituents of each index across bins of 10 stocks and over the relevant period. Russell 1000 Group includes the Russell 1000 and Russell MidCap (blend, value, and growth). Russell 2000 Group includes the Russell 2000 (blend, value, and growth).*

right cutoff $(1000 + c_2)$ and 3000.

The band, that is, the range of ranks between $(1000 - c_1)$ and $(1000 + c_2)$, is still based on a mechanical rule but it changes each year with the distribution of firm sizes around the cutoff.[23] Because of banding, the turnover between indices went down significantly, as intended.[24] We list the number of additions and deletions per year in Appendix, Table 7.

Because of this new assignment procedure, there is a market value region in which both Russell 1000 and Russell 2000 stocks are present. Panel (c) of Figure 4 plots the index

[23]Specifically, it is a 5% band around the cumulated market cap of the stock ranked 1000 in Russell 3000E universe on the rank date.

[24]Russell's analysis is available online: https://www.ftserussell.com/blogs/russell-2000-recon-banding-results-lower-turnover.

weights around the cutoffs on the rank day (May 31^{st}) in 2012. In that year, the band is between ranks 823 and 1243. The discontinuity is still apparent: Russell 2000 stocks (in grey) have higher index weights. BMI mirrors the new pattern of the weights: the plot for Russell 2000 stocks lies above that for the Russell 1000 (Panel (d) of Figure 4).

In contrast with the literature, which typically accounts only for the Russell 1000 (blend) and Russell 2000 (blend), we will consider all nine indices on both sides of the cutoff. These indices include the Russell 1000 (blend, value, and growth) and Russell MidCap (blend, value, and growth) to the left of the cutoff and the Russell 2000 (blend, value, and growth) to the right of it.[25] Style funds (i.e., value and growth) have historically had a larger market share on the Russell 1000 side of the cutoff, while blend funds have been more important on the Russell 2000 side. Moreover, we include funds benchmarked to the Russell MidCap – an index which spans stocks smaller than rank 200 within the Russell 1000. It assigns a higher weight to the stocks near the cutoff than the Russell 1000 index because it excludes its 200 largest constituents. Market share of funds benchmarked to the Russell MidCap in our sample is almost as high as that of all Russell 2000 funds (Figure 1 in the Introduction).

3.3.2 Identification approach

Our main goal is to test the relationship between benchmarking intensity (BMI) and long-run returns, identified by our theory. For that, we need exogenous variation in BMI. The Russell cutoff provides a convenient setup. Given the random assignment of stocks around the cutoff on the rank day in May, index membership in June is a valid instrument for BMI. Stocks in the neighborhood of the cutoff share similar properties as an idiosyncratic market value shock on the rank day can put them to one or the other side of the cutoff.

Were we able to observe the exact running variable (the total market value rank used by Russell), we would be able to deploy a sharp regression discontinuity design (RDD). Even though the measure we received from Russell replicates index assignment very closely, it is not error-free. We show this formally with Table 8 in the Appendix. Imperfect fit indicates that we do not have a sharp rule even though the resulting prediction significantly outperforms those in the existing studies.[26]

Hence, a fuzzy RDD fits our setup better as it allows for noise in the treatment status around the cutoff and we can assume conditional exogeneity of the instrument for BMI. In a fuzzy RDD setup, it is common practice to use 2SLS to estimate the treatment effect (Lee and Lemieux (2010) and Hahn, Todd, and Klaauw (2001)). In the first stage, the treatment

[25] This set does not include Russell indexes that do not exhibit discontinuity in weights near the 1000/2000 cutoff. These are, for example, Russell 3000, Russell 2500, and Russell Small Cap Completeness.

[26] To be precise, 34 (4) additions and 16 (10) deletions are misclassified across years before 2007 (after). These numbers are before we apply filters.

dummy (Russell membership) is regressed on the predicted treatment and running variable (rank in May) controls. In the second stage, the outcome variable of interest is regressed on the predicted treatment, with the same controls. Since our goal is to instrument BMI, not just index membership, we add an intermediate stage, in which we regress BMI on the predicted treatment. In the last stage, we will hence use the predicted BMI to estimate the coefficient of interest. We perform our analysis for additions and deletions samples separately. If the usual IV assumptions are satisfied, it allows us to estimate the causal effect of BMI on long-run returns.[27]

We use the following stock-level three-stage specification to estimate β_0:

$$D_{it}^{R2000} = \alpha_0 \tau_{it} + \sum_n \alpha_n RV_{it}^n + \delta_0' \bar{X}_{it} + \varepsilon_{0t} \tag{5}$$

$$BMI_{it} = \gamma_0 \hat{D}_{it}^{R2000} + \sum_n \gamma_n RV_{it}^n + \delta_1' \bar{X}_{it} + \varepsilon_{1t} \tag{6}$$

$$Y_{i\,t+h} = \beta_0 \widehat{BMI}_{it} + \sum_n \beta_n RV_{it}^n + \delta_2' \bar{X}_{it} + \varepsilon_{2t} \tag{7}$$

In the above specification, τ_{it} is 1 when stock i is on the correct side of the cutoff on the rank day in May of year t, D^{R2000} is 1 when stock i is in the Russell 2000 on the reconstitution day in June of year t. $Y_{i\,t+h}$ is an average long-run return of stock i from September of year t over the investment horizon h. Specifically, we consider the 12-, 24-, 36-, 48-, and 60-month excess returns, which are not risk-adjusted. We also consider periodic returns, i.e., the average of 0-12, 12-24, 36-48, and 48-60 months. Variables with hats, $\hat{D}_{it}^{R2000}$ and $\widehat{BMI}_{it}$, are the fitted values from the preceding stage, (5) and (6), respectively. RV is the logarithm of total market value, i.e., the running variable as of May provided by Russell. The main specification will only include RV but following the practice in the literature, we explore robustness to polynomials of RV up to order n. $\bar{X}$ is a vector of other controls that include: 5-year monthly rolling β^{CAPM} computed using the CRSP total market value-weighted index, Russell float factor (proprietary liquidity measure affecting index weight), 1-year monthly rolling average bid-ask percentage spread, and stock return over year $t-1$.[28]

[27] In unreported analysis, we compare our results to that of a 2SLS procedure, which excludes the prediction step for the index dummy. The results are almost identical, which supports the identification strategy in Appel, Gormley, and Keim (2016) and Appel, Gormley, and Keim (2019b).

[28] Typically, a valid RD does not require covariate adjustment (Lee and Lemieux (2010)). However, we choose a wider band of 300 stocks for our baseline specification and report results with and without covariates for consistency with the model. Our theoretical prediction for the expected return (4) highlights that stocks may have different fundamental exposure through β_i – so we add β^{CAPM}. We include the float factor and bid-ask spread to address the liquidity hypothesis for index effect. Past return is included because we see that it is imbalanced for the treated and control samples in the covariate tests, currently unreported, and it may affect long-run returns through momentum. We document covariate-free estimation results for the reduced form specification in the Appendix.

We perform estimation separately for additions and deletions. That is, we first estimate specifications (5)-(7) only for stocks that belonged to the Russell 1000 in the previous year. In this case, D^{R2000} distinguishes stocks that got added to the Russell 2000 (treated) from stocks that stayed in the Russell 1000 (control). Similarly, we run the test for the sample of Russell 2000 stocks only and compare stocks that stayed in the Russell 2000 (treated) with stocks that got moved to the Russell 1000 (control). This is consistent with Chang, Hong, and Liskovich (2014).

Our dependent variable spans horizons from 12 months to 5 years. There is some ambiguity about what the long run is in the literature. The IPO performance literature (following Ritter (1991)) typically defines it as three years. The long-run reversal literature (started by De Bondt and Thaler (1985)) uses horizons from 18 months to five years. In our case, an additional problem is posed by flippers, i.e., stocks that switch from one benchmark to the other during the horizon that we are considering. Our model requires the stock's BMI to remain largely unchanged for the expected return result to play out as predicted. We comment on flippers further in Section 3.4.3.

We use a local linear regression appoach, i.e., our samples are restricted to the neighborhood of the cutoff (rectangular kernel).[29] Our default bandwidth is 300 stocks around the cutoff and we discuss robustness with respect to this choice variable in Section 3.4.2.

For the period up to 2006, the cutoff rank around which we center the analysis is 1000. For each year starting from 2007, we compute the left and right cutoffs based on the Russell methodology. Market value levels for the cutoffs we compute are reported in Table 7 in the Appendix, we almost fully match historical values reported by Russell.[30]

3.3.3 Instrument Strength

In this section we show that both first and second stage instruments in our specification are strong and relevant.

As was already discussed in the previous section, index membership is well-instrumented by the assignment prediction dummy τ. Table 9 in the Appendix explicitly documents the estimation results for specification (5).

Predicted index membership is a valid instrument for benchmarking intensity. Results of the second stage regression of the benchmarking intensity on Russell 2000 membership and controls are presented in Table 2. Russell 2000 membership is associated with a considerable increase in the benchmarking intensity: the estimates range between 0.017 and 0.061. This

[29] In an unreported analysis, we experiment with triangular kernels and get similar results.

[30] Published on the website: https://www.ftserussell.com/research-insights/russell-reconstitution/market-capitalization-ranges.

represents a large change as the sample standard deviation of BMI within each index around the cutoff is 0.05 (Table 10 in the Appendix). High F-statistics support the relevance of our instrument.[31]

Table 2: Second stage regression results

	Benchmarking intensity							
	1998-2006 sample		*2007-2018 sample*		*1998-2006 sample*		*2007-2018 sample*	
	Additions	*Deletions*	*Additions*	*Deletions*	*Additions*	*Deletions*	*Additions*	*Deletions*
$\hat{D}_{it}^{R2000}$	0.033***	0.027***	0.017***	0.030***	0.035***	0.037***	0.037***	0.061***
	(12.58)	(12.78)	(7.19)	(12.19)	(9.18)	(11.05)	(9.93)	(11.61)
Band width	300				100			
Running variable ($logMV$)	Yes				Yes			
Other controls ($\bar{X}$)	Yes				No			
Observations	2035	2438	1230	1955	649	947	354	635
F-statistic	244	194	110	182	51	62	49	72
Adjusted R^2, %	43	46	47	54	13	15	28	22

This table reports the results of the second stage regression (6) for stocks in the pre-banding (1998-2006) and the post-banding (2007-2018) samples. The dependent variable is the normalized benchmarking intensity of stock i as of September in year t, BMI_{it}. The key independent variable, $\hat{D}_{it}^{R2000}$, is the Russell 2000 index membership dummy predicted in the first stage. We include only stocks that were in the Russell 1000 (additions) or Russell 2000 (deletions) in the previous year. Both discontinuities are positive as the treated are the firms that remained in the Russell 2000. Band width is 300 or 100 stocks around the relevant cutoffs (rectangular kernel). Other controls ($\bar{X}$) include a float factor control, a 5-year monthly rolling stock beta computed using the CRSP total market value-weighted index, a 1-year monthly rolling average bid-ask percentage spread, and stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are marked as: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$.

An interesting insight from the second stage regression is the asymmetry of the cutoffs after 2007. Looking at the specification with a narrower band, prior to the introduction of banding, Russell 2000 membership explained around 0.035 standard deviations difference in the intensity between stocks on different sides of the cutoff. After 2007, this number increased to 0.06 around the left cutoff (for deletions) and remained the same around the right cutoff (for additions). This observation mirrors the relative distance between red and grey lines in Panel (d) of Figure 4: the distance is larger for the left cutoff, i.e., for the stock ranked around 825, as opposed to than the right cutoff, i.e., the stock ranked around 1250.

3.3.4 BMI and Long-Run Returns

We now present the results of the third stage regression. This stage tests whether a higher benchmarking intensity leads to lower returns in the long run. Specifically, we use $\widehat{BMI}_{it}$, the benchmarking intensity instrumented with stock index membership, to show that stocks with a higher intensity in year t significantly underperform up to year $t+5$.

[31] Based on critical levels in Stock and Yogo (2002).

Results of estimating the third stage regression in the full sample period (1998-2018) are documented in Table 3. As the coefficient on BMI is significantly negative, stocks with higher benchmarking intensities have lower returns in the future. The effect persists for up to 5 years into the future for additions to the Russell 2000 and 2 years for deletions from it.[32]

The magnitude of this effect is economically significant. In order to interpret the magnitude for an average added or deleted stock in our sample, we need to take into account the second stage coefficient or refer to the reduced form regressions. The reduced form regression results are included in Tables 16 and 17 in the Appendix. In the 1998-2006 sample period, addition to the Russell 2000 results in between 60bps and 150bps lower return per month[33] while deletion from it leads to a 40bps-140bps higher return per month. After 2007, the magnitudes decrease: addition to the Russell 2000 results in between 70bps and 100bps lower return while deletion leads to a 50-70bps higher return.

Consistent with the model, this analysis shows that an increase in the size of the preferred habitat has a long-lasting[34] effect on stock returns. In other words, inelastic demand from the benchmarked institutions does indeed lower the expected returns of the stocks. This result can also be interpreted as a negative long-term return of a long-short portfolio that buys stocks with high BMI and sells stocks with low BMI.[35]

It is striking that despite the predicted change in BMI being higher after 2007, the effect on returns is lower and even insignificant in some specifications. Before the introduction of banding, the effect is symmetric and strong for up to 5 years following index reconstitutions.[36] After 2007, the effect for deletions is smaller in magnitude and the effect for additions lasts for 24 months only. After 2013, the effect is short-lived in both samples and is weaker for deletions.

We identify several reasons why we see a weaker relationship between BMI and long-

[32] Even though it might seem from Panel A that most of the effect is concentrated in the first 12 months after index reconstitution, the negative relationship is long-term. To confirm this, we report Panel B in Table 3, which uses average returns over a future period as the dependent variable. It shows that the returns are lowest in the 0-12 months period, and they are significantly lower for the periods between 12 and 24, 24 and 36 as well as 36 and 48 months. Second, we show in the Appendix that the effect is almost evenly negative for the full five-year horizon in the 1998-2006 sample. We explain why the introduction of banding by Russell from 2007 onwards weakens cutoffs-based tests in Section 4.4.

[33] As discussed earlier, the magnitudes depend on whether we include or exclude flippers so we report a range. Moreover, note that all our results are relative to the control group.

[34] Permanent, as long as the stock stays in the benchmark.

[35] To our knowledge, while the literature has argued that the index effect lasts for over two weeks/months, no one has documented a long-run (up to 5 years) effect of index inclusion on stock returns. This is probably because this effect is hard to tease out by studying index (most commonly, the S&P 500) inclusions and using the market portfolio as a control group. In our quasi-experiment, the control group consists of stocks around the Russell cutoff, which are more similar to treated stocks.

[36] Results for sub-samples are in Table 11 in the Appendix.

Table 3: Benchmarking intensity and long-run returns in 1998-2018

Panel A: Excess returns, average over horizon

	Additions to Russell 2000					*Deletions from Russell 2000*				
Horizon (months)	12	24	36	48	60	12	24	36	48	60
Panel A1: All controls (baseline)										
$\widehat{BMI}_{it}$	**-0.64*****	**-0.37*****	**-0.23*****	**-0.19*****	**-0.11*****	**-0.10*****	**-0.06*****	-0.02	-0.00	-0.01
	(-4.49)	(-4.43)	(-3.63)	(-3.72)	(-2.64)	(-3.21)	(-2.77)	(-1.07)	(-0.28)	(-0.50)
Observations	3096	2842	2602	2380	2156	4149	3750	3401	3094	2836
Panel A2: Only the running variable, *logMV*										
$\widehat{BMI}_{it}$	**-0.80*****	**-0.51*****	**-0.30*****	**-0.22*****	**-0.12*****	**-0.08*****	**-0.04****	-0.00	0.00	0.00
	(-4.02)	(-4.04)	(-3.75)	(-3.73)	(-3.02)	(-2.66)	(-1.68)	(-0.18)	(0.27)	(0.14)
Observations	3294	3021	2756	2509	2268	4415	3988	3616	3286	3002

Panel B: Excess returns, average in the period (months)

	Additions to Russell 2000					*Deletions from Russell 2000*				
Period (months)	0-12	12-24	24-36	36-48	48-60	0-12	12-24	24-36	36-48	48-60
Panel B1: All controls (baseline)										
$\widehat{BMI}_{it}$	**-0.63*****	**-0.49*****	**-0.41*****	**-0.31*****	-0.03	**-0.09*****	-0.02	0.08	0.05	-0.00
	(-4.37)	(-3.91)	(-3.51)	(-2.70)	(-0.30)	(-2.90)	(-0.51)	(2.53)	(1.41)	(-0.13)
Observations	3096	2842	2601	2383	2155	4149	3750	3401	3096	2836
Panel B2: Only the running variable, *logMV*										
$\widehat{BMI}_{it}$	**-0.80*****	**-0.70*****	**-0.49*****	**-0.35*****	-0.04	**-0.05***	0.02	0.07	0.05	0.00
	(-3.87)	(-3.73)	(-3.58)	(-2.69)	(-0.36)	(-1.51)	(0.66)	(2.06)	(1.39)	(0.10)
Observations	3294	3021	2509	2756	2268	4415	3988	3616	3286	3002

This table reports the results of the third stage regression (7) for the full sample period. We include only stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. We limit the sample to 300 stocks around the cut-offs (rectangular kernel). The dependent variable in Panel A is an average monthly excess return from September in year t over the respective horizon. The dependent variable in Panel B is an average monthly return in the respective period, e.g., 12-24 months after reconstitution. Panels A1 and B1 use all baseline controls while Panels A2 and B2 only include the running varriable, $logMV$. The baseline controls include log total market value, a float factor control, a 5-year monthly rolling β^{CAPM}, a 1-year monthly rolling average percentage bid-ask spread, and stock return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters by stock are in parentheses. Significance levels are based on a one-sided test and marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

run returns in the latter part of the sample. Firstly, with the introduction of banding, incentives to align closely with the benchmark changed for funds holding stocks around the cutoff: they became stronger for the left cutoff and weaker for the right. As we show in Section 4 below, funds benchmarked to the Russell 1000 and Russell MidCap are more likely to rebalance additions to their benchmarks. In the same section, we explain why these observations are consistent with the industry-wide practice of optimized sampling. Secondly, a range of Vanguard passive funds has switched to CRSP indexes in 2012. These indices have a cutoff which overlaps with the Russell cutoff. We discuss both in detail in Section 4.

3.3.5 BMI and Index Effect

In this section, we show that a higher benchmarking intensity change leads to the larger price pressure (short-term return) upon an index event. This corresponds to Prediction 2 of our model. It is a natural result since we use predicted index membership as an instrument for BMI. This is not a new result. It has been documented for the Russell 1000/2000 index cutoff in Chang, Hong, and Liskovich (2014). But our suggestion that the size of the index effect can be linked to a stock's BMI and our explanation for why the result is weaker in the latter part of their sample are novel.

For consistency with Section 3.3.2, we estimate the following specification:

$$\begin{aligned} \Delta BMI_{it} &= \gamma_0 \tau_{it} + \sum_n \gamma_n RV_{it}^n + \delta_1' \bar{X}_{it} + \varepsilon_{1t} \\ Ret_{it}^{June} &= \beta_0 \widehat{\Delta BMI}_{it} + \sum_n \beta_n RV_{it}^n + \delta_2' \bar{X}_{it} + \varepsilon_{2t} \end{aligned} \tag{8}$$

In the above specification, τ_{it} is 1 when stock i is on the correct side of the cutoff on the rank day in May of year t. Ret_{it}^{June} is the excess return of stock i in June of year t (as in Chang, Hong, and Liskovich), winsorized at 1%. ΔBMI_{it} is a difference between the BMI of stock i in May of year t and its deflated BMI in June of the same year. Deflated BMI is computed using index AUM shares in June but weights as of May; that is, it accounts for the new index membership of stock i but not its return in June. We deflate BMI because otherwise the actual June index weights will include the (post-announcement) price pressure and exhibit a positive relationship with June returns.[37] RV is the logarithm of total market value, the running variable as of May provided by Russell. $\bar{X}$ is a vector of other controls that we include for consistency, they are the same as in our long-run analysis.[38] We estimate

[37] Results are robust to alternative, shorter, deflators and to using May's AUM shares.

[38] 5-year monthly rolling β^{CAPM} computed using the CRSP total market value-weighted index, Russell float factor (proprietary liquidity measure affecting index weight), 1-year monthly rolling average bid-ask percentage spread, stock's return over year $t-1$.

this specification for additions and deletions separately.

Table 4: Change in BMI and price pressure in June

	Excess return in June			
	Additions	*Deletions*	*Additions*	*Deletions*
$\widehat{\Delta BMI}_{it}$	0.080***	0.015**	0.045**	0.024**
	(5.89)	(2.20)	(2.52)	(2.76)
Band width	300		100	
Running variable ($logMV$)	Yes		Yes	
Other controls, $\bar{X}$	Yes		No	
Observations	3632	4915	1089	1677
Adjusted R^2, %	17	16	2	2
First-stage coefficient	0.39***	0.59***	0.41***	0.60***
	(31.09)	(34.67)	(17.10)	(21.38)
First-stage F-statistic	521	652	292	457
First-stage R^2, %	26	38	36	37

*This table reports the results of specification (8) for stocks in the full sample (1998-2018). The dependent variable is the winsorized return of stock i in June in year t. The key independent variable, $\widehat{\Delta BMI}_{it}$, is the predicted change in BMI between June and May deflated to May prices. Other controls $\bar{X}$ are our baseline controls from Table 3. We include only stocks that were in Russell 1000 (additions) or Russell 2000 (deletions) in the previous year. Band width is 300 or 100 stocks around the cutoffs (rectangular kernel). t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are marked as: *p<0.10; **p<0.05; ***p<0.01.*

Estimation results for both stages are presented in Table 4. We see a significant and positive relationship between the BMI change and price pressure upon reconstitution. Consistent with our model's Prediction 2, price pressure will be highest for stocks experiencing the largest increase in BMI, all else equal. Therefore, in contrast with the existing literature which looks at the average index effect, our analysis suggests that the size of index effect can be linked to the stock's BMI.

This result is in line with the findings of Chang, Hong, and Liskovich. Both additions and deletions experience price pressure upon an index event. Moreover, our price pressure estimates based on index dummy instead of ΔBMI lie within the ranges documented in the Internet Appendix to their paper.

3.3.6 Remarks on Exclusion Restriction

Recent literature has similarly exploited the Russell 1000/2000 cutoff to document a number of corporate implications of institutional ownership (e.g., Appel, Gormley, and Keim (2019b) and references therein), some in conflict with each other. In this subsection, we discuss whether the findings of this literature may provide an alternative explanation for our findings and/or challenge our identification strategy. We comment on the exclusion restriction, ownership discontinuities, and the direction of long-run returns.

Table 5 lists the most common empirical approaches to exploit the Russell cutoff used in the literature. This table highlights that the pioneering work of Chang, Hong, and Liskovich (2014), sparked active research in this area, with many papers exploiting the cutoff to answer a variety of questions, primarily in corporate finance. The biggest challenge to identification in this body of work is that the true ranking variable is not available to researchers,[39] which led some researchers to use fuzzy RDD and some to question the conditional exogeneity assumption when a simple IV approach is used (Wei and Young (2017)). One point the literature now agrees on is that that the June weights cannot be used for assignment. Some of the early papers used the June weights and it is not known whether the results are robust to using end-of-May weights instead.

Table 5: A summary of empirical methods exploiting the Russell 1000/2000 cutoff

Methodology	**Sample**	**Instrument**	**Example**
Fuzzy RDD	1996-2012	Index dummy	Chang, Hong, and Liskovich (2014)
Fuzzy RDD with IV (3SLS)	1998-2006	ETF ownership	Ben-David, Franzoni, and Moussawi (2018)
IV approach with logMV	1998-2006	Passive IO	Appel, Gormley, and Keim (2016)
IV with logMV and band controls	2008-2014	Benchmarked passive IO	Appel, Gormley, and Keim (2019b)
IV with ranks	1991-2006	Total IO	Crane, Michenaud, and Weston (2016)
IV for additions/deletions	1993-2010	Change in the passive IO	Schmidt and Fahlenbrach (2017)
Cohort difference-in-differences	2004-2017		Heath, Macciocchi, Michaely, and Ringgenberg (2018)

This table shows empirical approaches most frequently used to exploit the cutoff between the Russell 1000 and Russell 2000 indices. RDD stands for regression discontinuity design, IO – institutional ownership, 3SLS - three-stage least squares, IV – instrumental variable, logMV – log market value (running variable). The layout is borrowed from Glossner (2018).

Our analysis poses no conflict with the documented evidence. First, a discontinuity in BMI is not conflicting with any of the documented discontinuities. That is, discontinuities in total institutional ownership (IO), passive IO, benchmarked IO, and ETF ownership are implicitly assumed in our measure. They are also assumed to be time-varying since the amount of capital linked to indices varies and new indices emerge. Therefore, if BMI is a comprehensive measure, we *expect* to see discontinuities in all aforementioned variables: whether the discontinuity is identified in a particular sample depends on the distribution of assets between benchmarks. Hence, it does not imply a violation of the exclusion restriction in our analysis. Second, we address the debate on identification in the literature in the following way. We use a proprietary running variable provided by Russell which minimizes the concern of the violation of conditional exogeneity of the Russell index dummy.[40] Moreover, we use

[39] Most papers reconstruct it using CRSP data. Ben-David, Franzoni, and Moussawi (2018) propose a new method to construct the variable for a higher assignment accuracy from public data.

[40] As we discussed above, the assignment prediction quality is very high.

a prediction step to orthogonalize any remaining measurement error (similar to Ben-David, Franzoni, and Moussawi (2018)).

We see no threat to our results coming from the findings of other papers either. First, the majority of documented results are short-term: they are measured in the year following the reconstitution, while our main focus is on long-term returns.[41] Second, and more importantly, most of the findings document improvements to future cash flows for stocks to the right of the cutoff, which would imply that future realized returns should be larger for those stocks, while we find the opposite.

In Table 15 in the Appendix, we perform reduced-form tests for the fundamental variables associated with cash flows average over the three years following the reconstitution.[42] We generally find little evidence that any of them is significantly different for the treated and control groups. However, one variable that stands out is stock repurchases (before 2007); they increase for additions to the Russell 2000. Repurchases are a part of the payout to shareholders. Prior research has argued that firms that enter the Russell 2000 are better monitored than firms in the Russell 1000 (Crane, Michenaud, and Weston (2016)), which leads to increased cash flows to the shareholders and, therefore, higher realized returns. In contrast, we find that the realized returns are lower. Hence, the significant effect on repurchases cannot explain our findings.

In an unreported analysis, we check if the risk factor loadings change with Russell index membership. We find no robust changes in either Fama-French-Carhart, Fama-French 5-factor, or Pástor and Stambaugh (2003) loadings.[43] In some specifications, SMB loading increases upon addition to the Russell 2000. Since this change should be associated with higher long-run returns, it could only prevent us from finding the result.

We discuss further alternative explanations of our results in the robustness section (Section 3.4.5).

[41] For example, Schmidt and Fahlenbrach (2017) focus on acquisitions in one year after an index switch. Since we find a longer-term effect, our results are at best complementary.

[42] Apart from the reported characteristics, we inspected all characteristics summarized in Lewellen (2015). We find some support for lower asset growth of stocks added to (or not deleted from) the Russell 2000 but that would only bias us against finding our main result. Furthermore, we observe higher turnover of these stocks so we link it to rebalancing of benchmarked funds (see Section 4). However, stock liquidity, as measured by ILLIQ of Amihud (2002) and bid-ask spread, deteriorates. In the litrature, this is associated with higher expected returns, which is opposite to what we find.

[43] This analysis involves estimating our reduced form specification with the future loadings on the left-hand side and controlling for lagged (pre-event) loadings. All loadings are 5-year computed from monthly rolling regressions of stock excess returns on factor returns available from Ken French's website or WRDS, with a minimum of 2 years of data required for estimation.

3.4 Robustness

3.4.1 Time Fixed Effects and Double-Clustering

Our baseline specification does not include year fixed effects, even though our goal is to compare the returns in the cross-section. We exclude them mainly because our analysis separates additions and deletions samples and there is an insufficient number of treated firms by year within them (as shown in Appendix, Table 7).

At the same time, for the year fixed effects to play a role in our experiment, there has to be a differential impact on the treated and control groups. That is, since we are comparing firms that switched indices with the firms that stayed, the estimate will be biased without year fixed effects only if the treated group is affected in a systematically different way than the control group. We discuss below that the covariates are balanced, which indicates that this concern is not very strong in our sample.[44]

For illustration only, we widen the band to report the results of estimating a reduced form specification with year fixed effects in Table 16 in the Appendix. The significance is weaker but the sign is negative for all horizons, in line with our main specification. The reduced form for our specification is (7) with τ_{it} (predicted assignment) used instead of $\widehat{BMI}_{it}$.

The reported t-statistics are based on standard errors clustered by stock. Double-clustering does increase the standard errors but most of the results remain significant at 5%. Since the prediction of the model is that the sign of β_0 is negative, our inference is based on a one-sided test with $H_a : \beta_0 < 0$. Of course, the one-sided tests do not change the reported t-statistics.

3.4.2 Bandwidth Selection

Our default bandwidth is 300 stocks around the cutoff. We also report the main estimation results with the bandwidth of 100 in Table 12 in the Appendix.

We check the MSE-optimal bandwidths as well (first suggested in Imbens and Kalyanaraman (2011)). The symmetric (asymmetric) MSE-optimal bandwidths are typically slightly below (above) the width of 300 which we use in the main results. They also vary slightly by return horizon. With these bands, the statistical significance of β_0 is weaker but the sign remains consistent with the reported results.

[44]Nevertheless, because our theory suggests the role for differential fundamental loadings (β_i in (2.1)), we report the estimation results for CAPM abnormal long-run returns in the Appendix (Table 13). It alleviates the concern that our results are driven by differential market exposure of the control and treated groups.

3.4.3 Flippers

We considered excluding stocks that moved between the Russell 1000 and 2000 indexes more than once in five years – the so-called 'flippers'. Our theoretical predictions concern stocks that joined a set of indices and stayed in them until the end of the investment horizon. Our results are considerably stronger, both statistically and in magnitude, if we drop stocks that moved between the Russell 1000 and 2000 indexes more than once in the relevant horizon.[45] Economically, our theory would predict that the BMI of such flippers would change when they move index again and hence distort the long-run returns upwards. Moreover, fund managers, especially those who can take on more tracking error risk, could be able to identify stocks that move back and avoid trading them in the first index reconstitution to avoid transaction costs.

From the econometrics perspective, however, excluding flippers introduces a selection bias. A stock which was added to the Russell 2000 index has to appreciate in value to come back to the Russell 1000 the next year. Therefore, by excluding flippers, we naturally exclude stocks with the most positive return realizations, which biases our β_0 estimate downward.

We believe that an analysis of any long-run variable using the Russell cutoff has to weigh these potential biases. At least, it has to take into account future index membership changes. In our case, the main reported results do not exclude flippers from the sample and hence represent the upper bound on β_0 coefficient.

We also use other filters. Consistent with the literature (e.g., Schmidt and Fahlenbrach (2017)), we exclude stocks that move more than 500 ranks in one year. Our results are robust to this filter but we prefer to keep it in place to ensure the comparability of stocks.

Since we study long-run returns, stocks that leave the sample within a certain horizon will be dropped from the respective regression[46]. All returns we use are adjusted for delisting.

3.4.4 Covariate Balance

In this section, we show that the other observed characteristics are smooth for firms around the cutoff. That is, we test for the differences in fundamental firm characteristics determined prior to the Russell reconstitution. We do it by estimating specification (7) with index dummy D^{R2000} instead of BMI (baseline controls are included) separately for additions and deletions samples. We ensure that the data we use is released to the public by the rank day in May.

Results are in Table 14 in the Appendix. None of the imbalances is robustly sig-

[45] Reduced form regression results are reported in Tables 16 and 17.

[46] Results remain qualitatively unchanged if we keep the available returns of these stocks (incomplete year).

nificant. Moreover, we cannot think of an economic story why, say, repurchases should be significant for additions but not for deletions and enter with different signs before and after 2007. Nonetheless, in unreported analyses, we control for each of the imbalances and find no change to our results.

Apart from the reported characteristics, we considered additional characteristics shown to predict returns for U.S. stocks in the cross-section (summarized in Lewellen (2015)) as well as factor loadings (CAPM, Fama-French-Carhart, Fama-French 5, Pastor-Stambaugh, and standalone benchmark betas, e.g. with respect to the Russell 1000) and liquidity measures (ILLIQ of Amihud (2002), Return-to-Turnover of Florackis, Gregoriou, and Kostakis (2011), and effective spread of Abdi and Ranaldo (2017)). The only measures that appear imbalanced are the bid-ask spread, stock volatility, past year stock return, and CAPM beta. We, therefore, include them as controls in our baseline specification, dropping volatility as it turns out to be insignificant and have no effect on our estimates.

3.4.5 Alternative Explanations

One of the alternative explanations for our results is that returns of firms that have transitioned to the Russell 2000 are lower because these firms have fallen on hard times and their cash flows are deteriorating. If this momentum continues, it is not surprising to see that the firms added to the Russell 2000 have lower future returns relative to firms that stayed in the Russell 1000. Our baseline controls (specifically, past returns) and the reported covariate tests are designed to alleviate this concern. Nonetheless, we took further steps to ensure this explanation is ruled out.

In addition to the covariate imbalance tests, we have checked explicitly whether any of the firms moving to the Russell 2000 are in financial distress. In our dataset, treated and control firms have similar Altman Z-scores and the scores do not change upon index reconstitution. Moreover, excluding firms classified by Altman Z-score as being ‘in distress’ or ‘in the grey zone’ does not change either the significance or magnitude of our results. We have also experimented with excluding firms that had a rapid deterioration in their market value rank prior to reconstitution. While our baseline analysis excludes jumps of 500 ranks, we have tried excluding firms that lost even as little as 100 ranks. Our results remained qualitatively unchanged, albeit the magnitude of the effect was smaller.

4 Benchmarking Intensity and Trends in Institutional Ownership

Starting from Gompers and Metrick (2001), empirical literature documented a range of effects of institutional trading and ownership for stock prices. A recent strand of literature looks into the effects of ownership on corporate outcomes. There has been no research, however, on the benchmarking-induced ownership.

Benchmarking intensity reflects the incentives elicited by the contracts of asset managers, both active and passive. In this section, we show that both investor types have a considerable fraction of holdings concentrated in their benchmarks and that they rebalance stocks relevant for *their* benchmarks around the Russell cutoffs. That is, we document a heterogeneity of investor habitat dictated by their benchmarks.

We also show that the change in Russell's reconstitution methodology in 2007 (i.e., the introduction of banding) has altered funds' incentives to rebalance. It mostly affected the buying of deletions from the Russell 2000. In the light of this change, we discuss how portfolio construction based on optimized sampling trades off benchmarking incentives with transaction costs.

Finally, we describe other index groups, CRSP and S&P, and how their reconstitutions may affect studies on the Russell cutoff.

4.1 Benchmarks as Funds' Habitat

As Robert Stambaugh points out in his AFA Presidential Address (Stambaugh (2014)), U.S. mutual funds' tracking errors have been going down. In our dataset, this trend is drastic. A simple average tracking error of active funds went down from 7% per annum in the early 2000s to below 4% in 2010s. For passive funds, these numbers have been below 2% and closer to 0.5%, respectively. Given that the share of passive funds grew significantly over the past two decades,[47] the overall industry tracking error is at its historical low.

Exploiting the granularity of our dataset, we also compute the percentage of fund AUM invested in its benchmark stocks and the number of benchmark stocks held. Over our sample period (1998-2018), the AUM share in the benchmark stocks has risen from 75% to 82% for active funds. The number of benchmark stocks they hold has also risen from 60% to 80% of the total number of stocks in their portfolios. Both figures have consistently been close to 100% for passive funds.

[47] The assets of stock index mutual funds and ETFs now match that of active funds, according to: https://www.bloomberg.com/news/articles/2019-09-11/passive-u-s-equity-funds-eclipse-active-in-epic-industry-shift.

These trends suggest that benchmarks define funds' preferred habitats.[48] In the following section, we document that funds actually rebalance stocks added to or deleted from their benchmarks.

4.2 Net Purchases of Index Additions and Deletions

Earlier studies documented that Russell index funds and ETFs buy additions to and sell deletions from their benchmarks. We argue that this list is incomplete and that active managers engage in the same behavior but detecting it requires granular data on their benchmarks.

In order to see which funds rebalance additions and deletions, we estimate the following specification at a stock level:

$$D_{it}^{Index} = \alpha_0 \tau_{it}^{Index} + \sum_n \alpha_n RV_{it}^n + Own_{i,j,t-1} + \delta_0' \bar{X}_{it} + \epsilon_{1t}$$
$$Own_{i,j,t} = \beta_0 \hat{D}_{it}^{Index} + \sum_n \beta_n RV_{it}^n + Own_{i,j,t-1} + \delta_2' \bar{X}_{it} + \epsilon_{2t}$$

In the above specification, τ_{it}^{Index} is 1 when stock i is on the correct side of the cutoff on the rank day in May of year t for membership in the respective index, the Russell 1000 or Russell 2000. $Own_{i,j,t}$ is the percentage of outstanding shares of stock i owned by fund group j at the end of September of year t. The funds are grouped by benchmark and type (active/passive). We perform analysis on September holdings data because: (1) it allows for delayed rebalancing after June reconstitution[49], (2) it is based on quarterly holdings[50], and (3) it is in line with most of the previous studies (e.g., Appel, Gormley, and Keim (2016)). RV is the logarithm of total market value, the running variable as of the rank day in May provided by Russell. $\bar{X}$ is a vector of other controls from our long-run analysis that include: 5-year monthly rolling β^{CAPM} computed using CRSP VW index, Russell float factor (proprietary liquidity measure affecting index weight), 1-year monthly rolling average Bid-Ask percentage spread, and stock return over year $t-1$.[51]

As in our earlier analysis, we use a 2SLS estimator. It allows us to identify the effect of

[48] All our analysis is conditional on the benchmark in manager's contract. Our model does not take a stand on how end investors pick the benchmark or fund to invest in. Possible rational explanations include the need to hedge endowment shocks of a particular type or to hedge displacement risk. Behavioral explanations include psychological foundations for why investors prefer growth over value, over-reaction, and extrapolation of past returns.

[49] In undocumented analysis, we see that after 2007 a considerable fraction of rebalancing of additions and deletions happens in July.

[50] These holding records are more complete because their filing is mandatory on a quarterly basis.

[51] This specification does not include year and industry (SIC-1) fixed effects for the same reason of insufficient variation within additions and deletions samples. Results are similar if we include them.

addition to or deletion from an index and alleviate a concern that an omitted variable might be driving both membership in the index and the level of ownership of funds benchmarked to that index. We perform this analysis on additions and deletions separately, at an index level, and distinguish between active and passive funds benchmarked to that index. For example, we estimate a separate regression for the ownership share of the active Russell 1000 funds in stocks that were in the Russell 1000 on the rank day in May. In this example, the interpretation of β_0 on D^{R2000} is the change in their ownership share due to the stock's addition to the Russell 2000 index (and its deletion from the Russell 1000 index group – i.e., the Russell 1000 blend, Russell MidCap blend, and their value and growth counterparts.).

Table 6 documents that both passive and active funds rebalance additions and deletions. Consistent with the literature, we find highly significant stock ownership changes for passive funds in line with their benchmarks. For example, passive funds benchmarked to the Russell 2000 increase their ownership in stocks added to the Russell 2000 by 103bps. These funds also sell deleted stocks in similar proportions. At the same time, we see that active funds benchmarked to the Russell 2000 also sell deletions, decreasing their ownership share by 100bps.

Table 6 reveals that active funds engage in rebalancing additions and deletions even more after banding was introduced in 2007. Active funds benchmarked to the Russell 1000 and Russell MidCap increase their ownership shares in stocks deleted from the Russell 2000 by 42bps and 108bps respectively. They also sell additions to the Russell 2000 (23bps and 55bps, respectively). It is important to note the asymmetry in active funds' trading of additions and deletions after 2007. We will elaborate on this point when discussing optimized sampling in Section 4.4 below.

Table 6 includes CRSP benchmarks after 2013 as well. Those regressions feature dummies for CRSP US Large and CRSP US Small membership (not instrumented), $D_{it}^{CRSP-Large}$ and $D_{it}^{CRSP-Small}$, respectively. CRSP Large and Mid Cap funds (counterparts of the Russell 1000 and Russell MidCap) buy members of the CRSP Large and Mid Cap indices and sell members of the CRSP Small Cap indices. CRSP Small Cap funds do the opposite. The sheer size of these funds makes them large investors in the market cap region around the Russell cutoff. Furthermore, there is a potential conflict between the Russell and CRSP cutoffs, which we explain in Section 4.5 below.

Table 6: Rebalancing of additions and deletions, by benchmark and fund type

Summary of separate regressions on additions and deletions

	Ownership by Benchmark Group and Fund Type, %									
	Stocks ranked < 1000							Stocks ranked > 1000		
Benchmark	Russell 1000		Russell MidCap		S&P 500		CRSP LM	Russell 2000		CRSP S
Fund type	Active	Passive	Active	Passive	Active	Passive	Passive	Active	Passive	Passive
Panel A: Pre-banding sample (1998-2006)										
$\hat{D}_{it}^{R1000}$	0.10***	0.04***	0.13	0.02***	0.06	-0.01		-0.67***	-0.26***	
	(3.59)	(21.16)	(1.33)	(17.39)	(0.67)	(-1.54)		(-4.45)	(-12.05)	
$\hat{D}_{it}^{R2000}$	0.08	-0.04***	-0.16	-0.02***	-0.07	0.01		0.14	0.48***	
	(1.36)	(-14.76)	(-1.27)	(-14.16)	(-0.55)	(0.24)		(0.99)	(21.60)	
Number of funds (2006)	224	10	133	3	343	37		254	12	
Panel B: Post-banding sample (2007-2018)										
$\hat{D}_{it}^{R1000}$	0.41***	0.28***	1.08***	0.34***	0.65***	0.00	0.02	-1.54***	-2.48***	0.07
	(8.11)	(36.51)	(7.90)	(43.60)	(4.87)	(0.46)	(1.85)	(-9.06)	(-55.31)	(0.83)
$\hat{D}_{it}^{R2000}$	-0.23***	-0.27***	-0.55***	-0.33***	0.06	-0.18***	0.04**	-0.33	2.37***	-0.03
	(-3.00)	(-22.29)	(-3.29)	(-30.89)	(0.40)	(-2.87)	(3.03)	(-1.38)	(32.49)	(-0.31)
$D_{it}^{CRSP-Large}$							3.17***			-1.84***
							(114.33)			(-7.62)
$D_{it}^{CRSP-Small}$							-0.73***			1.69***
							(-3.87)			(12.36)
Number of funds (2013)	326	23	181	8	378	67	6	305	22	3
Panel C: Full sample (1998-2018)										
$\hat{D}_{it}^{R1000}$	0.23***	0.10***	0.68***	0.12***	0.38***	-0.01		-1.00***	-1.39***	
	(9.77)	(21.68)	(10.25)	(21.48)	(5.70)	(-1.23)		(-10.93)	(-38.49)	
$\hat{D}_{it}^{R2000}$	-0.03	-0.10***	-0.37***	-0.12***	-0.05	-0.03		0.01	1.03***	
	(-0.75)	(-16.19)	(-3.91)	(-16.69)	(-0.53)	(-1.18)		(0.10)	(21.90)	

This table reports the discontinuities in rebalancing for the pre- and post-banding sample periods. Estimation is performed at investor group level (by benchmark and fund type). The coefficients come from separate regressions: on stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. Band width is 500 stocks around the cutoffs (rectangular kernel). The dependent variables are ownership shares in stock i as of September in year t of the respective investor group. CRSP funds are only available from 2013. All regressions include one-year lagged ownership, year and industry fixed effects, polynomial of log total market value of order 1 and all other controls in $\bar{X}$. t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

4.3 Value and Growth Indices

Disaggregating investor groups by style (value or growth), we document additional discontinuities by benchmark. When a stock moves from the Russell 1000 to Russell 2000, it also enters the Russell 2000 Value and Growth indices.[52] In an analysis similar to the previous section, we show that active value funds rebalance value stocks and growth funds rebalance growth stocks.

In order to perform a discontinuity test as in earlier sections, we would need to control for variables that define assignment to value and growth indices. This assignment is not as easy for researchers to predict as that to market cap indices. Using a proprietary database of I/B/E/S forecasts, B/P, and sales growth, Russell runs a custom probability algorithm to define a share of stock's market cap as value or growth. Therefore, we cannot ensure the exogeneity of style dummies, e.g., $D^{R2000\,Value}$ and $D^{R2000\,Growth}$, and our results in this section should be viewed as suggestive.

As Table 18 in the Appendix reports, both active and passive funds rebalance in line with their benchmarks.[53] For example, passive Russell MidCap Growth funds buy additions to the Russell 1000 Growth universe and sell additions to Russell 2000 universe, and more so for additions to the Russell 2000 Growth universe.

Similarly to the previous section, active funds rebalance deletions after 2007 the most. For example, Russell 1000 Value funds buy additions to the Russell 1000 Value, Russell MidCap Growth funds – to the Russell 1000 Growth, while active Russell 2000 funds sell deletions.

Again, in the post-banding sample (after 2006), we detect fewer discontinuities for active funds around the right Russell 1000/2000 cutoff: they do not trade additions to the Russell 2000 at a similar scale. This emphasizes the asymmetry we documented above for the analysis at a market cap index level. In the following section, we offer a potential explanation for this phenomenon.

[52] Russell methodology is such that most of the stocks belong to both indices, i.e., some part of market value is assigned to value and some – to growth. In other words, a stock is rarely a pure value or growth. Russell has special indices for pure style stocks that are rather small in AUM.

[53] In these regressions, a coefficient on a style dummy should be summed with the coefficient on market cap dummy, e.g., a coefficient on $D^{R2000\,Value}$ should be summed with D^{R2000} to get a change in ownership for stocks that entered the Russell 2000 Value. Market cap dummy can be interpreted as the change for the pure style opposite the style dummy, or Growth in our example, because it shows the rebalancing when $D^{R2000\,Value} = 0$.

4.4 Optimized Sampling

In this section, we propose an explanation for the asymmetry in funds' net purchases of index additions and deletions following the introduction of banding in 2007. We have documented that funds benchmarked to the Russell 1000 and Russell MidCap are more likely to rebalance additions to their benchmarks than deletions and they seem to do so more robustly than before 2007. These observations are consistent with the industry-wide practice of optimized sampling.

Optimized sampling is a portfolio construction technique in which ex ante tracking error is balanced with expected transaction costs. In our model, the tracking error concern of the manager is driven by the relative performance component $R_j - B_j$ in her contract. The higher the relative performance sensitivity b, the lower the tracking error the fund (i.e., the manager demands more shares of the benchmark and fewer shares of the mean-variance portfolio). Our model abstracts from transaction costs, whereas in practice, transaction costs are an important consideration. Not buying an asset in the benchmark saves on transaction costs but increases the manager's tracking error relative to the benchmark. Optimized sampling addresses this trade-off.[54] Figure 8 in the Appendix illustrates how funds describe this portfolio construction approach in their prospectuses.

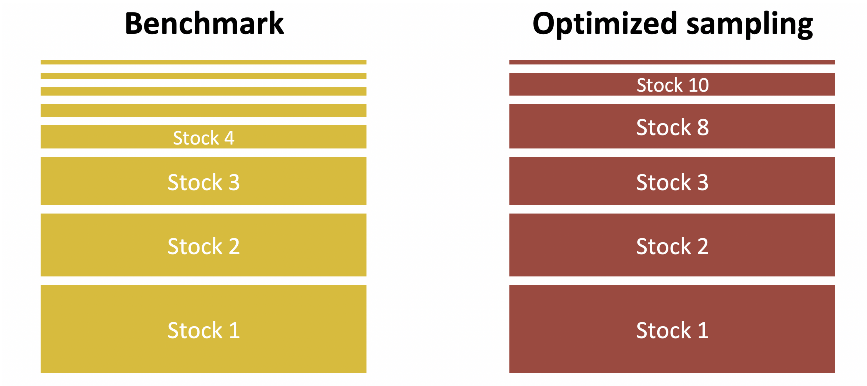
Optimized sampling directly interferes with the incentives to hold the benchmark portfolio. In the presence of transaction costs, funds no longer hold benchmark securities proportionally to benchmark weights. Rather, they typically hold the largest stocks with benchmark weights, completely omit the smallest and some mid-range stocks, and overweigh most of the mid-range stocks. An example of portfolio with benchmark weights and weights under optimized sampling is illustrated in Figure 5.

In an unreported numerical analysis, we modify the fund manager's optimization problem by introducing fixed transaction costs for trading each stock and adding a constraint that the fund's tracking error cannot exceed a realistic upper bound. Solving such a problem for the Russell 1000, Russell MidCap, and Russell 2000 yields portfolio weights that underweight the lowest-cap stocks in each index while overweighting the mid-cap stocks. This, in turn, changes how the weight discontinuities align with the cutoff after 2007.[55] The right pane of Figure 6 plots benchmarking intensity computed using such weights. There is essentially no discountinuity in BMI at the left cutoff, while the right cutoff continues to

[54] In practice, portfolio construction software typically allows additionally for further constraints like matching dividend yield of the benchmark, its B/M, industry exposures, etc.

[55] The effect of optimized sampling on the cutoff before 2007 is opposite: since the smallest stocks in either the Russell MidCap or Russell 1000 are not purchased due to higher transaction costs and the largest Russell 2000 stocks are purchased close to index weights, the discontinuity is larger than the one implied by benchmark weights.

Figure 5: Benchmark portfolio weights vs. optimized sampling weights

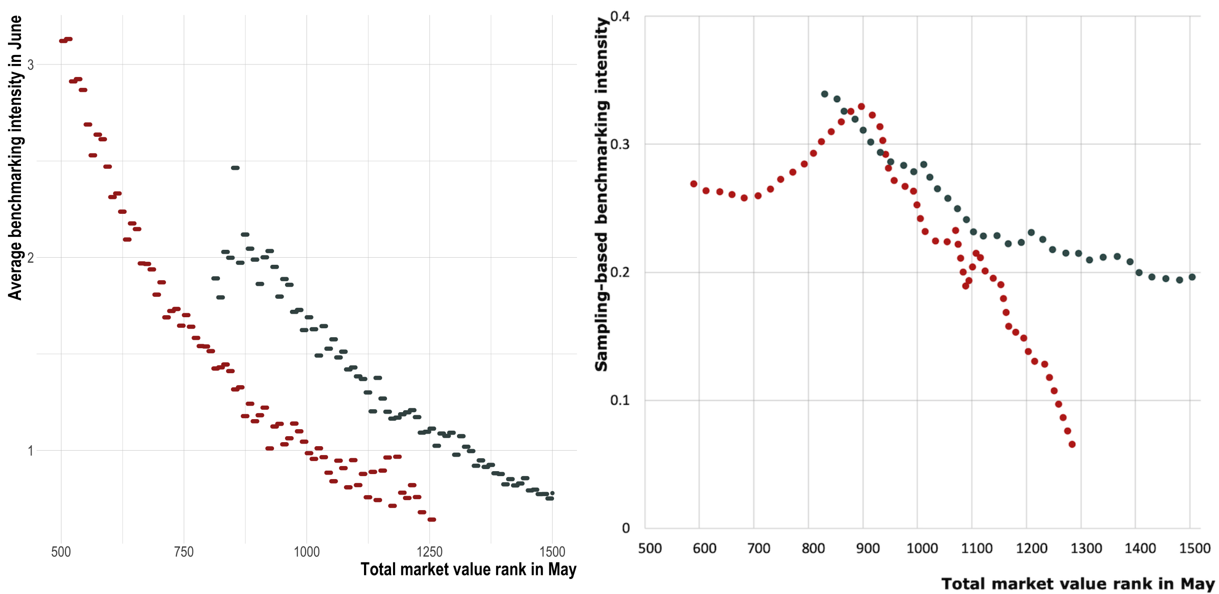


This figure illustrates the differences between a pure benchmark portfolio (left) and a portfolio constructed using optimized sampling (right). Horizontal bars represent stocks and their heights represent weights of these stocks in the respective portfolios.

display a discountinity.

Figure 6: Benchmarking intensity in 2007-2018 without and with optimized sampling

(a) BMI with no sampling (b) BMI with sampling



This figure plots average BMI in 2007-2018 based on benchmark weights (a) and BMI based on portfolio weights using optimized sampling (b) against total market value rank in May. Red (lower) plot indicates Russell 1000 constituents and grey (upper) plot – Russell 2000.

Results on the rebalancing of deletions after 2007 in Table 6 are in line with this illustration. When a stock gets added to the Russell 1000 (and therefore to Russell MidCap), it has a rank of around 800, while the ranks of existing index constituents range up to 1300. This addition now contributes to funds' tracking errors significantly more than smaller stocks at the bottom of the index and it is not as expensive to trade. In other words, funds benchmarked to the Russell 1000 and Russell MidCap are now more likely to purchase this addition. At the same time, additions to the Russell 2000 obtain a rank of around 1300. Because the existing constituents now have ranks starting from 800, the contribution of these additions to funds' tracking errors is, on average, lower compared to the pre-banding period. Even though passive funds benchmarked to the Russell 2000 will still trade these stocks, active funds are less likely to do so. Therefore, the incentives to hold stocks around the cutoffs changed with the introduction of banding, which must contribute to the performance of BMI in tests of long-run returns.

The change of these incentives provides an alternative explanation to the reduction of the index effect over time, documented in Chang, Hong, and Liskovich (2014). The authors hypothesize that the alleviation of limits to arbitrage over time made demand curves more elastic. We provide a different explanation: the introduction of banding made funds benchmarked to the Russell 1000 and Russell MidCap participate in index rebalancing almost at par with Russell 2000 funds. For example, the stocks that are being deleted from the Russell 2000 and experiencing selling pressure from Russell 2000 funds will also experience relatively higher buying pressure from Russell 1000/MidCap funds. In other words, we suggest evening out of the price pressure from buying and selling.

4.5 CRSP Indexes

A range of Vanguard passive funds switched from MSCI to CRSP indexes in 2012.[56] The switch concerned 9 funds that invest in stocks around the Russell cutoff. As Figure 1 in Introduction shows, at the time of the switch, these funds' AUM represented around 7% of assets in the 75%-95% of market capitalization range (around the cutoff). By 2018, this share grew to 15%, which is too high a number for researchers to ignore.

CRSP indexes have a different construction and reconstitution methodologies.[57] They do not have a fixed number of constituents and, instead, they include stocks that represent certain percentages of the US equity market capitalization. Moreover, rebalancing of CRSP indexes happens quarterly and over a 5-day period, when an index adds a 20%-fraction of the market value of a newly included stock on each day. The purpose of such a slow transition

[56]FTSE indices as well, for international funds. Media coverage is available online: https://www.ft.com/content/fa60b8b0-6655-11e2-919b-00144feab49a.

[57]Methodology guides are publicly available online: http://www.crsp.org/indexes-pages/crsp-us-equity-indexes-methodology-guide.

is to reduce expected rebalancing costs for the end investors.

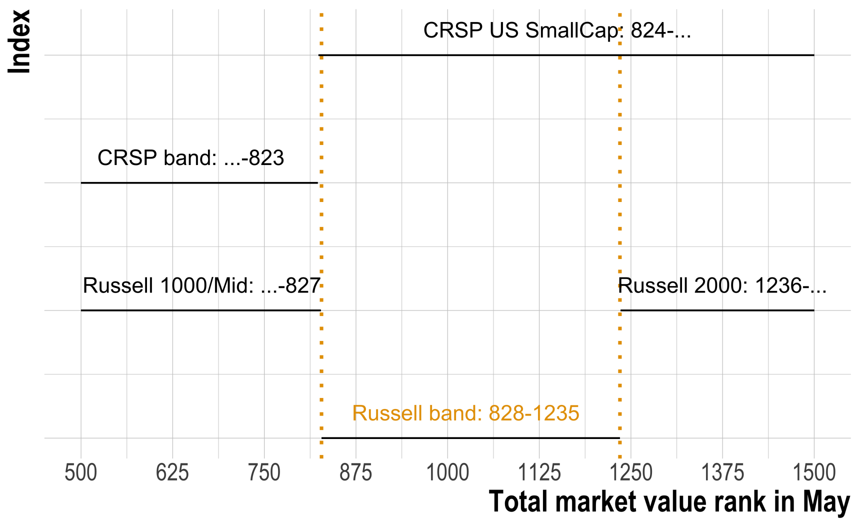
For index reconstitution, CRSP indexes use banding and 'pocketing' methodologies. The former is similar to Russell's band between the Russell 1000 and Russell 2000: it implies that a stock has to move further than the actual market value cutoff to be assigned to the new index. Pocketing is unique to CRSP indexes and it features partial, or 'pocketed', assignment to an index, in which only 50% of the market value of a stock gets added to the index once the stock passes the banding cutoff for that index. The next 50% get added if the stock remains beyond the cutoff until the next reconstitution. Both methodologies ensure the stability and representativeness of the index constitution.

CRSP indexes have a cutoff which overlaps with Russell's upper cutoff and may introduce confounding. Figure 7 illustrates this conflict in 2012. As stocks get reclassified to the CRSP US Small Cap, their BMI goes up. As former Russell 1000 stocks, these stocks become a control group in the test we perform on additions to the Russell 2000 index (provided they move sufficiently close to the lower Russell cutoff). With higher BMI, these stocks are subject to lower returns, which in turn brings returns of the control sample down and makes it less likely for our tests to identify the effect. In general, the existence of another cutoff right at the point of the Russell cutoff may violate the exclusion restriction. In our case, the restriction is satisfied as long as BMI accounts for the change and still exhibits discontinuity. Since CRSP reconstitution happens quarterly (and the share of the CRSP indexes keeps growing), it is more likely that a stock's BMI will change even if the stock remains in the neighborhood of the Russell cutoff. We see it as another reason of weaker results after 2013.

Another index present in the neighborhood of the cutoff is the S&P MidCap 400. It represents the next 400 most important companies in the US after the S&P 500. This may suggest that it has a cutoff around stocks with the rank of 900, but this is not quite the case. The methodology of S&P indices is different to that of Russell and CRSP: constituent selection is at the discretion of the Index Committee and sector balance is as important as market capitalization for inclusion.[58] Hence, the S&P 400 has a wide span of ranks (in the Russell rank terms). Instead of occupying ranks 501-900, it ranged from 172 until 2550 across all years in our sample.

[58]See https://us.spindices.com/documents/methodologies/methodology-sp-us-indices.pdf to access the S&P methodology publicly available online.

Figure 7: Overlap Between the CRSP and Russell Cutoffs



This figure depicts the Russell and CRSP cutoffs (band thresholds) effective at the end of June 2012. All thresholds are expressed in Russell ranks.

5 Conclusion

In this paper, we theoretically derive a measure which reflects the size of the preferred habitat investors in a stock – benchmarking intensity. Exploiting a discontinuity in the benchmarking intensity of stocks at the bottom of the Russell 1000 and the top of the Russell 2000 index, we document that stocks with higher benchmarking intensities have higher prices and lower expected returns.

Our measure reflects inelastic demand for their benchmarks of both active and passive funds. According to our preferred habitat view, active funds are not genuinely active investors. Rather, they simply deviate from their benchmarks to a lesser extent than passive funds. In our sample, active funds own large fractions of shares outstanding, higher than passive funds, and that is why they contribute significantly to the aggregate inelastic demand for benchmark stocks. On average, a large part of active funds' holdings is in benchmark stocks, both in terms of the number of stocks and AUM share.

Studying the rebalancing around the Russell cutoff, we document that both active and passive managers buy additions to their benchmarks and sell deletions. Our results also highlight that active managers participate in this rebalancing more after the introduction of banding. We explain why this is consistent with the optimized sampling practice.

We also discuss how the growth of the CRSP indexes may affect research design based on the Russell cutoff. The CRSP indexes have several cutoffs, which could potentially be exploited in research due to the mechanical reconstitution rules. This may inform the growing literature using the identification approach based on index cutoffs.

References

ABDI, F., AND A. RANALDO (2017): "A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices," *The Review of Financial Studies*, 30(12), 4437–4480.

AMIHUD, Y. (2002): "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets*, 5(1), 31–56.

APPEL, I., T. GORMLEY, AND D. KEIM (2016): "Passive Investors, not Passive Owners," *Journal of Financial Economics*, 121(1), 111–141.

——— (2019a): "Identification Using Russell 1000/2000 Index Assignments: A Discussion of Methodologies," *SSRN Electronic Journal*.

——— (2019b): "Standing on the Shoulders of Giants: The Effect of Passive Investors on Activism," *The Review of Financial Studies*.

BARBERIS, N., AND A. SHLEIFER (2003): "Style Investing," *Journal of Financial Economics*, 68(2), 161–199.

BASAK, S., AND A. PAVLOVA (2013): "Asset Prices and Institutional Investors," *American Economic Review*, 103(5), 1728–1758.

BEN-DAVID, I., F. FRANZONI, AND R. MOUSSAWI (2018): "Do ETFs Increase Volatility?," *The Journal of Finance*, 73(6), 2471–2535.

BRENNAN, M. (1993): "Agency and Asset Pricing," Working paper, UCLA.

BUFFA, A. M., AND I. HODOR (2018): "Institutional Investors, Heterogeneous Benchmarks and the Comovement of Asset Prices," *SSRN Electronic Journal*.

BUFFA, A. M., D. VAYANOS, AND P. WOOLLEY (2014): "Asset Management Contracts and Equilibrium Prices," *SSRN Electronic Journal*.

CHANG, Y.-C., H. HONG, AND I. LISKOVICH (2014): "Regression Discontinuity and the Price Effects of Stock Market Indexing," *Review of Financial Studies*, 28(1), 212–246.

CHEN, H., G. NORONHA, AND V. SINGAL (2005): "Index Changes and Unexpected Losses to Investors in S&P500 and Russell 2000 Index Funds," *SSRN Electronic Journal*.

CRANE, A. D., S. MICHENAUD, AND J. P. WESTON (2016): "The Effect of Institutional Ownership on Payout Policy: Evidence from Index Thresholds," *Review of Financial Studies*, 29(6), 1377–1408.

CUOCO, D., AND R. KANIEL (2011): "Equilibrium Prices in the Presence of Delegated Portfolio Management," *Journal of Financial Economics*, 101(2), 264–296.

De Bondt, W., and R. Thaler (1985): "Does the Stock Market Overreact?," *The Journal of Finance*, 40(3), 793–805.

Doshi, H., R. Elkamhi, and M. Simutin (2015): "Managerial Activeness and Mutual Fund Performance," *Review of Asset Pricing Studies*, 5(2), 156–184.

Floracki, C., A. Gregoriou, and A. Kostakis (2011): "Trading Frequency and Asset Pricing on the London Stock Exchange: Evidence From a New Price Impact Ratio," *Journal of Banking & Finance*, 35(12), 3335–3350.

Glossner, S. (2018): "The Effects of Institutional Investors on Firm Outcomes: Empirical Pitfalls of Quasi-Experiments Using Russell 1000/2000 Index Reconstitutions," *SSRN Electronic Journal*.

Gompers, P. A., and A. Metrick (2001): "Institutional Investors and Equity Prices," *The Quarterly Journal of Economics*, 116(1), 229–259.

Hacibedel, B., and J. van Bommel (2007): "Do Emerging Markets Benefit from Index Inclusion?," Money Macro and Finance (MMF) Research Group Conference 2006 128, Money Macro and Finance Research Group.

Hahn, J., P. Todd, and W. Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–209.

Harris, L., and E. Gurel (1986): "Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures," *The Journal of Finance*, 41(4), 815–829.

Heath, D., D. Macciocchi, R. Michaely, and M. C. Ringgenberg (2018): "Passive Investors are Passive Monitors," *SSRN Electronic Journal*.

Imbens, G., and K. Kalyanaraman (2011): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, 79(3), 933–959.

Kashyap, A., N. Kovrijnykh, J. Li, and A. Pavlova (2018): "The Benchmark Inclusion Subsidy," *NBER Working Paper 25337*.

——— (2020): "Is There Too Much Benchmarking in Asset Management?," Working paper, University of Chicago.

Kaul, A., V. Mehrotra, and R. Morck (2000): "Demand Curves for StocksDoSlope Down: New Evidence from an Index Weights Adjustment," *The Journal of Finance*, 55(2), 893–912.

Lee, D. S., and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2), 281–355.

Lewellen, J. (2015): "The Cross-section of Expected Stock Returns," *Critical Finance Review*, 4(1), 1–44.

Ma, L., Y. Tang, and J.-P. Gómez (2019): "Portfolio Manager Compensation in the

U.S. Mutual Fund Industry," *The Journal of Finance*.

McCrary, J. (2008): "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142(2), 698–714.

Pástor, Ľ., and R. F. Stambaugh (2003): "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111(3), 642–685.

Pástor, Ľ., R. F. Stambaugh, and L. A. Taylor (2015): "Scale and Skill in Active Management," *Journal of Financial Economics*, 116(1), 23–45.

Ritter, J. R. (1991): "The Long-Run Performance of Initial Public Offerings," *The Journal of Finance*, 46(1), 3–27.

Schmidt, C., and R. Fahlenbrach (2017): "Do Exogenous Changes in Passive Institutional Ownership Affect Corporate Governance and Firm Value?," *Journal of Financial Economics*, 124(2), 285–306.

Scholes, M. S. (1972): "The Market for Securities: Substitution Versus Price Pressure and the Effects of Information on Share Prices," *The Journal of Business*, 45(2), 179.

Sensoy, B. A. (2009): "Performance Evaluation and Self-Designated Benchmark Indexes in the Mutual Fund Industry," *Journal of Financial Economics*, 92(1), 25–39.

Shleifer, A. (1986): "Do Demand Curves for Stocks Slope Down?," *The Journal of Finance*, 41(3), 579–590.

Stambaugh, R. F. (2014): "Presidential Address: Investment Noise and Trends," *The Journal of Finance*, 69(4), 1415–1453.

Stock, J., and M. Yogo (2002): "Testing for Weak Instruments in Linear IV Regression," *NBER Technical Working Paper 284*.

Vayanos, D., and J.-L. Vila (2009): "A Preferred-Habitat Model of the Term Structure of Interest Rates," *NBER Working Paper 15487*.

Wei, W., and A. Young (2017): "Selection Bias or Treatment Effect? A Re-Examination of Russell 1000/2000 Index Reconstitution," *SSRN Electronic Journal*.

A Appendix

A.1 Construction of the Historical Benchmarks Data

We manually assemble a dataset of historical mutual funds benchmarks from the following sources:

1. Snapshot of benchmarks (*'primary_prospectus_benchmark* field) in Morningstar as of September 2018.

2. Database of historical fund prospectuses available on the website of the U.S. Securities and Exchange Commission (SEC)[59].

3. SEC Mutual Fund Prospectus Risk/Return Summary[60] data sets (MFRR). Benchmarks are mentioned in the annual return summary published in prospectuses.

We use the *crsp_fundno*-CIK mapping from CRSP to link CIK, SEC identifiers, back to *crsp_fundno*. To map CRSP and Morningstar, we mostly follow the procedure in Pástor, Stambaugh, and Taylor (2015), details are below in Section A.3.

A.1.1 Scraping the EDGAR and Building Text-Based Series

Mutual funds are required to reqularly submit filings to the SEC. The SEC's EDGAR system stores filings in electronic archives since 1994. Even though the SEC Rule S7-10-97[61] required funds to report their benchmark (or a 'reference broad market index') in prospectuses from December 1, 1999, some funds voluntarily did so prior to that (Sensoy (2009)). Reporting of manager compensation contracts was required by the SEC Rule S7-12-04[62] starting in October, 2004. Therefore, the procedure discussed below will cover the history of filings for any particular fund back to 1998.

The filings that include information on fund benchmark and manager compensation are: N-1A/485 (registration statement including a prospectus), 497K (summary prospectus), 497 (fund definitive materials) and 497J (certification of no change in definitive materials). All of these can be accessed via package 'edgarWebR' available in R.[63] Since the holdings

[59] Follow SEC's mutual fund search page: https://www.sec.gov/edgar/searchedgar/mutualsearch.html

[60] Follow the MFRR page: https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets.

[61] Available on: https://www.sec.gov/rules/final/33-7512r.htm.

[62] Available on https://www.sec.gov/rules/final/33-8458.htm.

[63] Description is available on: https://cran.r-project.org/web/packages/edgarWebR/index.html.

data set is already linked to CRSP fund identifiers (*fundno*), we will use all CIK codes[64] available in the mapping file *crsp_cik_map*. For each CIK, we retrieve a list of all historical filings (485 and 497/497K/497J forms) using *company_filings()* function. Then we parse the filings into raw text format using *parse_filing()* function.

Having obtained the filings for each CIK and each filing date, we re-organize the data set into a panel: quarterly text files for each fund. To do so, we assign observations with a 497J form a 'no-change' tag. Moreover, after looking at the text data, we assign a 'no-change' tag to 497 forms with no reference to benchmark or manager compensation.[65]

Before extracting the data, each of the filings is tokenized (we work with both tokenized text and string formats) and de-capitilized, punctuation and certain stop words are removed.[66] All these steps are done using NLTK[67] module in Python. Afterwards, we classify all 485 and 497K documents as prospectuses, while we have to look into the content of 497 filings to classify them into prospectuses or statements of additional information (SAI). Typically, funds specify the type of the document in the header, we therefore search for the exact match ('prospectus' or 'statement of additional information') in the first 100 characters of the filing.

There are a few challenges we face when extracting the fund benchmark from prospectus text. Even though all funds are required to disclose the benchmark, they tend to do it in a very different manner. Some funds explicitly say that the performance can be evaluated against a particular market index, some only report the index performance below the required performance tables (as implicit benchmark). If referring to the benchmark in the text, funds do not use standardized language: some may say 'benchmark', some 'market index' or 'reference index' and some may omit the term and only use a phrase similar to 'performance is measured against'. Moreover, some funds may define a mixture of indices as their benchmark, e.g., "60% Russell 1000, 40% Russell 2000". Therefore, we are faced with the task of extracting information from unstructured text.

Finally, in some cases we need to first isolate the text to extract the benchmark name from. Fund families may choose to submit one prospectus for many funds. Within one prospectus document, many funds can either share the same section or each fund can

[64] The Central Index Key (CIK) is used as the main identifier of the filing entities on the SEC's EDGAR and available per fund, fund series, and fund company. We first use series CIK as benchmarks differ at this level, then we use company CIK to fill in any missing observations.

[65] Since fund prospectus is a legal document and fund clientele supposedly depends on it, we se that prospectuses are relatively 'sticky' and hence the time series for most of the funds looks like 'prospectus' definition at early date and then at most 1-2 changes for the fund history.

[66] Numerical data and special characters cannot be removed though as they are included in benchmark names. Moreover, we retain negation.

[67] Official page is: http://www.nltk.org/.

have a separate section. We therefore extract the fund-relevant part of prospectus whenever possible (typically in the second case only). To do so, we search for fund name and fund ticker in the text. Most commonly, the relevant section starts with a ticker/name and has it repeated on each page throughout the section. We hence extract the part of the text with the highest density of tickers/fund names.

When extracting benchmarks from (isolated) text, we use a set of rules that maximizes the chance of the algorithm picking up the benchmark correctly. The set of rules includes but is not limited to:

- Search for a benchmark provider name from the list (de-capitilized already): {*s&p, russell, crsp, msci, dj, dow jones, nasdaq, ftse, schwab, barclays, wilshire, bridgeway, guggenheim, calvert, kaizen, lipper, redwood, w.e. donoghue, essential treuters, barra, ice bofaml, bbgbarc, cboe*}.[68] If a benchmark from the list is found, retrieve subsequent 40 characters to extract the full benchmark name. Match the full names using the list from Morningstar (for example, *russell 1000 value tr usd*).

- If several matches are established, we record the number of matches and each benchmark name (with subsequent characters, as above).

- We also search for words from the list (*context words*): {*index, benchmark, reference, performance, relative, return, measure, evaluate, assess, calculate*}. We use these words in two ways. Firstly, if a benchmark name match is established, we check if any of these *context words* is present within 100 characters around the name. Secondly, if no match is established, we record pairwise distance in letters between benchmark names and *context words* and return the pair with minimum distance. This second approach is based on string format of the text and required if the match was not established due to imprecision in tokenization.

We manually clean the extracted data to remove typos and map it to full benchmark names. In the resulting sample of quarter-fund-benchmarks, we manually verify all funds that got matched with several benchmarks or that had a benchmark change. Subsequently, we validate a random sample of funds through manual analysis of prospectus' text. We also compare the benchmarks as of September 2018 with a snapshot we obtained from Morningstar database and manually resolve any mismatch. Furthermore, we compare a time series we get with a series available for a small sample of funds in MFRR.

[68] This list has been compiled using the Morningstar benchmark snapshot. It is survivorship-bias free. According to Morningstar, the first three providers take over 90% of the market and the first five - around 98%.

As expected, prospectuses are relatively sticky. In the entire sample over 1998-2018, we observe 1,208 changes at a share class level (around 300 at master fund level). The largest benchmark change in terms of tracking assets for passive funds in Vanguard's move from MSCI to CRSP indexes in 2012 and 2013. For active funds, it is T. Rowe Price's change from the S&P 500 to Russell 1000 Value and Growth indexes in 2018.

A.2 CRSP and Thomson Reuters S12 Merge Procedure

We use Mutual Fund Links (MFLINKS) to merge CRSP and TRS12 similar to the procedure described in Doshi, Elkamhi, and Simutin (2015).

Firstly, we prepare TRS12 holdings:

- keep last holdings report for each fund in a given month,
- match WFICN number from MFLINKS to fundno, rdate, and fdate in TRS12 file,
- when there are duplicate reports for the same date, keep the fund with largest assets,
- pull CRSP stock files and adjust reported number of shares by the correct adjustment factor - as of rdate.

Then, we prepare CRSP holdings:

- clean the data based on portnomap to ensure that only one portno is valid for a particular date for any fund (remove overlaps in the data due to mergers),
- match WFICN number from MFlinks to crsp_fundno,
- clean overlaps in wficn-portno mapping,
- keep the last report for every month.

Finally, we stack the two parts and remove duplicate entries from CRSP (at a fund level).

A.3 CRSP and Morningstar Merge Procedure

The merge procedure is a slight modification of Pástor, Stambaugh, and Taylor (2015).[69]

A.4 Asset Validation

TNA and holdigns data are generally validated by MFLINKS (only funds with sufficient match quality are linked). However, we additionally validate the TNA in order to ensure better match with the holdings. In case of CRSP, we use the sum of assets across share classes and weigh share class level data such as equity percentage by the fraction of

[69]Details are available upon request.

total assets this share class represents. Because TRS12 reports only equity and CRSP reports all assets, we multiply the most recent equity percentage by CRSP assets. We use the following for validation:
- compare the total dollar sum of holdings in the merged file with the assets reported by TRS12 and CRSP and call the difference 'unexplained',
- if difference between TRS12 and CRSP is smaller than 1%, we use CRSP,
- if CRSP has lower unexplained or TRS12 does not report assets, we use CRSP and otherwise TRS12.

A.5 Filtering

In the final sample, we keep only funds that:
- have fund-quarter entries where I validated the assets at 20% precision;
- are either active or passive domestic equity funds that did not change its style or objective over their history (see details below in Section A.6);
- have an average common equity percentage between 50 and 120%;
- have more than USD 1 million in assets.

A.6 Active and Passive Domestic Equity Funds

We follow the major steps of the procedure described in Doshi, Elkamhi, and Simutin (2015) to filter out active domestic equity funds and augment it to identify passive funds better.

We use *crsp_obj_cd* (CRSP objective code) to identify 'equity', 'domestic', 'cap-based or style' and exclude 'hedged' and 'short' and remove those funds that changed their objectives. I also only keep funds with 'ioc' variable in TRS12 file (investment objective) not in (1,5,6,7). Active funds are identified as those without '$Index\_fund\_flag'$ or with 'B' (index-based funds) and without 'et_flag'. I also exclude funds that have a range of words in their names, as per the list below.

List of n-grams to exclude from active funds names (all in lower case).

1. Generic and index provider names: index, indx, ' idx ', s&p, ' sp ' (with spaces), nasdaq, msci, crsp, ftse, barclays, ' dj ', ' dow ', jones, russell, ' nyse ', wilshire, 400, 500, 600, 1000, 1500, 2000, 2500, 3000, 5000

2. Passive management names: ishares, spdr, trackers, holdrs, powershares, streettracks, ' dfa ', 'program', etf, exchange traded, exchange-traded

3. Target fund names: target, retirement, pension, 2005, 2010, 2015, 2020, 2025, 2030, 2035, 2040, 2045, 2050, 2055, 2060, 2065, 2070, 2075

Our sample of passive funds consists of index funds and ETFs available on CRSP. Index funds are those with '$index\_fund\_flag'$ of 'D' or 'E' and those that include a range of words in their name:

1. Generic and index provider names: index, indx, ' idx ', s&p, ' sp ' (with spaces), nasdaq, msci, crsp, ftse, barclays, ' dj ', ' dow ', jones, russell, ' nyse ', wilshire, 400, 500, 600, 1000, 1500, 2000, 2500, 3000, 5000
2. Passive management names: ishares, ' dfa ', 'program'

ETFs are those with not missing 'et_flag' or having $'etf'$, '$exchange - traded'$, '$exchangetraded'$ in their name:

1. Passive management names: spdr, trackers, holdrs, powershares, streettracks, etf, exchange traded, exchange-traded

Target funds are those with target years in the name, e.g., '2015' and '2075', or 'retirement', 'target'. Creating a clean sample of target funds potentially requires different treatment of objective codes (see CRSP Style Guide). Since we only aim to exclude them, we remove fund with the following n-grams in their names:

1. Target fund names: target, retirement, pension, 2005, 2010, 2015, 2020, 2025, 2030, 2035, 2040, 2045, 2050, 2055, 2060, 2065, 2070, 2075

We exclude all leverage and inverse funds by identifying the following n-grams in the names: $'leverage'$, $'inverse'$, $'2x'$, $'1.5x'$, $'1.25x'$, $'2.5x'$, $'3x'$, $'4x'$.

If we apply the rules above, some of the funds in the sample will include both active and passive share classes. We clean the resulting sample of funds with share classes of different type as per the rule: (a) Put ETF shares of index funds as ETFs (passive type maintained). (b) When missing flag for otherwise index funds and portno is the same, set to index. (c) If $portno/cl\_grp$ are different, exclude.

The remaining funds are further filtered based on the common equity percentage as discussed in A.5.

A.7 Russell Reconstitution

Table 7: Historical Details on Russell 2000 Reconstitution

			Russell 1000		Russell 2000	
Year	Additions	Deletions	Smallest	Smallest w/banding	Largest w/banding	Largest
1987	38	37	0.4			0.4
1988	22	35	0.3			0.3
1989	21	9	0.4			0.4
1990	47	28	0.3			0.3
1991	56	22	0.4			0.4
1992	55	26	0.5			0.5
1993	63	26	0.6			0.6
1994	66	36	0.7			0.7
1995	48	39	0.8			0.8
1996	65	38	1.0			1.0
1997	62	57	1.1			1.1
1998	57	54	1.4			1.4
1999	59	70	1.4			1.4
2000	50	48	1.6			1.5
2001	86	104	1.4			1.4
2002	78	73	1.3			1.3
2003	43	56	1.2			1.2
2004	49	38	1.6			1.6
2005	61	58	1.8			1.7
2006	49	68	2.0			1.9
2007	5	15	2.5	1.8	3.1	2.5
2008	31	38	2.0	1.4	2.7	2.0
2009	36	39	1.2	0.8	1.7	1.2
2010	14	25	1.7	1.3	2.2	1.7
2011	23	35	2.2	1.6	3.0	2.2
2012	27	32	2.0	1.4	2.6	1.9
2013	27	30	2.5	1.8	3.3	2.5
2014	28	24	3.1	2.2	4.1	3.1
2015	48	20	3.4	2.4	4.3	3.4
2016	48	34	2.9	2.0	3.9	2.9
2017	40	31	3.4	2.3	4.5	3.4
2018	35	48	3.7	2.5	5.0	3.7

This table reports the number of additions to and deletions from Russell 2000. We only report deletions which moved to Russell 1000, not those that moved down in the ranking. The last for columns report the market value (in billions USD) of smallest and largest stocks in the indices.

A.8 Assignment Prediction

Table 8: Quality of the assignment prediction

	1998-2006 sample		*2007-2018 sample*	
	D^{Ru1000}	D^{Ru2000}	D^{Ru1000}	D^{Ru2000}
$\tau_{i,t}$	0.82^{***}	0.84^{***}	0.67^{***}	0.68^{***}
	(35.20)	(43.97)	(15.59)	(21.87)
F-statistic	2,239	3,214	594	1,142
Adjusted R^2, %	90	90	82	84
Observations	712	1,022	386	660

This table reports the results of the assignment prediction regression: $D_{it}^{index} = \alpha_{0l} + \alpha_{1l}(Rank_{it} - c) + \tau_{it}(\alpha_{0r} + \alpha_{1r}(Rank_{it} - c))$ *(Chang, Hong, and Liskovich (2014)). Indicator* τ *is 1 if the stock is on the right side of the cutoff* c *to be assigned to the index. We include only stocks that were in the Russell 1000 (for additions) or Russell 2000 (deletions) in the previous year. The dependent variables are, respectively: Russell 2000 membership dummy,* D^{RU2000}*, and Russell 1000 membership dummy,* D^{RU1000}*. Bandwidth is 100. t-statistics based on HAC-robust standard errors with clusters at a firm level are in parentheses. Significance levels are marked as:* $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$.

A.9 First Stage Results

Table 9: First stage regression results

	D_{it}^{R2000}: **stock $\in$ Russell 2000 index**							
	1998-2006 sample		*2007-2018 sample*		*1998-2006 sample*		*2007-2018 sample*	
	Additions	*Deletions*	*Additions*	*Deletions*	*Additions*	*Deletions*	*Additions*	*Deletions*
τ_{it}	0.963***	0.964***	0.964***	0.949***	0.953***	0.950***	0.923***	0.906***
	(110.32)	(118.41)	(98.93)	(86.19)	(92.37)	(81.67)	(59.86)	(42.82)
Band width	300				100			
Running variable ($logMV$)	Yes				No			
Observations	2,181	2,652	1,343	2,096	958	657	650	356
F-statistic	30,147	25,953	7,223	14,982	8,532	6,670	3,583	1,834
Adjusted R^2, %	95	96	90	91	91	89	82	79

*This table reports the results of the first stage regression (5) for stocks in the pre-banding (1998-2006) and the post-banding (2007-2018) samples. The dependent variable is the dummy for Russell 2000 membership of a stock i as of June in year t. We include only stocks that were in Russell 1000 (additions) or Russell 2000 (deletions) in the previous year. Band width is 300 or 100 stocks around the cutoffs (rectangular kernel). t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are marked as: *p<0.05; **p<0.01; ***p<0.001.*

A.10 Descriptive Statistics

Table 10: Descriptive statistics around the cut-off

	Stocks in Russell 1000					*Stocks in Russell 2000*				
	Obs.	Mean	St.Dev.	Min	Max	Obs.	Mean	St.Dev.	Min	Max
BMI	5,354	0.11	0.05	0.00	0.68	6,178	0.11	0.05	0.00	0.43
Average long-run excess return, % (winsorized at 1%):										
12-month	5,091	1.02	2.67	-11.18	12.28	5,847	1.00	2.71	-11.18	12.28
24-month	4,605	0.94	1.86	-7.11	8.35	5,255	1.00	1.93	-7.11	8.35
36-month	4,149	0.94	1.50	-4.86	6.34	4,706	1.00	1.50	-4.86	6.34
48-month	3,720	0.96	1.26	-3.76	5.51	4,211	0.98	1.28	-3.76	5.51
60-month	3,316	1.00	1.11	-3.04	4.83	3,798	0.99	1.13	-3.04	4.83
Average periodic excess return, % (winsorized at 1%):										
0-12 months	5,084	0.58	2.70	-15.04	9.27	5,840	0.56	2.82	-15.04	9.27
12-24 months	4,605	0.21	2.94	-14.26	8.84	5,261	0.30	2.95	-14.25	8.84
24-36 months	4,151	0.20	2.91	-13.39	8.73	4,712	0.31	2.97	-13.39	8.73
36-48 months	3,723	0.36	2.83	-12.73	8.51	4,216	0.23	2.97	-12.73	8.51
48-60 months	3,324	0.38	2.81	-11.98	8.30	3,806	0.38	2.90	-11.98	8.30
Bid-ask spread, %	5,370	0.14	0.14	0.00	2.00	6,200	0.15	0.16	0.00	4.68
β^{CAPM} (winsorized at 1%)	5,104	1.14	0.69	-0.08	3.56	5,876	1.16	0.71	-0.08	3.56
Market value (Russell)	5,419	2663.2	1094.8	826.2	6193.8	6,272	1925.7	870.2	778.9	5043.6
Last-year return, % (winsorized at 1%)	5,255	5.67	35.08	-82.05	246.90	6,117	20.31	44.73	-82.05	246.90

This table reports the descriptive statistics of the main stock-level variables used in the analysis – by index the stock belongs to in the current year. These statistics are calculated on 300 stocks around the cut-off. All returns are monthly.

A.11 Third Stage Results Before and After Introduction of Banding

Table 11: Third stage results, by sample subperiod

Long-run excess return

Panel A: Excess returns, average over horizon (months)

	Additions to Russell 2000					*Deletions from Russell 2000*				
Horizon (months)	12	24	36	48	60	12	24	36	48	60
Panel A1: Pre-banding (1998-2006)										
$\widehat{BMI}_{it}$	**-0.30*** **	**-0.31*** **	**-0.23*** **	**-0.13*** **	-0.03	**-0.24*** **	**-0.35*** **	**-0.22*** **	**-0.16*** **	**-0.09*** **
	(-4.53)	(-5.41)	(-5.20)	(-3.76)	(-0.95)	(-3.49)	(-6.21)	(-5.24)	(-4.31)	(-2.80)
Observations	1921	1811	1713	1628	1548	2285	2159	2033	1930	1833
Panel A2: Post-banding (2007-2018)										
$\widehat{BMI}_{it}$	**-0.63*** **	**-0.18***	-0.01	-0.03	-0.09	**-0.14** **	**-0.08** **	**-0.06** **	**-0.06** **	**-0.07*** **
	(-2.62)	(-1.39)	(-0.09)	(-0.28)	(-0.77)	(-2.10)	(-1.95)	(-1.71)	(-2.03)	(-2.68)
Observations	1175	1031	888	752	607	1864	1591	1368	1164	1003
Panel A3: After CRSP switch (2013-2018)										
$\widehat{BMI}_{it}$	**-0.60*** **	-0.01	0.02	0.01	-0.08	**-0.27*** **	0.06	0.07	0.05	-0.05
	(-3.57)	(-0.08)	0.25	(0.15)	(-0.73)	(-2.46)	(0.91)	(1.43)	(0.94)	(-1.14)
Observations	635	508	389	280	160	922	703	515	352	234

Panel B: Excess returns, average in the period (months)

	Additions to Russell 2000					*Deletions from Russell 2000*				
Period (months)	0-12	12-24	24-36	36-48	48-60	0-12	12-24	24-36	36-48	48-60
Panel B1: Pre-banding (1998-2006)										
$\widehat{BMI}_{it}$	**-0.25*** **	**-0.34*** **	**-0.38*** **	-0.01	0.30	**-0.13** **	**-0.44*** **	**-0.28*** **	**-0.17** **	0.10
	(-3.85)	(-4.02)	(-4.80)	(-0.10)	(3.37)	(-1.92)	(-5.51)	(-3.32)	(-2.07)	(1.21)
Observations	1921	1811	1713	1628	1548	2285	2159	2033	1930	1833
Panel B2: Post-banding (2007-2018)										
$\widehat{BMI}_{it}$	**-0.66*** **	**-0.39** **	-0.30	**-0.56** **	**-1.08*** **	**-0.12***	**-0.10***	-0.03	**-0.10** **	**-0.10** **
	(-2.70)	(-1.86)	(-1.15)	(-1.84)	(-2.61)	(-1.63)	(-1.51)	(-0.58)	(-1.73)	(-1.74)
Observations	1175	1031	888	752	607	1864	1591	1368	1164	1003
Panel B3: After CRSP switch (2013-2018)										
$\widehat{BMI}_{it}$	**-0.61*** **	0.13	-0.02	-0.01	**-0.66** **	**-0.29*** **	0.13	0.10	-0.09	**-0.35** **
	(-3.48)	(0.64)	(-0.09)	(-0.04)	(-1.76)	(-2.47)	(1.18)	(1.12)	(-0.91)	(-2.31)
Observations	635	508	389	280	160	922	703	515	352	234

This table reports the results of the third stage regression for the subsamples: 1998-2006 (Panels A1 and B1), 2007-2018 (Panels A2 and B2), 2013-2018 (Panels A3 and B3). The dependent variable in Panel A is an average monthly excess return from September in year t over the respective horizon. The dependent variable in Panel B is an average monthly return in the respective period, e.g., 12-24 months after reconstitution. We include only stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. Band width is 300 stocks around the cutoffs (rectangular kernel). All regressions include the baseline controls: log total market value, a float factor control, a 5-year monthly rolling β^{CAPM}, a 1-year monthly rolling average Bid-Ask Spread, %, stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters by stock are in parentheses. Significance levels are based on a one-sided test and marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

A.12 Narrow Band

Table 12: Third stage results, by sample subperiod and with a narrow band

	Additions to Russell 2000					*Deletions from Russell 2000*				
Panel A: Excess returns, average over horizon (months)										
Horizon (months)	12	24	36	48	60	12	24	36	48	60
Panel A1: Pre-banding (1998-2006)										
$\widehat{BMI}_{it}$	-0.215***	-0.209***	-0.148***	-0.082***	-0.009	-0.127**	-0.178***	-0.109***	-0.072**	-0.043*
	(-2.96)	(-3.68)	(-3.43)	(-2.40)	(-0.25)	(-2.03)	(-3.57)	(-2.80)	(-2.21)	(-1.47)
Observations	577	552	528	499	477	834	798	752	708	679
Panel A2: Post-banding (2007-2018)										
$\widehat{BMI}_{it}$	-0.232**	-0.040	-0.007	0.000	0.018	-0.014	0.006	0.007	0.006	-0.010
	(-1.87)	(-0.53)	(-0.09)	(0.01)	(0.27)	(-0.34)	(0.19)	(0.29)	(0.28)	(-0.51)
Observations	307	262	218	182	141	576	498	432	359	301
Panel B: Excess returns, average in the period (months)										
	Additions to Russell 2000					*Deletions from Russell 2000*				
Period (months)	0-12	12-24	24-36	36-48	48-60	0-12	12-24	24-36	36-48	48-60
Panel B1: Pre-banding (1998-2006)										
$\widehat{BMI}_{it}$	-0.171***	-0.232***	-0.226***	0.005	0.255	-0.071	-0.215***	-0.070	-0.086	0.069
	(-2.40)	(-2.58)	(-2.81)	(0.06)	(2.57)	(-1.15)	(-3.19)	(-0.98)	(-1.07)	(0.98)
Observations	577	552	529	500	478	834	798	752	708	680
Panel B2: Post-banding (2007-2018)										
$\widehat{BMI}_{it}$	-0.205*	-0.033	-0.125	-0.125	-0.233*	-0.010	0.042	-0.018	0.005	-0.042
	(-1.73)	(-0.26)	(-0.81)	(-0.62)	(-1.61)	(-0.24)	(0.94)	(-0.44)	(0.14)	(-0.94)
Observations	307	262	218	182	141	576	498	432	359	301

This table reports the results of the third stage regression for the subsamples: 1998-2006 (Panels A1 and B1), 2007-2018 (Panels A2 and B2). The dependent variable in Panel A is an average monthly excess return from September in year t over the respective horizon. The dependent variable in Panel B is an average monthly return in the respective period, e.g., 12-24 months after reconstitution. We include only stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. Band width is 100 stocks around the cutoffs (rectangular kernel). All regressions include the baseline controls: log total market value, a float factor control, a 5-year monthly rolling β^{CAPM}, a 1-year monthly rolling average Bid-Ask Spread, %, stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters by stock are in parentheses. Significance levels are based on a one-sided test and marked as: *$p<0.10$; **$p<0.05$; ***$p<0.01$.

A.13 Abnormal Returns

Table 13: Third stage results, by sample subperiod

Long-run abnormal return
CAPM abnormal returns, average over horizon (months)

	Additions to Russell 2000					*Deletions from Russell 2000*				
Horizon (months)	12	24	36	48	60	12	24	36	48	60
Panel A1: Pre-banding (1998-2006)										
$\widehat{BMI}_{it}$	-0.20***	-0.18***	-0.09**	0.01	0.06	-0.18***	-0.22***	-0.06*	0.02	0.02
	(-2.92)	(-3.20)	(-1.93)	(0.36)	(1.52)	(-2.81)	(-4.10)	(-1.48)	(0.42)	(0.49)
Observations	1921	1811	1713	1628	1548	2285	2159	2033	1930	1833
Panel A2: Post-banding (2007-2018)										
$\widehat{BMI}_{it}$	-0.80***	-0.24*	-0.00	0.07	0.14	-0.18***	-0.05	-0.02	-0.01	-0.00
	(-3.13)	(-1.60)	(-0.02)	(0.66)	(1.19)	(-2.84)	(-1.20)	(-0.66)	(-0.28)	(-0.09)
Observations	1175	1031	888	752	607	1864	1591	1368	1164	1003

Market model abnormal returns, average over horizon (months)

	Additions to Russell 2000					*Deletions from Russell 2000*				
Horizon (months)	12	24	36	48	60	12	24	36	48	60
Panel B1: Pre-banding (1998-2006)										
$\widehat{BMI}_{it}$	-0.17***	-0.16***	-0.08**	0.02	0.06	-0.15**	-0.24***	-0.07**	0.02	0.03
	(-2.54)	(-2.91)	(-1.95)	(0.56)	(1.82)	(-2.30)	(-4.53)	(-1.73)	(0.65)	(0.89)
Observations	1921	1811	1714	1628	1549	2287	2161	2034	1930	1834
Panel B2: Post-banding (2007-2018)										
$\widehat{BMI}_{it}$	-0.51***	-0.21*	0.00	0.02	0.04	-0.14**	-0.06*	-0.03	-0.02	-0.02
	(-2.60)	(-1.58)	(0.02)	(0.19)	(0.38)	(-2.30)	(-1.32)	(-0.95)	(-0.68)	(-0.67)
Observations	1169	1030	887	752	607	1860	1593	1369	1165	1004

This table reports the results of the third stage regression for the subsamples: 1998-2006 (Panels A1 and B1) and 2007-2018 (Panels A2 and B2). The dependent variable is an average monthly CAPM or market model abnormal return from September in year t over the respective horizon. We include only stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. Band width is 300 stocks around the cutoffs (rectangular kernel). All regressions include the log total market value and the baseline controls: a float factor, a 5-year monthly rolling β^{CAPM}, a 1-year monthly rolling average Bid-Ask Spread, %, stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters by stock are in parentheses. Significance levels are based on a one-sided test and marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

A.14 Covariate Imbalance Tests

Table 14: Covariate imbalance tests

	Leverage	ROA	Repurchase	Div. yield	ΔSales	Lerner	$\frac{Capex}{Sales}$	$\frac{M}{B}$	$\frac{R\&D}{Sales}$	$\frac{\Delta Debt}{Assets}$	$\frac{\Delta Stock}{Assets}$	Assets	Acquisitions	Tobin's Q
Panel A: Pre-banding period (1998-2006)														
Panel A1: Additions (Russell 1000 stocks)														
D^{RU2000}	0.00	0.01	**0.01****	0.00	-0.01	0.01	-0.00	0.03	-0.01	0.01	-0.01	123.520	**0.071***	1.436
	(0.15)	(1.50)	(2.02)	(1.23)	(-0.11)	(0.29)	(-0.32)	(0.26)	(-0.71)	(0.95)	(-1.17)	(0.40)	(1.96)	(0.67)
Obs	2034	2034	1780	2031	2030	2034	2034	2034	2034	1899	1893	2034	2035	1682
Panel A2: Deletions (Russell 2000 stocks)														
D^{RU2000}	-0.02	**0.01****	0.00	-0.00	0.03	0.72	-0.01	**-0.17***	-0.10	0.00	**-0.02****	39.884	0.003	-1.221
	(-1.49)	(2.35)	(1.55)	(-1.03)	(0.61)	(0.98)	(-0.30)	(-1.77)	(-0.77)	(0.43)	(-2.60)	(0.24)	(0.09)	(-0.64)
Obs	2437	2437	2053	2434	2432	2437	2437	2437	2437	2274	2272	2437	2437	1978
Panel B: Post-banding period (2007-2018)														
Panel B1: Additions (Russell 1000 stocks)														
D^{RU2000}	-0.03	-0.02	**-0.01***	-0.01	0.00	-0.32	**0.01****	0.14	0.23	-0.01	0.01	-764.652	-0.047	-0.621
	(-1.41)	(-1.44)	(-1.85)	(-1.61)	(0.03)	(-1.30)	(2.01)	(1.56)	(1.31)	(-1.27)	(1.12)	(-0.93)	(-1.16)	(-0.09)
Obs	1230	1230	1178	1230	1226	1229	1230	1230	1229	1230	1230	1230	1230	1017
Panel B2: Deletions (Russell 2000 stocks)														
D^{RU2000}	0.01	-0.00	0.00	0.00	**-0.24*****	7.85	-0.01	**-0.21***	-3.97	0.01	**-0.02*****	**732.667***	0.022	-41.453
	(0.68)	(-0.48)	(1.64)	(0.52)	(-3.48)	(1.19)	(-1.23)	(-1.65)	(-1.30)	(0.74)	(-3.15)	(1.79)	(0.66)	(-0.98)
Obs	1952	1952	1817	1950	1941	1945	1952	1952	1945	1950	1947	1952	1956	1596

This table reports the covariate imbalance tests by sample subperiod. We include only stocks that were in the Russell 1000 (Panel A1 and B1) or in the Russell 2000 (Panel A2 and B2) in the previous year. We limit the sample to 300 stocks around the cutoffs (rectangular kernel). All tests include the baseline controls: log total market value, a float factor control, a 5-year monthly rolling β^{CAPM}, a 1-year monthly rolling average Bid-Ask Spread, %, stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters by stock are in parentheses. Significance levels are marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

A.15 Tests on Long-Run Return Drivers

Table 15: Tests on long-run return drivers (financial characteristics)

	Leverage	ROA	Repurchase	Div. yield	ΔSales	Lerner	$\frac{Capex}{Sales}$	$\frac{M}{B}$	$\frac{R\&D}{Sales}$	$\frac{\Delta Debt}{Assets}$	$\frac{\Delta Stock}{Assets}$	Assets	Acquisitions	Tobin's Q
Panel A: Pre-banding period (1998-2006)														
Panel A1: Additions (Russell 1000 stocks)														
τ^{RU2000}	-0.00	0.00	**0.01****	0.00	-0.04	-0.06	0.00	-0.03	0.05	0.01	**-0.01*****	-343.35	-0.02	-1.42
	(-0.04)	(0.55)	(2.42)	(0.88)	(-1.48)	(-1.12)	(0.84)	(-0.36)	(0.91)	(1.11)	(-3.30)	(-1.05)	(-0.57)	(-1.12)
Obs	1720	1720	1492	1716	1720	1720	1720	1720	1720	1630	1626	1720	1720	1430
Panel A2: Deletions (Russell 2000 stocks)														
τ^{RU2000}	-0.00	0.00	**0.01*****	0.00	**-0.05***	0.02	0.00	**-0.31*****	-0.01	0.00	**-0.02*****	-345.11	0.01	-2.32
	(-0.09)	(0.94)	(4.66)	(0.19)	(-1.79)	(0.46)	(0.78)	(-3.60)	(-0.45)	(0.54)	(-4.54)	(-1.52)	(0.42)	(-1.47)
Obs	2039	2039	1679	2034	2037	2037	2039	2039	2037	1916	1917	2039	2039	1635
Panel B: Post-banding period (2007-2018)														
Panel B1: Additions (Russell 1000 stocks)														
τ^{RU2000}	-0.01	0.00	-0.00	0.00	-0.03	0.00	0.00	**0.17****	0.00	0.01	-0.00	**-1429.91***	-0.04	3.62
	(-0.50)	(0.34)	(-0.02)	(0.10)	(-1.10)	(0.01)	(0.26)	(2.24)	(0.09)	(1.06)	(-0.28)	(-1.88)	(-0.84)	(1.23)
Obs	800	800	739	800	798	800	800	800	800	800	800	800	800	653
Panel B2: Deletions (Russell 2000 stocks)														
τ^{RU2000}	0.01	-0.00	-0.00	0.00	**-0.13****	-2.97	**-0.01***	0.03	1.37	0.00	-0.01	-726.78	-0.01	-5.87
	(0.45)	(-0.41)	(-0.01)	(0.33)	(-2.67)	(-1.00)	(-1.86)	(0.17)	(0.96)	(0.52)	(-0.95)	(-1.04)	(-0.19)	(-0.46)
Obs	1236	1236	1110	1235	1229	1233	1236	1236	1233	1234	1232	1236	1236	1013

This table reports the validity tests by sample subperiod. We include only stocks that were in the Russell 1000 (Panel A1 and B1) or in the Russell 2000 (Panel A2 and B2) in the previous year. We limit the sample to 300 stocks around the cutoffs (rectangular kernel). Dependent variable is a 3-year average of the respective variable after the reconstitution. All tests include the baseline controls: log total market value, a float factor control, a 5-year monthly rolling β^{CAPM}, a 1-year monthly rolling average Bid-Ask Spread, %, stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters by stock are in parentheses. Significance levels are marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

A.16 Reduced Form Regressions

Table 16: Reduced form results for 1998-2006

	Long-run excess return									
	Additions to Russell 2000					*Deletions from Russell 2000*				
	12-month	24-month	36-month	48-month	60-month	12-month	24-month	36-month	48-month	60-month
Panel A: All stocks, bandwidth is 100, only $logMV$ included										
τ_{it}	-0.009***	-0.008***	-0.006***	-0.003***	-0.001	-0.004*	-0.006***	-0.004***	-0.003***	-0.001*
	(-3.53)	(-4.30)	(-3.72)	(-2.54)	(-0.75)	(-1.77)	(-3.85)	(-3.01)	(-2.60)	(-1.56)
Adjusted R^2, %	1.96	5.94	5.38	3.29	0.30	1.06	4.63	5.91	4.00	0.76
Observations	612	584	558	525	503	888	845	791	744	711
Panel B: Flippers excluded, bandwidth is 100, only $logMV$ included										
τ_{it}	-0.020***	-0.022***	-0.020***	-0.015***	-0.012***	-0.015***	-0.020***	-0.018***	-0.016***	-0.012***
	(-7.03)	(-9.86)	(-10.51)	(-8.96)	(-7.27)	(-6.47)	(-10.56)	(-11.41)	(-10.89)	(-7.75)
Adjusted R^2, %	10.63	27.03	31.10	27.34	21.95	7.34	21.46	31.60	32.98	22.69
Observations	426	342	292	256	237	554	433	354	299	271
Panel C: All stocks, bandwidth is 300, only $logMV$ included										
τ_{it}	-0.010***	-0.010***	-0.008***	-0.004***	-0.001	-0.005***	-0.009***	-0.006***	-0.004***	-0.002***
	(-5.02)	(-6.41)	(-6.16)	(-4.16)	(-1.16)	(-3.11)	(-7.20)	(-6.14)	(-4.96)	(-2.87)
Adjusted R^2, %	1.64	4.40	5.59	3.86	1.04	0.43	3.47	3.33	2.24	0.60
Observations	2019	1901	1796	1700	1619	2443	2299	2161	2050	1944
Panel D: All stocks, bandwidth is 300, all controls included (baseline)										
τ_{it}	-0.010***	-0.010***	-0.007***	-0.004***	-0.001	-0.006***	-0.009***	-0.006***	-0.004***	-0.002***
	(-4.84)	(-6.20)	(-5.85)	(-4.03)	(-0.98)	(-3.67)	(-7.40)	(-5.99)	(-4.69)	(-2.90)
Adjusted R^2, %	4.07	7.92	9.73	8.63	7.44	2.94	6.10	4.50	3.15	2.14
Observations	1921	1811	1713	1628	1548	2285	2159	2033	1930	1833
Panel E: All stocks, bandwidth is 300, all controls and year fixed effects included										
τ_{it}	-0.005**	-0.003*	-0.002*	-0.001	0.000	-0.002	-0.003**	-0.000	-0.000	-0.000
	(-2.24)	(-1.57)	(-1.41)	(-0.48)	(0.38)	(-1.10)	(-2.09)	(-0.46)	(-0.08)	(-0.60)
Within R^2, %	3.29	3.57	2.66	2.58	3.29	0.79	0.34	0.14	0.05	0.02
Observations	1921	1811	1713	1628	1548	2285	2159	2033	1930	1833
Panel F: Flippers excluded, bandwidth is 300, all controls and year fixed effects included										
τ_{it}	-0.019***	-0.019***	-0.017***	-0.013***	-0.011***	-0.017***	-0.018***	-0.014***	-0.012***	-0.010***
	(-7.70)	(-9.88)	(-10.14)	(-8.50)	(-7.33)	(-7.15)	(-9.42)	(-8.63)	(-7.78)	(-6.84)
Within R^2, %	6.74	12.51	14.80	13.34	12.39	8.14	14.70	14.83	15.25	14.39
Observations	1572	1301	1134	1021	928	1719	1367	1130	969	865

This table reports the results of the reduced form regression for the pre-banding sample period (1998-2006). The dependent variables are excess long-run returns of stock i from September in year t over the respective horizon. Panel A uses all stocks in the band of 100 around the cutoff, Panel B uses same band but excludes stocks moving back to the other index in the relevant horizon, Panel C uses all stocks in the band of 300 around the cutoff, Panel D adds all controls to the specification and sample of Panel C (our baseline), Panel E adds year fixed effects, Panel F adds year fixed effects and excludes stocks moving back to the other index in the relevant horizon. We include only stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are based on a one-sided test and marked as: $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 17: Reduced form results for 2007-2018

	Long-run excess return									
	Additions to Russell 2000					*Deletions from Russell 2000*				
	12-month	24-month	36-month	48-month	60-month	12-month	24-month	36-month	48-month	60-month
Panel A: All stocks, bandwidth is 100, only $logMV$ included										
τ_{it}	-0.007*	-0.001	0.000	0.003	0.001	-0.001	0.001	0.001	0.001	-0.003
	(-1.70)	(-0.59)	(0.04)	(0.66)	(0.20)	(-0.75)	(0.76)	(0.90)	(0.44)	(0.66)
Adjusted R^2, %	3.20	3.74	3.35	13.79	14.03	1.38	6.32	1.73	1.93	4.35
Observations	330	290	284	238	150	604	535	524	457	312
Panel B: Flippers excluded, bandwidth is 100, only $logMV$ included										
τ_{it}	-0.007**	-0.007***	-0.005**	-0.005**	-0.009***	-0.008***	-0.009***	-0.008***	-0.011***	-0.010***
	(-2.06)	(-2.39)	(-2.09)	(-1.79)	(-2.59)	(-3.12)	(-4.08)	(-3.87)	(-5.28)	(-4.76)
Adjusted R^2, %	1.33	4.28	7.00	8.74	13.91	2.08	4.82	5.74	13.89	17.33
Observations	216	164	116	75	53	420	310	233	172	130
Panel C: All stocks, bandwidth is 300, only $logMV$ included										
τ_{it}	-0.009**	-0.004**	-0.000	-0.001	-0.001	-0.005***	-0.003**	-0.001	-0.002**	-0.003***
	(-2.87)	(-1.82)	(-0.05)	(-0.53)	(-0.73)	(-2.65)	(-1.77)	(-1.21)	(-1.66)	(-2.53)
Adjusted R^2, %	4.25	4.79	5.63	11.21	12.75	2.19	2.61	1.49	3.11	4.27
Observations	1275	1120	960	809	649	1973	1690	1456	1237	1059
Panel D: All stocks, bandwidth is 300, all controls included (baseline)										
τ_{it}	-0.010***	-0.003*	-0.000	-0.000	-0.001	-0.004**	-0.003**	-0.002**	-0.002**	-0.003***
	(-2.87)	(-1.41)	(-0.09)	(-0.29)	(-0.80)	(-2.13)	(-1.97)	(-1.74)	(-2.07)	(-2.79)
Adjusted R^2, %	4.88	4.92	6.40	11.31	13.66	4.63	4.31	3.36	4.28	6.85
Observations	1175	1031	888	752	607	1865	1592	1369	1165	1004
Panel E: All stocks, bandwidth is 300, all controls and year fixed effects included										
τ_{it}	-0.003	0.001	0.002	0.003	0.001	-0.001	0.001	0.001	0.000	0.000
	(-0.85)	(0.32)	(0.98)	(1.37)	(0.68)	(-0.59)	(0.59)	(0.32)	(0.03)	(0.05)
Within R^2, %	0.42	0.54	0.19	0.81	1.63	0.62	1.14	0.86	0.98	1.72
Observations	1175	1031	888	752	607	1865	1592	1369	1165	1004
Panel F: Flippers excluded, bandwidth is 300, all controls and year fixed effects included										
τ_{it}	-0.009**	-0.008***	-0.006**	-0.006**	-0.007***	-0.012***	-0.012***	-0.010***	-0.011***	-0.009***
	(-2.27)	(-2.55)	(-2.13)	(-2.00)	(-2.49)	(-4.48)	(-5.00)	(-4.58)	(-5.15)	(-4.10)
Within R^2, %	2.17	1.28	3.24	1.84	2.65	1.40	3.67	6.63	8.53	8.95
Observations	908	684	519	374	277	1432	1076	829	670	539

This table reports the results of the reduced form regression for the post-banding sample period (2007-2018). The dependent variables are excess long-run returns of stock i from September in year t over the respective horizon. Panel A uses all stocks in the band of 100 around the cutoff, Panel B uses same band but excludes stocks moving back to the other index in the relevant horizon, Panel C uses all stocks in the band of 300 around the cutoff, Panel D adds all controls to the specification and sample of Panel C (our baseline), Panel E adds year fixed effects, Panel F adds year fixed effects and excludes stocks moving back to the other index in the relevant horizon. We include only stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are based on a one-sided test and marked as: $^{*}p<0.05$*;* $^{**}p<0.01$*;* $^{***}p<0.001$*.*

A.17 Value and Growth Funds' Rebalancing

Table 18: Who rebalances additions and deletions? By style

Summary of separate regressions on additions, deletions, and styles

Ownership by Benchmark Group and Type, %

Type	Active									Passive								
Benchmark	Russell 1000			Russell MidCap			Russell 2000			Russell 1000			Russell MidCap			Russell 2000		
Style	Blend	Value	Growth	Blend	Value	Growth	Blend	Value	Growth	Blend	Value	Growth	Blend	Value	Growth	Blend	Value	Growth
Panel A: Pre-banding sample (1998-2006)																		
D_{it}^{R1000}	0.00	-0.01	-0.02	0.02	-0.02	-0.27**	-0.36***	-0.21**	-0.17**	0.01***	-0.00	-0.001	0.01***	-0.00**	-0.00	-0.17***	-0.02**	-0.01
	(0.70)	(-0.33)	(-0.90)	(0.90)	(-0.35)	(-3.21)	(-3.67)	(-2.96)	(-2.74)	(10.17)	(-0.37)	(-1.23)	(17.97)	(-3.28)	(-1.24)	(-12.55)	(-2.75)	(-1.47)
$D_{it}^{R1000Value}$		0.05			0.04			-0.01			0.02***			0.01***			-0.06***	
		(1.86)			(0.71)			(-0.12)			(15.40)			(10.09)			(-10.42)	
$D_{it}^{R1000Growth}$			0.13***			0.65***			0.08			0.02***			0.01***			-0.04***
			(4.85)			(5.62)			(0.84)			(13.46)			(12.87)			(-9.45)
D_{it}^{R2000}	-0.00	-0.06	0.00	-0.01	-0.02	-0.14*	0.15	0.08	-0.26***	-0.01***	-0.02***	-0.01***	-0.01***	-0.01***	-0.01***	0.32***	0.00	-0.00
	(-0.48)	(-1.24)	(0.08)	(-0.32)	(-0.22)	(-2.18)	(1.90)	(0.76)	(-4.76)	(-5.59)	(-12.91)	(-7.30)	(-17.44)	(-5.40)	(-8.26)	(23.76)	(0.81)	(-0.29)
$D_{it}^{R2000Value}$		0.14*			-0.01			0.10			-0.01***			-0.00***			0.12***	
		(2.22)			(-0.10)			(0.76)			(-9.58)			(-5.71)			(15.49)	
$D_{it}^{R2000Growth}$			0.01			0.14			0.20*			-0.01***			-0.01***			0.11***
			(0.25)			(1.17)			(2.47)			(-6.63)			(-6.72)			(12.40)
Number of funds (2006)	19	98	107	30	36	67	109	60	85	2	3	5	1	1	1	8	2	2
AUM share (2006), %	1.0	10.8	4.4	1.3	2.3	4.1	4.1	1.6	2.6	0.0	0.1	0.1	0.0	0.0	0.0	0.1	0.0	0.0
Panel B: Post-banding sample (2007-2018)																		
D_{it}^{R1000}	0.03**	0.02	-0.07*	0.17***	-0.03	-0.08	-0.39***	-0.50***	-0.20	0.06***	-0.01***	-0.01***	0.18***	-0.00	-0.00***	-1.85***	-0.08***	-0.20***
	(2.64)	(0.96)	(-2.17)	(4.01)	(-0.58)	(-0.78)	(-4.14)	(-7.43)	(-1.71)	(25.86)	(-4.63)	(-6.54)	(44.60)	(-1.09)	(-4.34)	(-56.12)	(-8.75)	(-9.40)
$D_{it}^{R1000Value}$		0.10*			0.19			-0.22*			0.16***			0.12***			-0.26***	
		(2.35)			(1.86)			(-2.25)			(14.66)			(15.79)			(-11.63)	
$D_{it}^{R1000Growth}$			0.42***			1.02***			-0.46**			0.20***			0.15***			-0.30***
			(7.60)			(6.82)			(-3.10)			(33.81)			(33.15)			(-12.79)
D_{it}^{R2000}	-0.04*	-0.06	0.03	-0.09*	-0.20	-0.14	-0.22	-0.27	-0.21	-0.05***	-0.08***	-0.014**	-0.18***	-0.07***	-0.02***	1.73***	0.00	0.00*
	(-2.57)	(-1.38)	(0.61)	(-2.49)	(-0.98)	(-1.20)	(-1.55)	(-1.15)	(-1.87)	(-11.94)	(-5.94)	(-2.80)	(-29.22)	(-8.25)	(-4.86)	(32.43)	(0.22)	(2.22)
$D_{it}^{R2000Value}$		-0.13			-0.08			0.53			-0.10***			-0.06***			0.54***	
		(-1.89)			(-0.34)			(1.93)			(-6.76)			(-6.51)			(23.84)	
$D_{it}^{R2000Growth}$			-0.08			-0.08			0.17			-0.09***			-0.06***			0.42***
			(-1.23)			(-0.44)			(0.87)			(-7.62)			(-7.69)			(13.64)
Number of funds (2013)	30	142	154	38	61	82	123	87	95	5	10	8	5	1	2	17	3	2
AUM share (2013), %	1.2	6.5	5.8	0.8	2.7	2.9	2.6	1.7	2.1	0.3	0.6	0.6	0.3	0.1	0.1	0.9	0.1	0.1

This table reports the discontinuities in rebalancing for the pre- and post-banding sample periods. Estimation is performed at fund group level (by benchmark, style, and type). The coefficients come from separate regressions: on stocks that were in the Russell 1000 (additions) or in the Russell 2000 (deletions) in the previous year. Band width is 500 stocks around the cutoffs (rectangular kernel). The dependent variables are ownership shares in stock i as of September in year t of the respective investor group. All regressions include one-year lagged ownership, log total market value, Russell float factor, 5-year monthly rolling β^{CAPM}, 1-year monthly rolling average bid-ask percentage spread, and stock's return over year $t-1$. t-statistics based on HAC-robust standard errors with clusters at a stock level are in parentheses. Significance levels are marked as: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$.

A.18 Optimized Sampling in Prospectus

Figure 8: An extract from the prospectus of Fidelity's ZERO Large Cap index fund.

Principal Investment Strategies

- Normally investing at least 80% of assets in common stocks of large capitalization companies included in the Fidelity U.S. Large Cap Index℠, which is a float-adjusted market capitalization-weighted index designed to reflect the performance of U.S. large capitalization stocks. Large capitalization stocks are considered to be stocks of the largest 500 U.S. companies based on float-adjusted market capitalization.
- Using statistical sampling techniques based on such factors as capitalization, industry exposures, dividend yield, price/earnings (P/E) ratio, price/book (P/B) ratio, and earnings growth to attempt to replicate the returns of the Fidelity U.S. Large Cap Index℠ using a smaller number of securities.
- Lending securities to earn income for the fund.