ON INDEX INVESTING

Jeffrey L. Coles
David Eccles School of Business
University of Utah
jeff.coles@eccles.utah.edu

Davidson Heath
David Eccles School of Business
University of Utah
davidson.heath@eccles.utah.edu

Matthew C. Ringgenberg
David Eccles School of Business
University of Utah
matthew.ringgenberg@eccles.utah.edu

First Draft: March 15, 2017
This Draft: October 20, 2017*

*Comments welcome. The authors thank Todd Gormley, participants at the Olin Business School Wealth and Asset Management Research Conference, and seminar participants at the University of Nebraska. We also thank Russell for providing the index data. All errors are our own.

ON INDEX INVESTING

First Draft: March 15, 2017
This Draft: October 20, 2017

ABSTRACT

We quantify the impact of index investing on stock prices. Using a regression discontinuity analysis around yearly Russell index reconstitutions, we find that index investing introduces noise into stock prices, but does not impact long-term price efficiency or trading by arbitrageurs. Stocks with more index investors have prices that deviate more from a random walk and exhibit higher correlations with index price movements. However, these stocks have no difference in turnover, trading volume, or earnings response coefficients. In other words, index investing introduces noise into prices, but it does not impact the ability of arbitrageurs to impound information into prices.

Keywords: active management, index investing, market efficiency, passive management.

JEL Classification Numbers: G12, G14

Not literally everyone can index! Some people need to build and manage and run businesses, and some other people need to allocate capital to those businesses....If everyone indexed, nothing would get built.

-Matt Levine, Bloomberg

I. Introduction

There is now overwhelming evidence that active managers do not outperform the market after fees.[1] As a consequence, the amount of capital devoted to index investing has grown by more than $4 trillion dollars over the last 40 years (Bogle (2016)). But this increase in passive index investing is not without controversy. Passive investors are necessarily free-riding on the research and effort exerted by active managers. This suggests a trade-off: while passive management allows investors to earn index returns for low fees, some amount of active managers must exist to ensure that prices correctly reflect fundamental value. Put differently, not everyone can index, someone has to be *active*. The question is, how many active managers are enough to ensure that prices correctly reflect fundamental value?

In this paper we explore this trade-off using Russell Index reconstitutions as a source of exogenous variation in index investing. Over our 23-year sample period, we find that increased index investing leads to significant changes in stock prices, however, it does not alter the ability of arbitrageurs to impound information into prices. We start by examining the direct effects of index membership; after a change in Russell Index assignment, we find strong evidence of a shift in the composition of investors. When a stock is added to the Russell 2000 index, we find that ownership by passive funds (index funds and closet indexers) increases by approximately 0.4% of market cap during the 1993 to 2006 period, and approximately 0.8% of market cap during the period from 2007 to 2016. We also find

[1] There is a large literature examining the performance of active investors who attempt to "beat" the market. As far back as Cowles (1933), it has been shown that experts generally cannot pick winners and losers in the stock market. More recently, a number of papers have shown that active managers do not outperform after accounting for fees. See, for example, Jensen (1967), Carhart (1997), Sharpe (1991), French (2008), and Fama and French (2010).

that ownership by active mutual funds falls by a similar magnitude. In other words, index investors appear to *replace* active managers as owners in treated stocks.

Next, we examine whether index investing impacts asset prices. Consistent with the large literature on downward sloping demand curves (e.g., Shleifer (1986), Bagwell (1992), Chang, Hong, and Liskovich (2015), etc.) we find strong evidence of price effects from index re-balancing. On average, stocks experience a permanent price increase of approximately 0.5% following their addition to the Russell 2000 index. The results imply that our sample of U.S. equities has a price elasticity of demand of approximately -0.26. We also find that index investing is associated with a sharp increase in volatility, consistent with the theoretical predictions in Basak and Pavlova (2013) and the empirical findings in Ben-David, Franzoni, and Moussawi (2016). Moreover, we find strong evidence that a change in investor composition, from active investors to passive investors, leads to a degradation in weak-form price efficiency. Using variance ratio tests (Campbell and Mankiw (1987), Lo and MacKinlay (1988), Campbell, Lo, and MacKinlay (1997)), we find that increased index investing is associated with prices that significantly deviate from a random walk model. Following their addition to the Russell 2000 index, stocks experience an increase in their variance ratios of approximately 30%.

Theoretically, there are many ways in which the composition of investors might impact prices. For example, if active managers have limited risk-bearing capacity, then a large shift in the allocation of capital, from active funds to passive funds, might leave active funds with too little capital to trade on all mis-pricings. However, while index investing has increased dramatically over the last 40 years, U.S. mutual funds still have over $16 trillion in assets, most of which are allocated to active strategies (Investment Company Institute (2017)). As such, it seems unlikely that active funds have too little capital to correct all apparent mis-pricings. Another possibility it that an increase in passive ownership makes it more costly for active funds to trade on mis-pricings. For example, if an increase in passive ownership leads to a significant reduction in public float, then active investors might incur higher price impact to trade on mis-pricings. As a result, they might be less likely to trade and prices might be less efficient.[2] In a related point, it is also possible that increased index investing changes the incentives to acquire information (e.g., Brown and Davies (2017), Bond and Garcia (2017)) which could then impact price efficiency in the marketplace.

To explore the economic channel through which index investing impacts prices, we examine a variety of dependent variables proposed by theory. We first examine return volatilities and correlations. We find strong evidence that an increase in index investing leads to increased volatility at the stock level (consistent with the results in Ben-David et al. (2016)). Moreover, we also find that increased index investing generates higher correlations between stocks in the index. While increases in return volatilities and correlations may not necessarily be bad, it is possible that these effects could deter arbitrage. Accordingly, we next examine measures of trading and semi-strong form efficiency. However, we find no evidence that index investing deters arbitrage. When we examine turnover and trading volume, we find no effect. Interestingly, this result implies that index owners do not actually trade less than active owners, and as such, they may not be truly "passive." Similarly, when we examine earnings response coefficients and the Hou and Moskowitz (2005) measure of semi-strong form price efficiency, we find no effect. In other words, index investing does impact the price process, but it does not appear to significantly alter the ability of arbitrageurs to trade or impound information into prices.

Of course, one of the key challenges to understanding the impact of passive investing is that the quantity of passive capital allocated to a stock is not random. Several existing papers document a correlation between the composition of investors and price efficiency. For example, Boehmer and Kelley (2009) find that institutional ownership is associated with improved price efficiency. More recently, Israeli, Lee, and Sridharan (2016) and Glosten, Nallareddy, and Zhou (2016) both examine the relation between informational efficiency and the amount of capital invested in exchange traded funds (ETFs). Israeli et al. (2016)

[2] For example, Schnitzler (2016) finds that the price impact from adding a stock to the S&P500 is lessened when the stock has more supply (i.e., its free float is larger).

uses ordinary least squares (OLS) panel regressions and find that higher ETF ownership is associated with worse efficiency, as measured by future earnings response coefficients. Similarly, Glosten et al. (2016) finds that changes in quarterly ETF ownership are associated with a change in a firm's information environment. Specifically, they find that an increase in ETF ownership is associated with improved incorporation of accounting information into prices.[3] Our paper differs in that we focus on index investing, in general, instead of the amount of capital allocated to ETFs. Moreover, to the best of our knowledge, our paper is the first to use exogenous variation in index ownership in a regression discontinuity design (RDD) to identify the impact of passive investing on price efficiency.

Although our results suggest that index investing does change the price process in our sample of U.S. equities, we are careful to point out that our paper does *not* say that index investing decreases social welfare. First, the RDD methodology we employ estimates the *local average treatment effect* (LATE). In our sample, the average change in index investing amounts to approximately 0.3% of market capitalization pre-banding and 0.7% post-banding. Our estimates are not able to establish whether larger changes (or substantially smaller changes) in index investing would have similar effects to the results we document. Moreover, our estimate of the impact on price efficiency is for stocks assigned to the top of the Russell 2000 Index; as such, it is not clear if our estimate would apply to stocks that are substantially different from the largest stocks in the Russell 2000. Finally, we note that there is strong evidence that investors earn higher returns, after fees, by investing in passive index funds. As such, investors receive some benefit from the existence of passive index funds.

Overall, our paper makes several contributions. First, our paper documents clear evidence that index re-balancing *causes* index investors to *replace* active managers as owners in a stock. On average, index re-balancing significantly changes the composition of investors in

[3]Israeli et al. (2016) and Glosten et al. (2016) both focus on ETF ownership and they look at the incorporation of accounting information whereas we look at price efficiency using variance ratio tests. Moreover, our paper is the first to use the Russell Reconstitutions in a regression discontinuity design to account for the endogenous relation between index investing and firm characteristics. In Section IV, below, we present results from a Durban-Watson-Hausman chi squared test which show strong evidence that passive investing and price efficiency are endogenously determined.

a firm. Second, we confirm the findings in the large literature on downward sloping demand curves for stocks. Third, consistent with theoretical predictions (e.g., Basak and Pavlova (2013), we find that increased passive ownership is associated with higher volatility. Fourth, we find that passive index investing *changes* the correlation structure of asset prices and *causes* prices to deviate from a random walk. Finally, we show that index investing does not change trading behavior by informed arbitrageurs. Overall, our findings suggest that index investing changes the price process but it does not impact the ability of arbitrageurs to impound information into prices.

The rest of the paper proceeds as follows. Section II discusses the existing literature and motivates our empirical tests. Section III describes our sample and outlines our identification strategy. Section IV displays our main results. Section V discusses several possible interpretations of our results. Section VI concludes.

II. Background

In this section, we briefly discuss existing work on index investing and its impact on market outcomes. We then formalize the economic mechanisms introduced in the beginning of the paper.

A. Related Literature

While index investing has been around for nearly 40 years (Bogle (2016)), there are relatively few papers examining the economic impact of passive index investing. Theoretically, there are multiple ways in which a shift in the composition of investors, from active to passive, could impact asset prices. Stein (1987) models the impact of introducing speculators to commodity markets. In his model, new speculators can change the information content of prices in a manner that generates negative externalities for other investors, ultimately leading to welfare reductions. Goldstein and Yang (2017) also examine the impact of a change in investor composition in the commodity market, and they show that an increase in index investing can change price efficiency and risk sharing. Subrahmanyam (1991) develops a model of strategic trading by liquidity traders who are allowed to trade in a "basket" of securities (i.e., index securities). The model derives conditions under which index investing can lead to increases, or decreases, in price informativeness depending on the liquidity trading that occurs.

More recently, Basak and Pavlova (2013) specifically model passive investors who are incentivized to track an index. Their model predicts that passive index investors will generate significant price pressure in index assets. Moreover, their model also predicts an increase in volatility as a result of index investing. Cong and Xu (2016) develop a model in which investors can trade composite securities, like index ETFs. In their model, the creation of ETFs can lead to lower asset-specific information in security prices, but higher systematic information in prices. Baruch and Zhang (2017) examine the impact of index investing in a rational expectations framework and shows that it can change the correlation structures of prices and generate an increase in idiosyncratic volatility. Brown and Davies (2017) show that an increase in passive investing may decrease the incentives of active managers to exert effort. Similarly, Bond and Garcia (2017) develop a model based on the Grossman-Stiglitz-Hellwig framework. They show that an increase in index investing reduces noise trading in individual assets, making prices more informative. This reduces the incentive for active investors to acquire information about those assets, so they shift their attention towards acquiring systematic information. Overall, in their model, an increase in index investing may increase price efficiency at the stock level, but reduce overall welfare by distorting risk taking.

Empirically, several recent papers have investigated the impact of investing in ETFs. Perhaps most closely related to our study, Ben-David et al. (2016) examine firm volatility using Russell Index reconstitutions to generate variation in index ownership. They find that an increase in index investing is associated with higher volatility. Israeli et al. (2016)

examines the relation between informational efficiency and the amount of capital invested in exchange traded funds (ETFs). Using ordinary least squares (OLS) panel regressions, they find that higher ETF ownership is associated with worse efficiency, as measured by future earnings response coefficients. They also find decreases in liquidity for stocks with higher ETF ownership. In contrast, Glosten et al. (2016) finds that changes in quarterly ETF ownership are associated with a change in a firm's information environment. Specifically, they find that an increase in ETF ownership is associated with improved incorporation of accounting information into prices. Importantly, our paper differs from these on several dimensions. First, we examine the impact of passive investing, instead of ETF investing. While the two topics are closely related, some ETFs are active, and a number of mutual funds are passive index funds (or quasi-index funds). Moreover, while index ETF managers do not try to actively beat the market, some ETF *investors* are likely to be active investors. Israeli et al. (2016) and Glosten et al. (2016) both look at variation in ETF ownership, some of which is driven by trading from *active* investors. Our experimental setting is different: because we examine changes in ownership that result purely from Russell reconstitutions, our results are driven by variation in passive investment. As such, our paper investigates a different question.

In addition, Israeli et al. (2016) and Glosten et al. (2016) focus primarily on the incorporation of accounting information into stock prices. In contrast, we examine variance ratio tests, which can be viewed as a measure of weak form efficiency. Moreover, to the best of our knowledge, our paper is the first to use exogenous variation in passive ownership in a regression discontinuity design (RDD) to identify the impact of index investing on price efficiency. In Section IV, below, we present results from a Durban-Watson-Hausman chi squared test which shows strong evidence that index investing and price efficiency are endogenously determined. Accordingly, we believe our setting is an ideal laboratory for examining the impact of index investing.

Our empirical design makes use of the mechanical rules for Russell index assignment, as have several other recent papers.[4] Mullins (2014) uses Russell reconstitutions to examine the impact of institutional ownership on corporate governance. Chang et al. (2015) use Russell reconstitutions to carefully measure the price effects from index additions and deletions, consistent with the existing literature on downward sloping demand curves for stocks. Appel, Gormley, and Keim (2016) find that higher passive investment was associated with better governance and changes in the type of campaigns launched by activist investors. Crane, Michenaud, and Weston (2016) find that higher passive investment was associated with higher payout to investors. Finally, Schmidt and Fahlenbrach (2017) find that higher passive investment was followed by increases in CEO power and worse M&A outcomes.

III. Data and Research Design

To examine the impact of passive index investing on price efficiency, we use variation in Russell Index membership as an exogenous shock to passive index investing. Specifically, we combine daily stock data from the Center for Research in Security Prices (CRSP) with firm data from Compustat, ownership data from Thomson Reuters S12, and information on Russell index membership.

A. Index Assignment

In June of each, year Russell Investments reconstitutes their popular Russell 1000 and 2000 indexes. To determine index membership, Russell ranks all U.S. common stocks by their market capitalization as of the last business day in May. The list of new index memberships is released during June and the indexes are reconstituted at the close of the last business day in June.

Prior to 2007 index assignment followed a simple threshold rule: stocks ranked from

[4]While we are not the first to use index reconstitutions as a source of variation, we note that we use a different sample period than the existing literature. Because Russell changed their methodology for index reconstitutions in 2007, our sample period requires us to use a different methodology to identify changes in index membership. We discuss this in greater detail in Section III below.

1-1000 were assigned to the Russell 1000 while stocks ranked from 1001-3000 were assigned to the Russell 2000. In the years from 1993 to 2006, we make use of a simple regression discontinuity design (RDD) that uses index membership as an instrument for passive fund ownership.

Starting in June 2007 Russell implemented a new assignment regime ("banding") which replaced the simple threshold rule. The banding regime eliminated the discontinuity of index assignment at the 1000-rank threshold, which makes the simple RDD approach infeasible. However, we document that in addition to eliminating the discontinuity at the threshold, the banding regime replaced it with two *new* discontinuities conditional on a stock's previous index assignment. We develop a new research design which allows us to make use of Russell assignment post-banding as an instrument for passive ownership, and we show that it works in the sense that 1) it is well-motivated by the mechanical post-banding assignment rules, 2) it produces clear and strong variation in both levels and changes in passive ownership, and 3) it passes a battery of balance tests suggesting that treated and control stocks were ex ante similar.

B. Index Weights and Rankings

Each month, Russell computes a weight for each stock in its assigned index based on the stock's float-adjusted market capitalization. The float adjustment removes non actively traded holdings such as officers and directors' holdings, large block holdings, holdings by related firms, ESOP shares, and government holdings.[5] The float adjustment thus down-weights relatively illiquid, cross-owned, and inside-owned stocks. Russell then re-sorts the stocks within each index on their weights, yielding the official Russell rankings. Because the index weights, float adjustments, and Russell rankings are endogenous and potentially correlated with *ex ante* stock characteristics like liquidity and fund ownership, we do not make use of them in our research design.

[5] `http://www.ftse.com/products/downloads/Russell-US-indexes.pdf`

C. Research Design Pre-Banding

In the pre-banding regime (pre-2007), close to the yearly index threshold there was a strong discontinuity in index membership (See Figure 1 Panel A). Index membership was determined by a single mechanical threshold, and affected by changes in market capitalization for *all* firms near the threshold. Thus, the first-best research design would be to estimate a sharp regression discontinuity design (RDD) using the true Russell rankings on end-of-May market cap.

Unfortunately, Russell does not disclose its end-of-May market cap rankings that determined index membership, and because they use their own proprietary methods to compute the rankings we do not observe the true forcing variable that determines yearly index assignment. We compute a proxy market capitalization variable, CAP, for each stock at the end of May each year using CRSP and Compustat data following Chang et al. (2015). We then sort all Russell 3000 stocks in each year to produce a proxy ranking $CAPrank$ and impute the yearly index cutoff between the Russell 1000 and 2000 indexes.

Our proxy ranking $CAPrank$ is very close to the true Russell rankings. Pre-banding, from 1993-2006 we correctly impute index assignment for 99.2% of Russell 3000 stocks and 94.4% of stocks in the RDD sample (within a +/-100 rank window of the threshold).

Since our proxy ranking does not predict treatment status perfectly, it would seem natural to estimate a fuzzy RDD. However, a fuzzy RDD requires that the likelihood of treatment jumps at the threshold by less than 100% because of noncompliance. By contrast, in this setting treatment status is known to be *perfectly* discontinuous at the threshold. As a result, using predicted treatment status would understate the true effects of index membership.[6]

Thus, our research design pre-banding uses stocks' actual treatment status in a narrow window around the yearly index threshold, and the RDD control function uses our proxy for

[6]Another way to see this is to note that what we want is not the effects of intention to treat but the effects of treatment on the treated, which requires that we use actual and not predicted treatment status.

the true forcing variable:

$$\Delta FundOwn_{it} = \beta R2000_{it} + \lambda R2000_{i,t-1} + f(CAPrank_{it}) + \gamma_t + \epsilon_{it}, \tag{1}$$

where $CAPrank$ is our proxy ranking and $R2000$ is a dummy variable that equals one if a stock was assigned to the Russell 2000 and zero if it was assigned to the Russell 1000. We also control for lagged index assignment $R2000_{i,t-1}$ because empirically we find, regardless of bandwidth around the threshold, that stocks assigned to the Russell 2000 were slightly more likely to be Russell 2000 members the previous year. Our results are all similar if we omit this control.

The main concern with this design is that the control function f uses the proxy ranking instead of the unobserved true ranking:

$$rankerror_{it} = CAPrank_{it} - Russellranking_{it}, \tag{2}$$

which introduces errors in f. If $rankerror$ is uncorrelated with $\Delta FundOwn$, we still recover an unbiased estimate of the treatment effect. However, if this is not the case then $rankerror$ might bias the control function and produce a biased estimate of the jump in institutional ownership at the threshold.

To examine this possibility, in the Appendix we perform extensive checks as to whether our results are affected by varying the construction of $CAPrank$, the window around the threshold, the control function, the weighting kernel, and using predicted treatment status instead of true treatment status as the independent variable. We find that all these variations produce similar estimates. In sum, we find no reason to suspect that $rankerror$ biases our results.

D. Research Design Post-Banding

Starting in June 2007, Russell implemented a new assignment regime ("banding"). After initially sorting stocks by their market cap, Russell computes an upper and lower band around the index threshold. The width of each band equals 2.5% of the total May market cap of the Russell 3000. Stocks within the bands do not switch their index assignment from last year. For example, in the pre-banding regime a stock ranked above the threshold but below the upper band would have been assigned to the Russell 1000. Post-banding, if that stock was in the Russell 2000 last year, it will stay in the Russell 2000 in the coming year.

Banding was intended to reduce the uncertainty in index membership. Russell's data suggest that it was successful; in the first seven years of banding (2007-2014), the total number of stocks added and deleted from the Russell 1000 and 2000 fell by 45%, down to 430 compared to 872 in the last seven years prior to banding (2000-2006).[7]

Figure 1 Panel B plots index assignments and the imputed index threshold and upper and lower bands for 2007, the first post-banding year. We see that close to the index threshold there was no discontinuity in index membership. Thus, an RDD for index assignment around the index threshold is no longer feasible because banding eliminated any variation in index membership for stocks near the threshold. However, Figure 1 Panel B also makes clear that banding replaced a single threshold for index membership with two thresholds for index *switching*. In the post-banding regime, there are clear discontinuities at the upper and lower bands for whether stocks switched indexes. This observation drives our empirical strategy in the post-banding period.

Consider a stock that was assigned to the Russell 2000 in 2006 (i.e. from from June 2006 to May 2007) *and* ranked near the upper band at the end of May 2007. This stock's 2007 index assignment depends on whether it ranked just above the upper band – in which case it would switch to the Russell 1000, or just below the upper band – in which case it would stay in the Russell 2000.

[7] `http://www.ftserussell.com/blog/russell-2000-recon-banding-results-lower-turnover`

The stock's position relative to the upper band depends on five parameters:

1. the stock's ranking in the Russell 3000, which is sensitive to small fluctuations in the market cap of both the focal stock and all the neighboring stocks in the ranking

2. the market cap of the rank-1000 stock, which determines the Russell 1000/2000 threshold

3. the total market cap of the Russell 3000 as calculated by Russell, which determines the width of the bands

4. the cumulative market cap as calculated by Russell of all the stocks ranked above the focal stock, which determines where the stock sits relative to the bands

5. whether the stock was in the Russell 1000 or 2000 last year, which was set by Russell 12 months prior.

The first four parameters are difficult to predict ex ante (and indeed, Russell does not disclose them ex post). All five parameters are either difficult or impossible to manipulate. This observation suggests that for stocks that were "potential switchers" (i.e. assigned to the opposite Russell index in the previous year) and ranked close to the relevant band this year were as good as randomly assigned. Indeed, post-banding, a stocks' index assignment conditional on being close to the upper or lower band, ex post, was arguably even harder to predict or manipulate *ex ante* than in the pre-banding regime.

This argument drives our post-banding research design. Specifically, we measure the effect of index assignment:

$$\Delta FundOwn_{it} = \beta R2000_{it} + I\{Upperband_{it}\} \times f(CAPrank_{it}) + I\{Upperband_{it}\} \times \gamma_t + \epsilon_{it}, \quad (3)$$

where the sample consists *only* of potential switcher stocks in windows around the upper and lower bands. That is, stock i is in the sample only if $R2000_{i,t-1} = 1$ and it is within

+/-100 ranks of the upper band in year t, or if $R2000_{i,t-1} = 0$ and it is within +/-100 ranks of the lower band in year t. The control function $f(CAPrank_{it})$ is fit separately across the upper and lower bands. The post-banding specification does not include a control for lagged index assignment because lagged index assignment is co-linear with the band-by-year fixed effects.

Our full procedure to construct the post-banding sample is as follows:

1. Rank all candidate stocks on their unadjusted market cap as before.

2. Impute the Russell index threshold as before.

3. Compute the running market cap up to each stock's ranking i.e. for stock #1 this is stock #1's market cap. For stock #3 this is the total market cap of stocks 1, 2, and 3. Call this function $RMC(n)$ where n is the stock's ranking.

4. Compute the total market cap of all candidate stocks (ranks 1-3000). Call this number $TMC = RMC(3000)$.

5. Compute an upper band, which is the ranking for which the running market cap is just below the running market cap at rank 1000 minus 2.5% of the total market cap for all stocks. Thus, the upper band in a given year is the largest n such that $RMC(n) < RMC(1000) - 0.025 \times TMC$.

6. Compute an analogous lower band, which is the ranking for which the running market cap is just *above* the running market cap at rank 1000 *plus* 2.5% of the total market cap for all stocks. Thus, the lower band in a given year is the smallest n such that $RMC(n) > RMC(1000) + 0.025 \times TMC$.

7. Stocks within the bands keep their index assignment from last year.

As in the pre-banding period, our proxy ranking $CAPrank$ is very close to the true Russell rankings. Post-banding, from 2007-2016 we correctly impute index assignment for

99.2% of Russell 3000 stocks and 96.3% of stocks in the RDD sample (within a +/-100 rank window of the upper and lower bands).

E. Data

Russell monthly index membership data come directly from Russell. Data on sample stocks is from CRSP and the merged CRSP-Compustat database. We use the most recent data for each firm and stock as of the last business day in May each year.

Institutional ownership data comes from the Thomson Reuters S12 database of mutual fund holdings. We compute the ownership of each sample stock by every fund in December of each year: prior to 2004 funds were only required to report twice a year, in June and December. We use the number of sole-voting shares held where available, otherwise the total shares held.

We classify mutual funds as active or passive using the approach of Cremers and Petajisto (2009). Specifically, in each December prior to index assignment, for each fund in the data we compute its "active share" relative to the index weights of the Russell 2000. Funds with an active share less than 0.6 (which Cremers and Petajisto suggest as a rough break-point for explicit or closet indexing) are classified as passive. Our results on fund ownership are effectively identical if we use a different active-share breakpoint such as 0.4 or 0.8.

Our measures of institutional ownership for each stock i as of December in year t are defined below, and are all expressed as a percent of the stock's total market capitalization.

- $FundOwn_{it}$: Total ownership by all mutual funds in the S12 database.
- $FundOwn_{t}^{PASSIVE}$: Total ownership by mutual funds with active share less than 0.6 relative to the Russell 2000.
- $FundOwn_{it}^{ACTIVE}$: Total ownership by mutual funds with active share greater than 0.6 relative to the Russell 2000.

Our pre-banding sample consists of all stocks in a +/-100 rank window (using our proxy ranking on May market capitalization) around the Russell index threshold, each year from 1993 to 2006. Our post-banding sample consists of all stocks in a +/-100 rank window around the upper and lower bands, each year from 2007 to 2016, that were potential switchers (i.e. stocks near the upper band that were in the Russell 2000 last year and stocks near the lower band that were in the Russell 1000 last year). Figure 2 shows the samples in the last pre-banding year (2006) and the first post-banding year (2007). To ensure that poor liquidity or market micro-structure issues do not affect our estimates we drop stocks that had a May share price under \$5 per share (see Asparouhova, Bessembinder, and Kalcheva (2015)). Our results are similar if we omit this filter.

Table I presents summary statistics for the pre-banding sample from 1993-2006. The firms are quite homogeneous in their May market capitalizations, as it is the basis of the index rankings. Thus, our sample consists of a tight bracket of mid-cap stocks with an average May market capitalization of \$1.3 billion in the pre-banding sample and \$2.6 billion in the post-banding sample. Sample stocks had higher turnover (monthly trading volume as a fraction of market cap) in the post-banding period, and very similar levels of monthly return volatility in the pre- and post-banding periods.

The remaining rows of Table I summarize the institutional ownership of our sample stocks as of the December *prior* to inclusion in the sample. Stocks arrived into the sample with average total mutual fund ownership of 13.0% of market capitalization in the prebanding period; 11.9% held by active funds and 1.1% held by passive funds. Ex ante mutual fund ownership was nearly twice as high in the postbanding period, at 23.1% of market capitalization in total; 17.6% held by active funds and 5.5% held by passive funds. The standard deviations and 10th and 90th percentiles indicate that the heterogeneity of mutual fund holdings across sample stocks was also higher in the postbanding period.

IV. Results

In this section we examine whether index investing has an effect on institutional ownership, stock returns, and price efficiency. We start by examining whether, and how, index rebalancing impacts the mix of active versus passive ownership. We then examine the relation between index investing and weak form price efficiency. Finally, we examine the relation between index investing and trading by arbitrageurs. Overall, our findings suggest that higher index ownership leads to a significant change in prices and weak form efficiency, but it does not appear to change trading behavior or semi-strong form price efficiency.

A. Effects on Fund Ownership

Table II presents our RDD estimates for fund ownership, which compare year-on-year (December-to-December) changes in mutual fund ownership across the Russell index discontinuities. The independent variable $R2000$ is a dummy variable for assignment to the Russell 2000. In Panel A (pre-banding) the specification includes quadratic control functions over our proxy ranking, and year fixed effects which remove any yearly aggregate changes in ownership. In Panel B (post-banding) the specification includes quadratic control functions over our proxy ranking, and band-by-year fixed effects which remove any yearly aggregate changes in ownership in each band separately. Panel C presents the pooled estimates across both periods.

In all three panels, we see that total mutual fund ownership ($FundOwn$) does not change significantly across Russell discontinuities. However, assignment to the Russell 2000 strongly alters the *composition* of institutional ownership. Ownership by passive index funds $FundOwn^{PASSIVE}$ increases by 0.48% of market cap in the pre-banding period and 0.79% of market cap in the post-banding period, an average of 0.54% of market cap in the pooled sample. Interestingly, we also see evidence that passive investors are buying from active investors (rather than retail traders or uncategorized investors): in all three estimates, off-setting the rise in passive investment, ownership by active mutual funds ($FundOwn^{ACTIVE}$) falls. The changes in holdings by passive funds are strongly significant because holdings by passive funds at the stock level are stable year to year and thus the standard error of the estimates is quite low. By contrast, although the point estimate of the change in holdings by active mutual funds is of a similar size, the standard error is an order of magnitude larger because holdings by active mutual funds are much more volatile. In short, our estimates broadly suggest that passive investors *replace* active investors as owners of a stock following its addition to the Russell 2000.

Table III compares the monthly returns of sample stocks centered on the index reconstitution in June of each year. There is no significant difference in returns prior to June, which suggests that stocks' treatment status was not systematically anticipated prior to index assignment in either regime. In the prebanding period, immediately after index reconstitution stocks assigned to the Russell 2000 had an average June return that was 2.5% higher than control stocks. The results are consistent with buying pressure from passive investors during June when the new index assignments were released and confirm the vast literature on downward sloping demand curves for equities (e.g., Shleifer (1986), Bagwell (1992), Chang et al. (2015), etc.). This price response for treated stocks in the pre-banding period was partly reversed by -2.0% in July. In August, post-assignment, there was again no statistically significant difference between treated vs control stocks' monthly returns. Overall, the results imply that our sample of U.S. equities has a price elasticity of demand of approximately -0.2, similar to the estimates in Chang et al. (2015).[8]

Interestingly, we see no significant response in June or July returns around the Russell discontinuities in the post-banding period (Panel B), although the sign of the estimates is the same. The lack of statistical significance appears to be due to a muted return movement and a larger standard error. We note that our sample size is significantly smaller in the post-

[8]Chang et al. (2015) report an elasticity that ranges from -0.3 to -1.4, depending on the assets used to measure changes in ownership. Similarly, Kaul, Mehrotra, and Morck (2000) calculate elasticities following a change in index weights. Petajisto (2009), footnote 13, calculates an elasticity of -0.3 using the estimates from Kaul et al. (2000).

banding period, but when we pool the two periods in Panel C, we again find a significant price response. Overall, the results in this section show that index assignment changes the composition of investors, from active to passive, which leads to a permanent increase in the stock price.

B. Weak-Form Price Efficiency

We next exploit the strong and sharply localized effects of Russell index assignments – increased passive and decreased active investment in treated stocks – to examine the effects of index assignment on weak-form price efficiency. To measure price efficiency for our sample stocks, we compute variance ratios of returns over horizons of q = 2, 4, and 8 trading days from July 1 to May 31 of each year.[9] Formally, we use the q-period bias-corrected variance ratio test of Lo and MacKinlay (1988):

$$VarRatio(q) = \frac{\hat{\sigma}^2(q)}{q \times \hat{\sigma}^2}, \tag{4}$$

where

$$\hat{\sigma}^2(q) = \frac{k}{(n-q+1)(k-1)} \sum_{t=q}^{n} (p_t - p_{t-q} - q\hat{\mu})^2, \tag{5}$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{t=1}^{n} (p_t - p_{t-1} - \hat{\mu})^2, \tag{6}$$

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^{n} (p_t - p_{t-1}), \tag{7}$$

and the data consists of $kq+1$ observations (for convenience we define $n = kq$). We calculate variance ratios separately for each firm and year using overlapping observations within the year. The efficient benchmark is a variance ratio equal to 1 – that is, returns over a q-day horizon had a variance that was q times the variance of daily returns. Our measure of (inverse) price efficiency is how far the stock's variance ratio was from 1, as in Boehmer and

[9]We omit June of each year, to avoid measuring the effects of trading in the month of index rebalancing.

Kelley (2009). In other words, we are measuring how far the stock price process deviated from a pure random walk. Formally, we examine the absolute value of the centered ratio:

$$AbsVarRatio_t(q) = abs(VarRatio_t(q) - 1) \tag{8}$$

Thus, $AbsVarRatio = 0$ defines perfect weak-form efficiency, and the larger the value of $AbsVarRatio$, the further the stock price process is from the random-walk benchmark.

Table IV presents our RDD estimates for weak form price efficiency around the Russell index discontinuities. In the pre- (Panel A) and post-banding (Panel B) periods, the coefficient estimates are all positive, although the statistical significance varies by the horizon. The point estimates indicate that stocks with increased index investment had significantly less efficient prices post-treatment. More specifically, these stocks had prices that were significantly less likely to follow a random walk. In panel B, the results are statistically insignificant at the 2-day and 4-day horizon, but as with Table III we note that we have only 700 observations. The pooled full-sample estimates (Panel C) again show positive coefficients at all horizons, and the added statistical power of the pooled estimates indicates that the decrease in treated stocks' price efficiency is statistically significant at all three horizons. In sum, increased index investing is associated with worse weak form price efficiency; following an increase in index investment, prices are significantly less likely to follow a random walk.

C. Other Characteristics

As discussed above, our results show that index investing changes the price process. After being added to the Russell 2000, stocks experience a permanent price increase and their prices are significantly less likely to follow a random walk. Several theoretical models suggest that an increase in index investment could generate additional effects, including changes to the correlation structure of prices (e.g., Baruch and Zhang (2017)), increased volatility (e.g., Basak and Pavlova (2013)), and changes to the information gathering and behavior of informed investors (e.g., Cong and Xu (2016), Brown and Davies (2017), Bond and Garcia (2017)). Accordingly, in this section we explore the effects of index investing on a variety of stock-level outcomes suggested by theory. To simplify the presentation of our results and maximize precision, henceforth we present only the pooled estimates that combine the pre-banding and post-banding regimes.[10]

We first investigate the effects on daily return volatilities and correlations. In columns (1) through (4) of Table V, we find that daily stock return volatility increased strongly and consistently at all horizons from 1 to 8 days, consistent with the theoretical predictions in Basak and Pavlova (2013). The magnitude of the RDD coefficient suggests that an increase in index investment of 1% of a stock's market capitalization is associated with a rise in its daily return volatility of 0.5%, which corresponds to 1/3 of one standard deviation. In other words, increased index investment results in a significant rise in stock-level volatility.

In column (5), we examine the impact of index investing on the correlation structure of returns. Specifically, we examine the correlation between a stock's return and the return on the Russell Index. The result shows a significant increase in the correlation between a treated stock's daily return and the daily return of the Russell 2000 index. Importantly, the increase is not mechanical; it also holds for each stock's correlation with the other individual stocks in the Russell 2000 (results not shown). Finally, column 6 shows that treated stocks' daily return autocorrelation (i.e. correlation with the stock's own lagged daily return) fell slightly relative to control stocks.

We next investigate the change in *monthly* return volatility and trading volume as a consequence of Russell index assignment. Table VI presents the RDD estimates. We see that index assignment did not alter firms' trading volume as measured by turnover or share volume (columns 2 and 3). In other words, an increase in index investing does not appear to significantly change trading behavior. Interestingly, this suggests that index investors may not be truly passive in the "buy and hold" sense. Indeed, anecdotal evidence suggests that

[10]Results broken out by pre- and post-banding periods are available from the authors upon request.

turnover in index ETFs is strikingly high. In 2017, the two most heavily traded assets in U.S. stock markets were the SPY and IWM ETFs, which are index ETFs benchmarked to the S&P500 and Russell 2000, respectively. Turnover in the SPY averaged $19 billion per day, while turnover in IWM averaged $3.5 billion per day (Vlastelica (2017)). By comparison, the most actively traded stock was Apple, with turnover just under $1 billion per day.

While index assignment does not appear to generate lower turnover or trading volume at the monthly level, we do find evidence that it changes monthly volatility, consistent with the daily estimates presented in Table V. Column (1) indicates that assignment to the Russell 2000 is followed by higher volatility of monthly returns. The magnitude of the RDD coefficient suggests that an increase in index investment of 1% of a stock's market capitalization is associated with a rise in its monthly return volatility of 1.6%. Compared to the average monthly volatility for our sample stocks of 10.7%, this suggests a sizeable effect of index investment on stock price volatility.

D. Post-Earnings Announcement Drift

Next, we examine whether increased index investing impacts the incentives and abilities of arbitrageurs to trade on value relevant information. To do this, we use a simple and well-documented source of return predictability that varies by stock: post-earnings-announcement drift (PEAD). Our measure is based on earnings announcements from 1 to 24 months post-index assignment for each sample stock-year. Specifically, for an earnings announcement on day t, we regress the cumulative stock return from day $t+3$ to day $t+63$ on the standardized unexpected earnings (SUE) of the announcement, which is scaled so a value of 1 represents a surprise in earnings per share equal to 1% of the stock price. The coefficient of returns on SUE, for each stock-year in the sample, is denoted $Beta_{SUE}$. A value of $Beta_{SUE} = 0$ corresponds to no post-earnings-announcement drift i.e. the stock incorporates all the information contained in the earnings announcement, on average, within 2 trading days after the announcement.

The mean $Beta_{SUE}$ in our sample is 4.3%; in other words, a 1% standard deviation positive surprise in the stock's earnings announcement is followed by a predictable post-announcement return drift of 4.3%. The sample mean has a standard error of 0.6% and so the mean is strongly different from zero; that is, in our sample of mid-cap stocks, there is significant PEAD on average. However, the sample stocks are also quite heterogeneous in their level of PEAD: the standard deviation of $Beta_{SUE}$ is 28%.

We next explore whether index investing impacts the incentives and abilities of arbitrageurs to gather fundamental information about a firm. Specifically, we examine whether Russell index assignment impacts PEAD. Table VII presents RDD estimates of the effects of Russell 2000 assignment on post-earnings-announcement drift. Column 1 shows that the treatment effect of an average 0.54% market capitalization increase in index investing is insignificant both statistically and economically (a one standard deviation increase in SUE is associated with a -0.8% decrease in return drift). Thus, although greater index investment increases stock price volatility and lowers weak-form stock price efficiency, our estimates indicate that it does not alter semi-strong form efficiency as measured by PEAD.

The insignificant treatment effect on PEAD that we find is unlikely to be due to a type-II error (i.e. it is not due to insufficient power). The standard error on the treatment coefficient is 1.6%, so our test has the power to detect a treatment effect on the order of one-tenth of one standard deviation in $Beta_{SUE}$. Table VII Column 2 presents the treatment effects when we winsorize $Beta_{SUE}$ at the 1% level to eliminate any extreme values. We see that the resolving power of the test is even stronger (standard error of the treatment coefficient falls to 1.2%), yet the estimated treatment effect is still insignificant.

Table VII Columns 3 and 4 present similar results when we examine the reaction of stocks' daily trading volume to a 1% earnings surprise. Again the reaction in terms of trading volume is slightly negative, very close to zero – and precisely estimated. Thus, neither the reaction of sample stocks' returns or trading volume are significantly affected by index inclusion and changes in passive investment.

To summarize, in this section we examine the effect of index investing on semi-strong form efficiency as measured by post-earnings-announcement drift. We find a zero treatment effect, which is precisely estimated, indicating that index investment does not significantly change treated stocks' price efficiency. Taken with our results on turnover and trading volume, the results suggest that increased index investing does not alter the trading behavior of informed agents.

E. Balance Tests

Finally, to ensure that our results are not due to a potential selection bias, we examine balance tests across a wide-variety of dependent variables. Table VIII presents estimates using *ex ante* (i.e., pre-treatment) levels of firm characteristics. Importantly, we see that in the year *prior* to index assignment there is no significant difference between treated and control stocks in any of the variables examined, except for a small difference in pre-treatment passive fund ownership which is marginally significant at the 10% level. Given that we run 15 balance tests in total, under the null hypothesis we expect to observe at least one result that is significant at the 10% level given a Type I error rate of 10%. More to the point, the imbalance on passive ownership is both extremely small – 0.03% of market cap – and in the opposite direction to our estimated treatment effect of 0.54%. Thus, the results in Table VIII suggest that our research design represents a clean setting that is free from selection bias. Moreover, the results suggest that the relation between index membership and our outcome variables – in particular, stock price volatility and price efficiency – is causal in nature.

The exclusion restriction in our setting requires that within each year, and conditional on the control functions of the forcing variable $CAPrank$ on either side of the Russell discontinuities, firms differed *only* on their subsequent index assignment. The finding that there was no pre-treatment difference in our outcome variables is consistent with this assumption. In unreported tests, we perform a variety of additional tests that further bolster confidence in our research design's validity. For example, in unreported placebo tests available upon request, we re-run the analyses using the "wrong" index threshold set at rank = 750 or 1250 and we find zero results across the board. Overall, the results suggest that our empirical methodology cleanly identifies the impact of increased index investing on firm-level outcomes.

V. Discussion

As discussed in the introduction, we are careful to note that our results do not provide welfare conclusions. Although our results show that index investing causes higher volatility and worse weak form price efficiency, we stress that we are documenting one aspect of the impact of passive investing. We do not attempt to evaluate the costs and benefits from passive investing in this paper. As such, future research should continue to explore the total welfare impact from passive investing.

In addition, we are careful to point out that, as with all RDD estimates, the external validity of our results remains unknown. The RDD methodology we employ estimates the *local average treatment effect* (LATE). In our sample, the average change in index investing amounts to approximately 0.4% of market cap pre-banding and 0.8% of market cap post-banding. Our estimates are not able to establish whether larger (or smaller) changes in index investing would have similar effects to the results we document. Moreover, our estimate of the impact on price efficiency is for stocks assigned to the top of the Russell 2000 index; as such, it is not clear if our estimate would apply to stocks that are substantially different from the largest stocks in the Russell 2000. Overall, our results provide strong evidence that index assignment leads to changes in the price process for stocks, but we leave welfare questions for future research.

VI. Conclusion

We examine the impact of the recent increase in index investing. If index investors are passive in the "buy and hold" sense, then they are, by definition, free-riding off the information production of active managers and this could change overall price informativeness. In other words, not everyone can index, some investors have to be *active* managers. The question is, how many active managers are enough to ensure that prices correctly reflect fundamental values? In this paper, we show that recent increases in index investing have changed the price process, but they have not significantly changed trading or semi-strong form price efficiency.

In the last few decades, index investing has grown dramatically as a fraction of total investing activity. Yet, to date, the existing literature has been unable to determine the consequences of this increase in index investing for price efficiency. On one hand, index investing reduces the 'float' that active traders are able to access and thus potentially reduces the payoff to informed trading. On the other hand, active investors are still a significant portion of the market. Similarly, it is possible that increased index investing changes the behavior of noise traders which then changes the incentives of active managers to invest in costly information acquisition. Overall, the net impact of increased index investing remains an open question.

Our paper provides a novel answer to this question. We measure the effects of index investing on price efficiency using a natural experiment. From the mid-1990s to 2006, the yearly assignment of stocks to the Russell 1000 and 2000 indexes followed a simple mechanical rule which created an exogenous discontinuity in the index ownership of stocks near the threshold. Subsequent to 2006, the "banding" regime broke the original index discontinuity, but created two new discontinuities in index switching. We exploit the subsequent variation over the period 2007 to 2016 to measure the causal effects of index assignment on firm-level characteristics.

We find that weak-form efficiency is reduced in stocks that experienced an exogenous increase in passive investment due to index assignment. Compared to an ex-ante identical set of control stocks, following index assignment, treated stocks' prices moved further away from the efficient random-walk benchmark. In addition, we find that these stocks experience higher daily and monthly volatility as well as increased correlations with other members of the Russell 2000 index. However, when we examine measures of trading and semi-strong form efficiency, we find no effects. Treated stocks do not experience any difference in turnover or trading volume, nor do they experience changes in semi-strong form efficiency as measured by post-earnings announcement drift. Overall, our results suggest that index investing does change the price process, but it does not significantly change the ability of arbitrageurs to impound information into prices.

References

Appel, I., Gormley, T. A., & Keim, D. B. (2016). Passive investors, not passive owners. *Journal of Financial Economics*, *forthcoming*.

Bagwell, L. S. (1992). Dutch auction repurchases: An analysis of shareholder heterogeneity. *Journal of Finance*, *47*, 71-105.

Baruch, S., & Zhang, X. (2017). Is index trading benign?

Basak, S., & Pavlova, A. (2013). Asset prices and institutional investors. *American Economic Review*, *103*, 1728-1758.

Ben-David, I., Franzoni, F. A., & Moussawi, R. (2016). Do ETFs increase volatility? *Working Paper*.

Boehmer, E., & Kelley, E. K. (2009). Institutional investors and the informational efficiency of prices. *Review of Financial Studies*, *22*, 3563-3594.

Bogle, J. C. (2016). The index mutual fund: 40 years of growth, change, and challenge. *Financial Analyst Journal*, *January/February*, 9-13.

Bond, P., & Garcia, D. (2017). The equilibrium consequences of indexing.

Brown, D. C., & Davies, S. W. (2017). Moral hazard in active asset management. , *1*, 1-15.

Campbell, J. Y., Lo, A., & MacKinlay, A. (1997). *The econometrics of financial markets*. New Jersey: Princeton University Press.

Campbell, J. Y., & Mankiw, G. (1987). Permanent and transitory components in macroeconomic fluctuations. *American Economic Review*, *77*, 111-117.

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

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

Cong, L. W., & Xu, D. X. (2016). Rise of factor investing: Asset prices, informational efficiency, and security design. *Working Paper*.

Cowles, A. (1933). Can stock market forecasters forecast? *Econometrica*, *1*, 309-324.

Crane, A. D., Michenaud, S., & Weston, J. P. (2016). The effect of institutional ownership on payout policy: Evidence from index thresholds. *Review of Financial Studies*, *29*, 1377-1408.

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

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

French, K. (2008). Presidential address: The cost of active investing. *Journal of Finance*, *63*, 1537-1573.

Glosten, L., Nallareddy, S., & Zhou, Y. (2016). ETF activity and informational efficiency of underlying securities. *Working Paper*.

Goldstein, I., & Yang, L. (2017). Commodity financialization: Risk sharing and price discovery in commodity futures markets. *Working Paper*.

Hou, K., & Moskowitz, T. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, *18*, 981-1020.

Investment Company Institute. (2017). *Trends in mutual fund investing January 2017.* Retrieved from `https://www.ici.org/research/stats/trends/trends_01_17`

Israeli, D., Lee, C. M., & Sridharan, S. A. (2016). Is there a dark side to exchange trade funds (ETFs)? An information perspective. *Working Paper*.

Jensen, M. C. (1967). The performance of mutual ffund in the period 1945-1964. *Journal of Finance*, *23*, 389-416.

Kaul, A., Mehrotra, V., & Morck, R. (2000). Demand curves for stocks do slope down: New evidence from an index weights adjustment. *Journal of Finance*, *55*, 893-912.

Lo, A., & MacKinlay, C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. (1), 41-66.

Mullins, W. (2014). The governance impact of index funds: Evidence from regression discontinuity. *Working Paper*.

Petajisto, A. (2009). Why do demand curves for stocks slope down? *Journal of Financial and Quantitative Analysis*, *44*, 1013-1044.

Schmidt, C., & Fahlenbrach, R. (2017). Do exogenous cchange in passive institutional ownership affect corporate governance and firm value? *Journal of Financial Economics*, *Forthcoming*.

Schnitzler, J. (2016). S&P500 inclusions and stock supply. *Working Paper*.

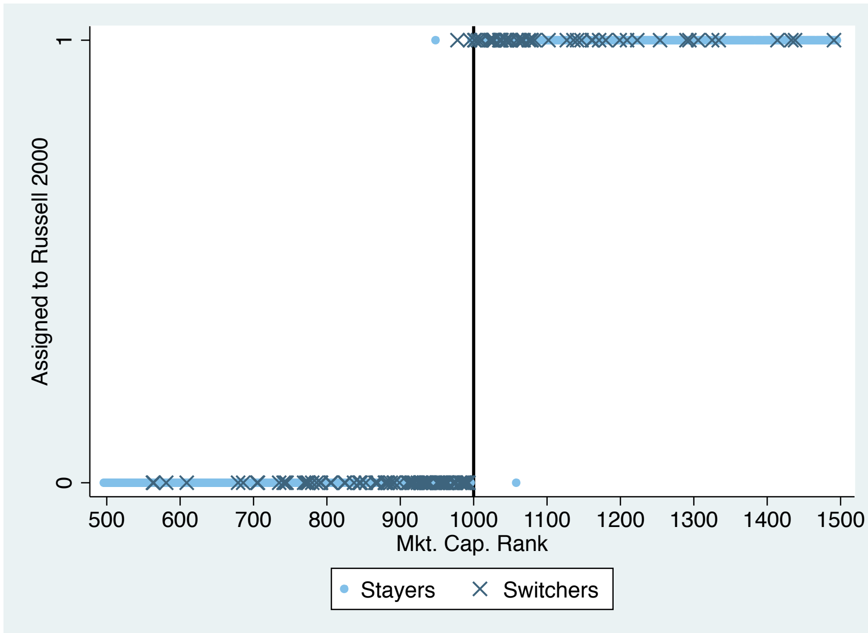
Sharpe, W. F. (1991). The arithmetic of active management. *The Financial Analysts' Journal*, *47*, 7-9.

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

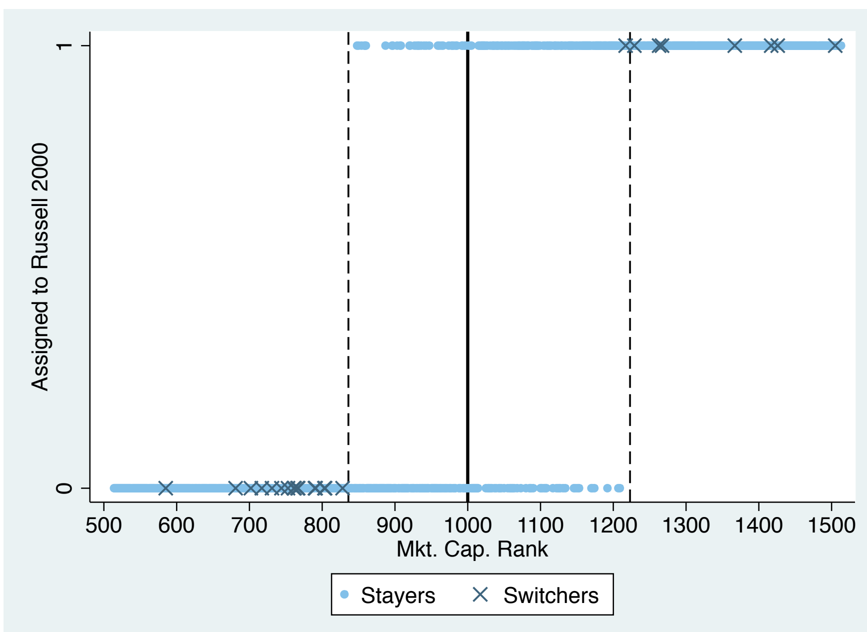
Stein, J. C. (1987). Informational externalities of welfare-reducing speculation. *Journal of Political Economy*, *95*, 1123-1145.

Subrahmanyam, A. (1991). A theory of trading in stock index futures. *The Review of Financial Studies*, *4*, 17-51.

Vlastelica, R. (2017). *Here's how much etfs are dominating the trading floor.* Retrieved from `http://www.marketwatch.com/story/heres-how-much-etfs-are-dominating-on-the-trading-floor-2017-01-10`

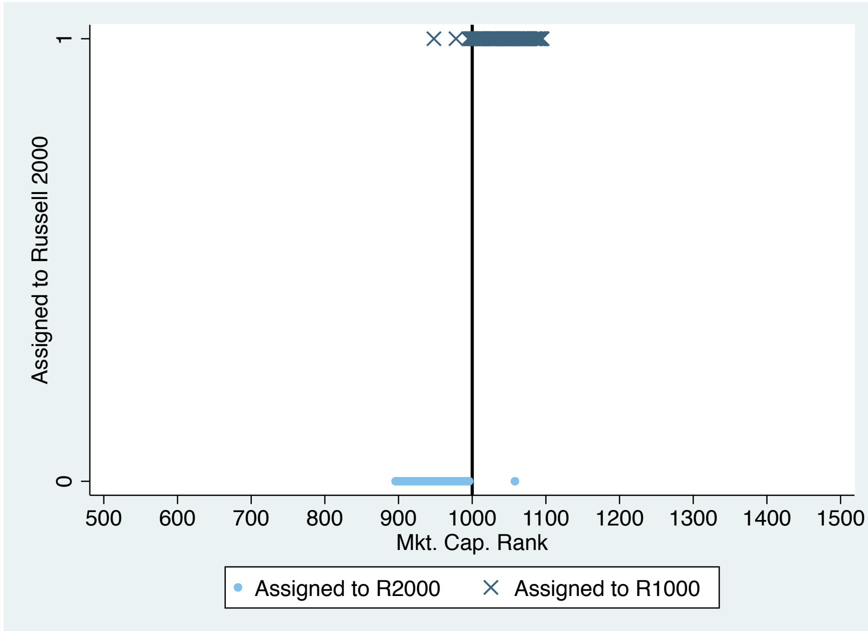


(a) 2006 Index Assignments

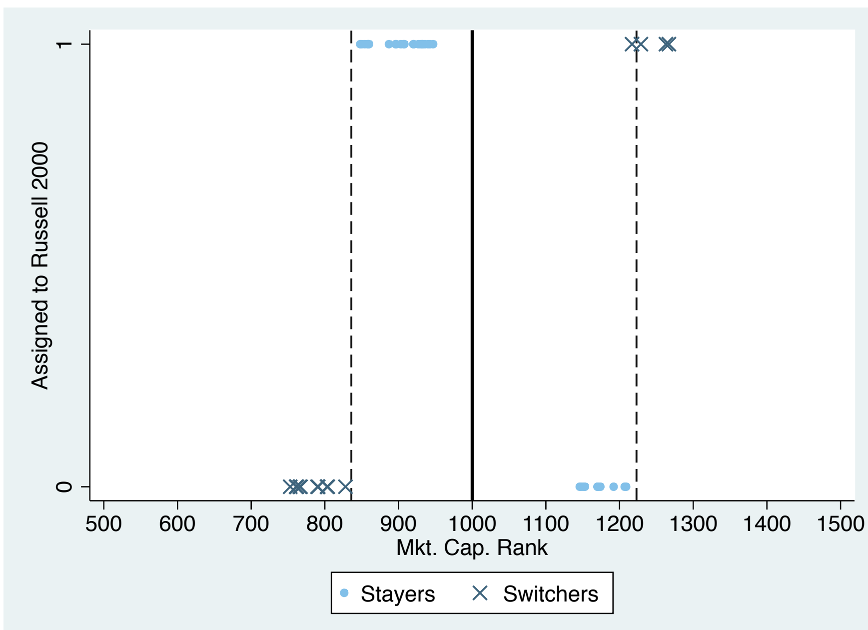


(b) 2007 Index Assignments

Figure 1. Index Assignment Pre- and Post-Banding
The figure plots stocks' assignments to the Russell 1000 and 2000 indexes in June of 2006 and 2007 against our proxy for Russell's proprietary market cap rankings. In 2006, the last year before banding was introduced, there is a sharp discontinuity in index assignment at the index threshold (solid line). In 2007, stocks near the threshold all stayed in their previous years' index, breaking the discontinuity in index assignment. Close to the estimated upper and lower bands (dashed lines), however, there are sharp discontinuities in index switching.



(a) 2006 Sample Stocks



(b) 2007 Sample Stocks

Figure 2. Sample Construction Pre- and Post-Banding
The figure plots the sample stocks' assignment to the Russell 1000 and 2000 indexes in June of 2006 and 2007 against our proxy for Russell's proprietary market cap rankings. In the pre-banding period, the sample is all stocks in a +/-100 rank window around the Russell index threshold. In the post-banding period, the sample is stocks in a +/-100 rank around the upper and lower bands that had the potential to switch indexes from the previous year.

Table I
Summary statistics

The table presents summary statistics for ex ante stock characteristics: market capitalization, monthly trading turnover, monthly volatility, and categories of mutual fund ownership. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016.

Panel A: Pre-Banding Sample, 1993-2006

	Mean	Std. Dev.	p10	p50	p90	N
	(1)	(2)	(3)	(4)	(5)	(6)
Market Cap ($M)	1284.4	412.6	669.0	1322.2	1789.4	2768
Turnover	1.50	1.61	0.27	0.99	3.31	2610
Rtn Volatility	0.111	0.060	0.051	0.090	0.176	2481
$FundOwn_{i,t-1}$	13.0	9.2	2.0	11.6	25.8	2768
$FundOwn_{i,t-1}^{ACTIVE}$	11.9	8.7	1.4	10.5	24.2	2768
$FundOwn_{i,t-1}^{PASSIVE}$	1.1	1.3	0.1	0.5	2.9	2768

Panel B: Post-Banding Sample, 2007-2016

	Mean	Std. Dev.	p10	p50	p90	N
	(1)	(2)	(3)	(4)	(5)	(6)
Market Cap ($M)	2598.9	964.5	1473.8	2500.6	3995.9	816
Turnover	2.86	2.03	1.08	2.43	4.99	814
Rtn Volatility	0.106	0.062	0.047	0.092	0.175	789
$FundOwn_{i,t-1}$	23.1	16.5	0.0	26.9	42.6	816
$FundOwn_{i,t-1}^{ACTIVE}$	17.6	13.5	0.0	19.2	33.5	816
$FundOwn_{i,t-1}^{PASSIVE}$	5.5	4.6	0.0	5.3	11.5	816

Table II
RDD Estimates: Institutional Ownership

The table presents estimates of the effects of Russell index assignment on mutual fund ownership expressed as a percentage (1=1%) of market capitalization. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

Panel A: Pre-Banding Sample, 1993-2006

	$\Delta FundOwn$	$\Delta FundOwn^{ACTIVE}$	$\Delta FundOwn^{PASSIVE}$
	(1)	(2)	(3)
$R2000$	-0.156	-0.635	0.479***
	(0.398)	(0.394)	(0.025)
Observations	2,585	2,585	2,585
R-squared	0.074	0.069	0.585
Lagged Index	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Panel B: Post-Banding Sample, 2007-2016

	$\Delta FundOwn$	$\Delta FundOwn^{ACTIVE}$	$\Delta FundOwn^{PASSIVE}$
	(1)	(2)	(3)
$R2000$	0.435	-0.357	0.792***
	(0.838)	(0.857)	(0.079)
Observations	816	816	816
R-squared	0.075	0.099	0.693
Band x Year FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Panel C: Pooled Sample, 1993-2016

	$\Delta FundOwn$	$\Delta FundOwn^{ACTIVE}$	$\Delta FundOwn^{PASSIVE}$
	(1)	(2)	(3)
R2000	-0.005	-0.545	0.539***
	(0.364)	(0.364)	(0.027)
Observations	3,401	3,401	3,401
R-squared	0.076	0.075	0.668
Lagged Index	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Table III
RDD Estimates: Monthly Stock Returns

The table presents estimates of the effects of Russell index assignment on monthly stock returns.The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

Panel A: Pre-Banding Sample, 1993-2006

Panel B: Post-Banding Sample, 2007-2016

	rtn_{April} (1)	rtn_{May} (2)	rtn_{June} (3)	rtn_{July} (4)	rtn_{Aug} (5)
$R2000$	0.001	-0.003	0.013	-0.009	0.019
	(0.012)	(0.012)	(0.010)	(0.014)	(0.014)
Observations	816	816	816	812	808
R-squared	0.349	0.266	0.327	0.239	0.322
Band x Year FE	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2

Panel C: Pooled Sample, 1993-2016

	rtn_{April} (1)	rtn_{May} (2)	rtn_{June} (3)	rtn_{July} (4)	rtn_{Aug} (5)
$R2000$	-0.002	-0.002	0.021***	-0.016**	0.003
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)
Observations	3,401	3,401	3,400	3,384	3,371
R-squared	0.213	0.182	0.136	0.175	0.300
Lagged Index	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2

Table IV
RDD Estimates: Distance from the Efficient Stock Price Benchmark

The table presents estimates of the effects of index assignment on stock price efficiency using the Russell index discontinuities. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

Panel A: Pre-Banding Sample, 1993-2006

	$AbsVarRatio_t^2$	$AbsVarRatio_t^4$	$AbsVarRatio_t^8$
	(1)	(2)	(3)
$R2000$	0.012***	0.011	0.018*
	(0.004)	(0.008)	(0.011)
Observations	2,561	2,559	2,557
R-squared	0.035	0.029	0.022
Lagged Index	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Panel B: Post-Banding Sample, 2007-2016

	$AbsVarRatio_t^2$	$AbsVarRatio_t^4$	$AbsVarRatio_t^8$
	(1)	(2)	(3)
$R2000$	0.001	0.019	0.031*
	(0.007)	(0.012)	(0.018)
Observations	722	722	722
R-squared	0.055	0.064	0.038
Band x Year FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Panel C: Pooled Sample, 1993-2016

	$AbsVarRatio_t^2$	$AbsVarRatio_t^4$	$AbsVarRatio_t^8$
	(1)	(2)	(3)
$R2000$	0.010***	0.012*	0.020**
	(0.004)	(0.006)	(0.009)
Observations	3,283	3,281	3,279
R-squared	0.044	0.040	0.029
Lagged Index	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Table V
RDD Estimates: Volatilities and Correlations

The table presents estimates of the effects of index assignment on treated stocks' daily return volatilities, correlation with the Russell 2000's daily index changes, and daily return autocorrelation using the Russell index discontinuities. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

	$rtnvol_{1day}$	$rtnvol_{2day}$	$rtnvol_{4day}$	$rtnvol_{8day}$	$corr(r_t, r_t^{R2000})$	$corr(r_t, r_{t-1})$
	(1)	(2)	(3)	(4)	(5)	(6)
$R2000$	0.0024***	0.0024***	0.0024***	0.0023***	0.055***	-0.011*
	(0.0008)	(0.0008)	(0.0009)	(0.0008)	(0.007)	(0.006)
Observations	3,285	3,285	3,285	3,285	3,384	3,384
R-squared	0.321	0.311	0.307	0.305	0.484	0.068
Lagged Index	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2	2

Table VI
RDD Estimates: Other Stock Variables

The table presents estimates of the effects of index assignment on monthly return volatility, turnover and share volume, using the Russell index discontinuities. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

	$Volatility_t$	$Turnover_t$	$ShareVolume_t$	$BidAsk_t(\%)$	$Amihud_t$
$R2000$	0.008*	0.069	14,839	0.015	-0.003
	(0.004)	(0.112)	(13,482)	(0.041)	(0.002)
Observations	3,159	3,159	3,159	3,395	3,401
R-squared	0.256	0.204	0.200	0.581	0.108
Lagged Index	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2

Table VII

RDD Estimates: Post-Earnings Announcement Drift

The table presents estimates of the effects of index assignment on post-earnings announcement drift using the Russell index discontinuities. The dependent variable is the coefficient of one-quarter-ahead stock returns post earnings announcement on the standardized unexpected earnings value (a surprise in earnings per share equal to 1% of the stock price) for all announcements from 1 to 24 months post-assignment. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

	$Beta_{SUE}$	Winsorized $Beta_{SUE}$	$Beta_{SUE}^{Volume}$	Winsorized $Beta_{SUE}^{Volume}$
	(1)	(2)	(3)	(4)
$R2000$	-0.008	-0.013	-0.16	-0.13
	(0.016)	(0.012)	(0.17)	(0.14)
Observations	1,922	1,922	1,922	1,922
R-squared	0.016	0.027	0.020	0.018
Lagged Index	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2

Table VIII
RDD Estimates: Balance Tests

The table presents estimates of potential selection bias in the sample by regressing *ex ante* stock price efficiency (Absolute Variance Ratio, AVR) and other stock characteristics on index assignment using the Russell index discontinuities. The sample consists of stocks within a +/- 100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

	AVR^2_{t-1} (1)	AVR^4_{t-1} (2)	AVR^8_{t-1} (3)	$logBookVal_{t-1}$ (4)	$logMktVal_{t-1}$ (5)	$corrR2000_{t-1}$ (6)	$rtnvol^{1mth}_{t-1}$ (7)	$turnover_{t-1}$ (8)
$R2000$	-0.002	0.004	0.003	-0.077	-0.025	0.000	0.003	-0.131
	(0.004)	(0.006)	(0.009)	(0.061)	(0.024)	(0.007)	(0.003)	(0.098)
Observations	3,285	3,285	3,285	3,327	3,324	3,384	3,259	3,399
R-squared	0.052	0.041	0.039	0.183	0.632	0.570	0.237	0.189
Lagged Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2	2	2	2

	$\Delta FundOwn$ (1)	$\Delta FundOwn^{ACTIVE}$ (2)	$\Delta FundOwn^{PASSIVE}$ (3)	$rtnvol_{1day}$ (4)	$rtnvol_{2day}$ (5)	$rtnvol_{4day}$ (6)	$rtnvol_{8day}$ (7)
$R2000$	-0.018	0.015	-0.032*	0.001	0.001	0.001	0.001
	(0.497)	(0.492)	(0.019)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	3,401	3,401	3,401	3,285	3,285	3,285	3,285
R-squared	0.470	0.446	0.790	0.304	0.284	0.274	0.271
Lagged Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2	2	2

VII. Appendix

This appendix[11] provides additional theoretical and empirical evidence to supplement the main text. In Section A we provide supplemental empirical analyses.

A. Alternate Control Functions

We first investigate how sensitive are our findings on the effects of Russell index assignment to the choice of control functions around the yearly Russell discontinuities. Our main specification uses a quadratic control function.

Table A1 presents the main results when we vary the degree of the control function from 1 (linear control function) to 3 (cubic) polynomial. The estimates are similar in all cases, indicating that our results are not sensitive to the choice of control function.

B. Alternate Windows Around the Russell Threshold

We next explore whether our results are sensitive to our choice of the window of stocks around the Russell discontinuities. The tradeoff is that when we widen the window, the treated stocks may differ from the control stocks systematically even after conditioning flexibly on $CAPrank$, whereas when we narrow the window, treated stocks will look more and more similar to control stocks even unconditionally, but we lose statistical power.

Table A2 presents the main results when we narrow the window of stocks around the yearly Russell discontinuities in our sample construction. We see that using narrower (N=75, N=50) windows, the estimates are similar to our main estimates in all cases, although statistical significance is reduced due to the reduced sample size. Thus, our estimates are not sensitive to the choice of window around the yearly Russell discontinuities.

[11]Citation format: Coles, Jeffrey L., Davidson Heath, and Matthew C. Ringgenberg, Appendix for "On Index Investing," 2017, Working Paper.

Table A1

RDD Estimates: Robustness I

The table presents estimates of the effects of index assignment on stock price efficiency (Absolute Variance Ratio, AVR) using the Russell index discontinuities, varying the degree of the RDD control functions. The sample consists of stocks within a +/-100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

	AVR_t^2	AVR_t^4	AVR_t^8	AVR_t^2	AVR_t^4	AVR_t^8
	(1)	(2)	(3)	(4)	(5)	(6)
$R2000$	0.011***	0.015**	0.022**	0.011***	0.015**	0.021**
	(0.004)	(0.006)	(0.009)	(0.004)	(0.006)	(0.009)
Observations	3,465	3,463	3,461	3,465	3,463	3,461
R-squared	0.043	0.039	0.026	0.043	0.039	0.027
Lagged Index	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	1	1	1	3	3	3

Table A2

RDD Estimates: Robustness II

The table presents estimates of the effects of index assignment on stock price efficiency (Absolute Variance Ratio, AVR) using the Russell index discontinuities, varying the bandwidth around the discontinuities. The sample consists of stocks within a +/-N rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-N rank window of the yearly Russell upper and lower bands from 2007-2016. Standard errors are HAC and clustered by stock.

	AVR_t^2	AVR_t^4	AVR_t^8	AVR_t^2	AVR_t^4	AVR_t^8
	(1)	(2)	(3)	(4)	(5)	(6)
$R2000$	0.012***	0.013*	0.021**	0.013**	0.013	0.017
	(0.004)	(0.007)	(0.010)	(0.005)	(0.009)	(0.013)
Observations	2,586	2,585	2,584	1,721	1,720	1,719
R-squared	0.043	0.038	0.028	0.054	0.039	0.027
Window	75	75	75	50	50	50
Lagged Index	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Band x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2	2

Table A3

RDD Estimates: Fund Holdings

The table presents estimates of the effects of index assignment on fund holdings of sample stocks using the Russell index discontinuities. The outcome variable is a dummy for whether a given mutual fund holds a given stock in December of the year post-treatment. The sample consists of stocks within a +/-100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2012. Standard errors are HAC and clustered by fund.

	(1) All Funds	(2) R2000 Passive	(3) All Others
$R2000$	0.004	0.533***	-0.006***
	(1.5)	(7.4)	(-3.7)
Observations	1,429,135	28,635	1,400,500
Adjusted R-squared	0.252	0.453	0.242
Year FE	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2

Table A4
RDD Estimates: Fund Holdings 2

The table presents estimates of the effects of index assignment on fund holdings of sample stocks using the Russell index discontinuities. The outcome variable is a dummy for whether a given mutual fund holds a given stock in December of the year post-treatment. The sample consists of stocks within a +/-100 rank window of the yearly Russell threshold from 1993-2006, and stocks within a +/-100 rank window of the yearly Russell upper and lower bands from 2007-2012. Standard errors are HAC and clustered by fund.

	(1) Hi Turnover	(2) Low Turnover	(3) Big Fund	(4) Small Fund	(5) Made Money	(6) Lost Money	(7) Low Vol	(8) High Vol
R2000	-0.001	-0.011***	-0.006***	-0.005**	-0.004**	-0.012***	-0.004**	-0.007***
	(-0.6)	(-4.3)	(-3.1)	(-2.3)	(-2.4)	(-4.2)	(-2.2)	(-3.7)
Observations	725,428	675,072	1,009,900	390,600	1,060,441	340,059	513,233	887,267
Adjusted R-squared	0.122	0.302	0.265	0.177	0.220	0.310	0.237	0.250
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ctrl Fn Deg	2	2	2	2	2	2	2	2