Excess Reconstitution-Day Volume*

Alex Chinco† and Marco Sammon‡

November 19, 2021

Abstract

Stocks that switch between the Russell 1000 and 2000 experience an economically massive spike in volume on Russell's reconstitution day. When a stock moves between these two indexes, ETFs benchmarked to either index must change their holdings. Such rebalancing typically increases an index switcher's reconstitution-day volume by 140% relative to its average daily volume during the previous year. Our main finding is that, for every share traded due to ETF rebalancing, an additional 3.1 shares get traded by somebody else. The dollar value of this excess reconstitution-day volume is on the same order of magnitude as the annual growth of Russell-benchmarked ETFs. Yet this phenomenon is entirely unexpected based on textbook asset-pricing reasoning, meaning that understanding excess reconstitution-day volume has the potential to yield theoretical dividends.

Keywords: Indexing, Passive Investing, Exchange-Traded Funds (ETFs), Russell Reconstitution Day, Trading Volume, Information-Based Asset Pricing

---

*We would like to thank Malcom Baker, Matthew Pritsker, Andrei Shleifer, and Luis Viceira for extremely helpful comments and suggestions.

Current version: http://www.alexchinco.com/excess-reconstitution-day-volume.pdf

†Baruch College, Zicklin School of Business. alexchinco@gmail.com

‡Harvard Business School. mcsammon@gmail.com

1 Introduction

On June 5th 2020, FTSE Russell made a preliminary announcement that Lumentum Holdings would move from the Russell 2000 to the 1000 on Monday June 29th. Following this announcement, everyone was fairly certain that the 13 exchange-traded funds (ETFs) benchmarked to the Russell 2000 would need to sell roughly 2.2m shares of Lumentum by market close on Friday June 26th. They could also guess that the 29 ETFs benchmarked to the Russell 1000 would need to acquire 300k shares.

A sizable chunk of Lumentum's 75m shares needed to change hands by the end of trading on June 26th. Yet, in the days following FTSE Russell's announcement, no one seemed very interested in getting this process started. Lumentum's average volume in the year leading up to Friday June 5th was 1.5m shares per day. It was still 1.5m shares per day from Monday June 8th through Thursday June 25th.

Given this initial lack of interest following FTSE Russell's June 5th announcement, you might expect that Lumentum's volume on June 26th would be at most $2.2m + 300k = 2.5m$ shares higher than usual. Authorized participants (APs) for Russell 2000 ETFs would sell 2.2m shares, and APs for Russell 1000 ETFs would buy 300k shares.

But this is not at all what happened. As shown in Figure 1, 12.6m shares of Lumentum were traded on Friday June 26th. Lumentum's excess volume,

$$\mathit{Excess\ Volume} \;=\; 8.6m \;=\; \underset{\left(\substack{\mathit{Total}\\ \mathit{volume}}\right)}{12.6m} \;-\; \underset{\left(\substack{\mathit{Avg\ volume}\\ \mathit{last\ year}}\right)}{1.5m} \;-\; \underset{\left(\substack{\mathit{ETF}\\ \mathit{volume}}\right)}{2.5m}, \tag{1}$$

was $8.6m / 2.5m \approx 3.4$ times larger than implied by ETF rebalancing alone. For every share traded on behalf of ETFs, 3.4 extra shares were traded by somebody else.

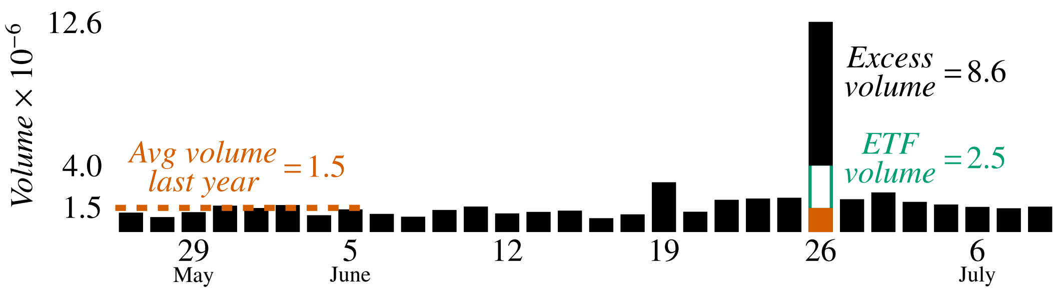

Figure 1. Lumentum's daily volume in the days around the June 26th 2020 Russell reconstitution event. Lumentum's switch was first announced on June 5th 2020.

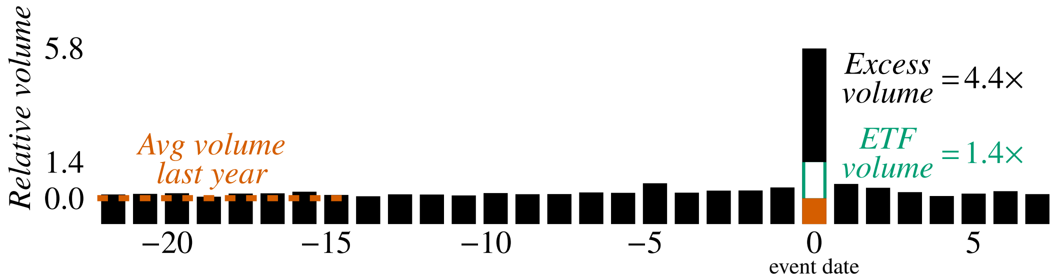

Figure 2. Volume relative to the past year for stocks that switched between the Russell 1000 and 2000 in the trading days around Russell's reconstitution date. Figure uses data from 2001 through 2020. x-axis shows event time in days relative to the reconstitution date each year, $\tau = 0$. *Relative volume*: A stock's volume on day τ divided by its average daily volume during the previous year. Relative volume = 0 when volume on day τ is equal to a stock's average daily volume over the past year. *ETF volume*: Number of shares traded on behalf of ETFs divided by a stock's average daily volume during the previous year. *Excess volume*: Volume on reconstitution day minus what would be expected due to ETF rebalancing and past experience (see Equation 1).

In this paper, we document that Lumentum's reconstitution-day experience is the rule not an exception. Figure 2 shows that, since June 2001, the average stock that has switched between the Russell 1000 and 2000 experienced an enormous spike in volume on reconstitution day just like Lumentum did on June 26th 2020.

The bar outlined in green shows that trading on behalf of Russell-benchmarked ETFs boosts the typical index switcher's volume by 140% on reconstitution day relative to its average daily volume during the previous year. However, on top of this predictable increase in volume coming from ETF rebalancing, other traders collectively exchange 4.4× an index switcher's average daily volume on reconstitution day. For every share traded by ETFs on reconstitution day, $4.4 / 1.4 \approx 3.1$ shares get traded by someone else. Spurred by this excess volume, Russell's reconstitution day is now "generally the single-biggest trading day in U.S. markets."[1]

By construction, this 4.4× increase in reconstitution-day volume is not coming from directly from trades made on behalf of ETFs. However, it is clearly linked to the existence of ETFs. Figure 3 shows that the spike in excess volume on reconstitution days starts in 2000 when Russell-benchmarked ETFs were first introduced.

[1] Rolf Agather, managing director of North America research for FTSE Russell, as quoted by Victor Reklaitis in "Why Friday could be the year's biggest trading day." *MarketWatch.* 06/26/2015.

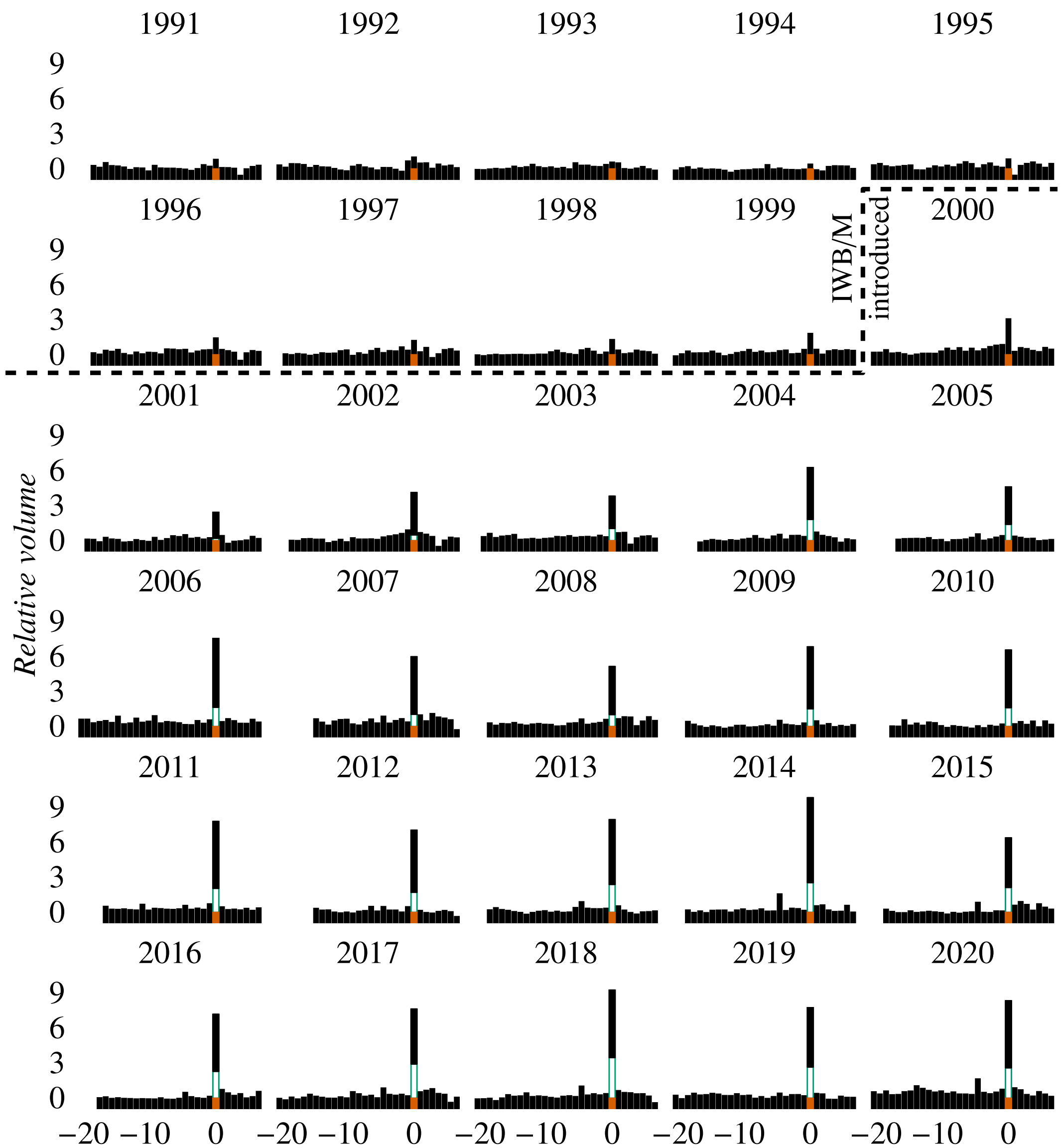

Figure 3. Average volume relative to the past year for stocks that switched between the Russell 1000 and 2000 in the days around Russell's reconstitution date. Figure uses data starting in 1991. Each panel represents data for the index-switchers in a single year. x-axis shows event time in days relative to the reconstitution date, $\tau = 0$. *Relative volume*: A stock's volume on event date τ divided by its average daily volume during the previous year minus one; i.e., $y = 0$ indicates volume equal to the average volume over the past year (red bars). Bar outlined in green: Number of shares traded due to ETF rebalancing divided by a stock's average daily volume during the previous year. The first Russell-benchmarked ETFs (iShares IWB and IWM) were introduced in May 2000. There are no Russell-benchmarked ETFs in the panels above the dashed line.

The existence of so much excess volume on Russell's reconstitution day changes the narrative about index investing. Right now, when a researcher thinks about index investing, the first thing he usually thinks about is the explosive growth of this industry. This was the topic of Robert Stambaugh's 2014 AFA address (Stambaugh, 2014). And, since 2014, the growth of the ETF industry in particular has been staggering. From 2014 to 2020, the combined assets under management (AUM) for Russell-benchmarked ETFs has increased by \$17.6*b* per year.

But if industry-wide growth of \$17.6*b* per year is worthy of investigation—i.e., something worth modeling—then the excess volume experienced by index switchers on Russell reconstitution days should really catch the researcher's eye. In a typical year, roughly 66 stocks switch between the Russell 1000 and 2000. The average dollar value each year of the excess reconstitution-day volume for just these 66 index switchers is \$14.5*b*. In other words, excess reconstitution-day volume is on the same order of magnitude as the annual growth of all Russell-benchmarked ETFs.

You would never predict the existence of this massive spike in volume based on textbook asset-pricing logic. Thus, excess reconstitution-day volume not only changes the narrative about index investing but does so in a way that has the potential to yield theoretical dividends. Explaining why there is so much excess volume on Russell reconstitution days may well reveal something deeper about how financial markets work. At the very least, it will involve developing new kinds of models. Existing information-based asset-pricing models must be tortured to fit the basic facts about excess reconstitution-day volume.

There are two broad classes of models in this literature. First, there are models built in the spirit of Grossman and Stiglitz (1980) and Admati (1985). These models study how an asset's equilibrium price aggregates the private information held by different members of a fixed population of traders in a static setting. None of this describes what takes place on Russell's reconstitution day. FTSE Russell publicly announces which stocks will switch indexes weeks ahead of time. It is not about private information held by some traders; traders learn this information from a press release not prices. Moreover, the existence of excess volume on reconstitution days implies that the population of traders is not fixed. And the excess volume experienced by index switchers on Russell's reconstitution day is not a static phenomenon. It is about one day being different.

There are also strategic-trading models in the spirit of Kyle (1985). This is the second broad class of models in the literature. These models are poorly suited to studying excess reconstitution-day volume for a different reason. The main takeaway from this breed of models is that informed traders should smooth out their demand to limit price impact. And the traders responsible for excess volume on reconstitution days do no such thing. "A record 2.37*b* shares listed on its exchange…traded in just 1.97 seconds on June 25th 2021 [at the] closing cross for this year's reconstitution."[2]

Financial economists are clever. We have every reason to believe that theorists will find some way to twist one or both of these modeling frameworks into a form that can explain the existence of excess volume on Russell's reconstitution day. However, in doing so, they will have to innovate on the underlying logic of these models. They will have to change the basic setup of a Grossman and Stiglitz/Admati-type model or reverse the main takeaway of a Kyle-type model. Either way, researchers stand to learn something important by explaining excess reconstitution-day volume.

Paper Outline

We begin our analysis in Section 2 by documenting the existence of excess volume for index switchers on Russell's reconstitution day. We give different ways of thinking about the economic magnitude of excess reconstitution-day volume. Regardless of which approach you prefer, this is an economically massive phenomenon.

There are two broad classes of information-based asset-pricing models that researchers typically use to analyze the effects index-linked investing. In Section 3, we describe how it is hard for both kinds of models to explain why there is so much extra trading in index switchers on Russell's reconstitution day.

In Section 4, we discuss a variety other considerations worth keeping in mind when examining excess reconstitution-day volume: tracking error, price changes, liquidity measures, algorithmic trading, mutual funds, closet indexing, and leveraged ETFs. None of these factors can account for our main empirical finding.

Finally, in Section 5, we look at what investors say about excess reconstitution-day volume. There is no guarantee that their reasoning will be economically sound. But, as pointed out in Chinco, Hartzmark, and Sussman (2021), knowing how investors are reasoning can be extremely useful in guiding model construction.

[2]Johnson, Steve. "FTSE Russell to review \$9*t* US index shake-up." *Financial Times.* 07/19/2021.

Related Literature

This paper adds to the literature on the economic consequences of index-linked investing (Wurgler, 2011).

People worry that the rise of indexing has made equilibrium prices less informative because fewer traders have incentives to learn about firm fundamentals. Papers such as Bond and García (2021), Buss and Sundaresan (2021), Chabakauri and Rytchkov (2021), Garleanu and Pedersen (2021), Glosten, Nallareddy, and Zou (2021), Haddad, Huebner, and Loualiche (2021), Lee (2021), and Sammon (2021) all examine the effect of the rise of indexing on price informativeness. This paper highlights a different as-yet-unappreciated consequence of indexing.

Existing papers have studied what happens when a stock gets added or dropped from an index. Harris and Gurel (1986), Beneish and Gardner (1995), Beneish and Whaley (1996), Kaul, Mehrotra, and Morck (2000), Wurgler and Zhuravskaya (2002), and Kashyap, Kovrijnykh, Li, and Pavlova (2021) look at the effect on a stock's price. Madhavan (2003) studies the price effect for stocks added to the Russell 3000. Other papers look at the effect of index additions and deletions on other attributes related to returns such as correlations and liquidity (Barberis, Shleifer, and Wurgler, 2005; Greenwood, 2008; Baker, Bradley, and Wurgler, 2011; Chang, Hong, and Liskovich, 2015; Burnham, Gakidis, and Wurgler, 2018; Brogaard, Ringgenberg, and Sovich, 2019). This paper analyzes volume rather than returns around index changes.

ETFs are now the main way that investors index. See Madhavan (2016) and Lettau and Madhavan (2018) for excellent overviews. Several recent papers have examined the consequences of ETF trading for underlying assets (Ben-David, Franzoni, and Moussawi, 2018; Da and Shive, 2018; Chinco and Fos, 2021). This paper shows that, when ETFs are present, index changes can lead to huge spikes in volume coming from other sources.

Other researchers have pointed out that passive funds can trade more often than you might think (Easley, Michayluk, O'Hara, and Tālis, 2021). And it is widely appreciated in industry that most of the \$460*b* in annual index trading takes place on the few days a year when major indexes are reconstituted (Novick, Cohen, Madhavan, Bunzel, Sethi, and Matthews, 2017). Our paper points out the puzzling nature of the excess volume that occurs on index reconstitution days. For every share traded on behalf of an ETF on one of these special days, 3.1 additional shares get traded by other investors.

2 Empirical Finding

This section presents our main empirical finding: stocks that switch between the Russell 1000 and 2000 experience a massive spike in volume on Russell's reconstitution day, and most of this volume is not coming directly from trades on behalf of ETFs. For every share of ETF trading on Russell's reconstitution day, an extra 3.1 shares get traded by other investors. Excess reconstitution-day volume only starts once ETFs begin to be benchmarked to Russell indexes in May 2000. The economic scale of this extra volume is massive. The dollar value of the excess volume experienced by index switchers (roughly 66 stocks) on Russell's reconstitution day (just 1 day a year) is on par with the annual growth of all Russell-benchmarked ETFs.

2.1 Data Description

It is not surprising that ETFs rebalance on Russell's reconstitution date. The excess volume we document is volume that cannot be explained this predictable bump in trading activity. We are as conservative as possible when estimating the scale of excess reconstitution-day volume. When faced with any uncertainty during the data construction process, we always select the choice which maximizes the amount of ETF rebalancing on the Russell reconstitution date. Thus, our estimates of excess reconstitution-day volume are lower bounds for the true amount of additional trading that takes place.

Index Membership

We obtain data on which stocks belong to the Russell 1000 and 2000 each year from FTSE Russell. There are three relevant dates to keep track for Russell reconstitution:

- #1) In late May each year, FTSE Russell ranks all stocks by market capitalization. This is the ranking date, t_{rank}.
- #2) After ranking stocks by market cap in late May, FTSE Russell keeps these rankings a secret until early June when it makes a preliminary announcement about which stocks will switch indexes. This is the announcement date, t_{ancmt}.
- #3) Finally, when FTSE Russell makes its preliminary announcement about which stocks will switch indexes, it also announces when these changes will go into effect. This is the reconstitution date, $t_{reconst}$.[3]

[3] Since 2018, FTSE Russell has reported the reconstitution date as the day a stock joined its new index.

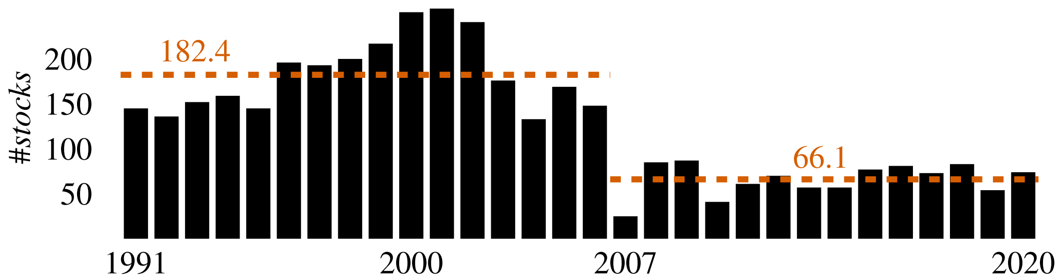

Figure 4. Number of stocks that switch from the Russell 1000 to the 2000 or vice versa each year. Horizontal dashed red lines denote average number of index switchers from 1991 through 2006 and from 2007 through 2020 respectively.

FTSE Russell's stated goal with this month-long three-step announcement process is to ensure that "stock-market trading [is] very orderly"[4] and totally predictable. Yet, FTSE Russell does not release the market cap data it uses to rank stocks in May of each year. So we follow the standard procedure to compute a proxy for the May market caps used by FTSE Russell (Chang, Hong, and Liskovich, 2015; Ben-David, Franzoni, and Moussawi, 2018), which involves matching FTSE Russell data to the CRSP/Compustat merged files based on the following variables in order of priority: (1) historical 8-digit CUSIP, (2) ticker and exchange, (3) 8-digit CUSIP, and (4) ticker. We correctly predict Russell index membership for 99.61% of stocks.

Prior to 2007, FTSE Russell assigned stocks ranked 1:1000 to the Russell 1000 index and stocks ranked 1001:3000 to the Russell 2000. However, this lead to some stocks bouncing back and forth between indexes in successive years (see Chang, Hong, and Liskovich, 2015). For example, a stock could be the 1001st largest stock in May 2002, the 995th largest stock in May 2003, and the 1010th largest stock in May 2004.

To address this issue, from 2007 onward, FTSE Russell implemented a bandwidth rule. Now, for a stock to be added to the Russell 1000 in June, its May market cap has to be larger than the smallest current Russell 1000 stock plus a margin of error, which is currently set to > 2.5% of the total cap of the Russell 3000E. This change has significantly reduced the number of index switchers each year.

Prior to 2018, the reconstitution date was listed as the last day a stock belonged to its previous index. The last day a stock belongs to an index is when all of the extra trading takes place, so we adjust the official Russell reconstitution date backward by 1 trading day in 2018, 2019, and 2020.

[4] Reklaitis, Victor. "Why Friday could be the year's biggest trading day." *MarketWatch.* 06/26/2015.

Figure 4 reports the number of stocks that switched between the Russell 1000 and 2000 each year from 1991 through 2020. Prior to 2007 when FTSE Russell used a simple market-cap ranking system, 182.4 stocks switched between the Russell 1000 and 2000 each year. Since 2007 when FTSE Russell imposed a minimum bandwidth requirement for additions and deletions, only 66.1 stocks have switched between indexes each year on average.

Stock Characteristics

We obtain data on daily trading volume, prices, returns, and total shares outstanding from the CRSP/Compustat database. Let $volume_n(t)$ denote the nth stock's volume on day t. Likewise, let $price_n(t)$, $return_n(t)$, and $\#shares_n(t)$ denote the nth stock's price, return, and number of shares outstanding on day t.

We want to compare trading volumes across stocks during the time window around the Russell reconstitution date each year. To make this possible, we will often normalize by the average daily trading volume of each stock during the 252 trading days leading up to FTSE Russell's preliminary announcement, t_{ancmt}:

$$avgVolumePastYear_n = \frac{1}{252} \cdot \sum_{\ell=1}^{252} volume_n(t_{ancmt} - \ell) \tag{2}$$

To be included in our analysis, a stock must have non-missing volume data for at least 151 of these 252 days.

The top panel of Table 1 reports summary statistics describing characteristics of stocks that switch between the Russell 1000 and 2000 on the before FTSE Russell makes its preliminary announcement, $t_{ancmt} - 1$. As expected, stocks that will be switching from the Russell 2000 to the 1000 have increased in size; whereas, those that will be switching in the opposite direction tend to have gotten smaller. Stocks at the top end of the Russell 2000 also get traded more often and have higher past returns.

ETF Holdings

Our data on ETF holdings comes from two different sources: ETF Global and Thompson S12. From 2012 through 2020, we use end-of-day ETF holdings from ETF Global. ETF Global's daily holdings data only goes back to 2012. So, from 1991 to 2011, we use quarterly holdings data from Thompson S12.

	All Switchers		1000 to 2000		2000 to 1000	
Announcement Date	Avg	Sd	Avg	Sd	Avg	Sd
Market Cap ($\$b$)	2.1	1.6	1.2	0.6	2.9	1.7
# of Shares Outstanding (m)	94.9	100.9	107.6	110.6	83.9	90.4
Average Daily Volume Past Year (m)	1.5	2.7	2.0	3.0	1.0	2.3
Return Past Year (%)	19.9	80.1	−37.7	33.2	66.9	76.4
#Obs	2,015		915		1,100	
Before Reconstitution						
# of ETFs Holding the Stock	8.8	3.3	9.0	3.7	8.5	2.7
# of Shares Held by ETFs (m)	2.1	3.2	1.8	3.6	2.4	2.6
Market Value of ETF Holdings ($\$m$)	70.5	74.7	19.6	26.5	123.3	71.8
Average Weight in ETF Portfolio (bp)	6.8	5.8	2.4	1.8	11.3	5.0
After Reconstitution						
# of ETFs Holding the Stock	11.5	3.5	10.7	3.0	12.4	3.7
# of Shares Held by ETFs (m)	2.2	3.4	3.4	4.0	1.0	1.9
Market Value of ETF Holdings ($\$m$)	42.0	36.6	38.4	20.2	45.7	47.8
Average Weight in ETF Portfolio (bp)	6.2	4.1	5.0	2.4	7.4	5.0
#Obs	634		323		311	

Table 1. Summary statistics describing characteristics and ETF holdings of stocks that switch between the Russell 1000 and 2000. All Switchers: Mean and volatility of stocks that switch between indexes. 1000 to 2000: Mean and volatility of the 1,100 stocks that drop out of the Russell 1000. 2000 to 1000: Mean and volatility of the 915 stocks that move up from the Russell 2000 to the Russell 1000. Top panel uses data from 2001 through 2020 at on the day before FTSE Russell makes its preliminary announcement, t_{ancmt}. Middle and bottom panels use data from ETF Global from 2012 through 2020. Middle panel reports results immediately before reconstitution, $t_{reconst}$. Bottom panel reports results immediately following reconstitution, $t_{reconst}$.

Let ETFs denote the set of all ETFs in the ETF Global database. We say that one of these ETFs is linked to the Russell 1000 or 2000 if it has "Russell 1000" or "Russell 2000" in its name or if ETF Global labels it as benchmarked to one of these indexes:

$$\mathsf{ETFs}_{Russell} = \left\{ f \in \mathsf{ETFs} \ : \ \text{fund } f \text{ linked to Russell 1000 or 2000} \right\} \tag{3}$$

Let $holdings_{n,f}(t)$ denote the number of shares that fund f holds of the nth stock on day t. Ideally, we could compute ETF volume for the nth stock on the Russell reconstitution day as $\sum_{f \in \mathsf{ETFs}_{Russell}} \left| holdings_{n,f}(t_{reconst} + 1) - holdings_{n,f}(t_{reconst} - 1) \right|$. However, there are sometimes reporting delays and data-entry errors in the ETF Global data. To account for these, we compute total ETF volume for stock n as the sum of the differences in ETF holdings ±5 days around the reconstitution date

$$ETFvolume_n(t) = \begin{cases} \sum_{f \in \mathsf{ETFs}_{Russell}} \left| \Delta_{\pm 5}[holdings_{n,f}(t)] \right| & \text{if } t = t_{reconst} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\Delta_{\pm 5}[holdings_{n,f}(t)] = holdings_{n,f}(t+5) - holdings_{n,f}(t-5)$. Thus, $ETFvolume_n(t)$ is defined so that all ETF-related trading occurs on the Russell reconstitution date.

Vanguard ETFs only report holdings at the end of each month because these funds are actually share classes of Vanguard mutual funds, giving them different disclosure requirements. We compute the contribution of Vanguard funds to $ETFvolume_n(t)$ as the difference between what these funds hold at the end of May each year and what they hold at the end of July.

From 1991 to 2011 or when we cannot match to ETF Global, we calculate $ETFvolume_n(t)$ using Thompson S12 as the sum of absolute differences in fund holdings of index-switcher n between March and September of each year. This approach assumes all changes in ETF portfolios occur on $t_{reconst}$ each year. In the online appendix, we show that, from 2012 to 2020 when both ETF Global and Thompson S12 data are available, $ETFvolume_n(t)$ is the same regardless of which data source we use to calculate it.

The bottom two panels of Table 1 reports summary statistics describing ETF holdings of index switchers immediately before and after reconstitution. The right-most columns of these panels reveal that switching from the Russell 1000 to the 2000 is not just the mirror image of switching from the Russell 2000 to the 1000. Because the

Russell 1000 and 2000 are value-weighted indexes, the largest stock in the Russell 2000 gets a larger weight in the Russell 2000 than the smallest stock in the Russell 1000 does in that index. So, if this stock switches from the Russell 2000 to the 1000, then ETFs benchmarked to the Russell 2000 will have to get rid of more shares than ETFs benchmarked to the Russell 1000 have to buy. The opposite will be true for any stock switching from the Russell 1000 to the 2000. In our empirical analysis below, we verify that this difference in buying-vs-selling pressure is not driving our results.

Our approach to calculating $ETFvolume_n(t)$ for each index switcher likely overstates the amount of trading on behalf of ETFs on Russell's reconstitution date, $t_{reconst}$. When using ETF Global data, we assume all trading on behalf of ETFs in a 11-day window happens on the reconstitution date itself. When using Thompson S12 data, we make the same assumption for a 6-month window. Our approach also does not account for changes in ETF holdings arising from creations/redemptions. This is a purposeful choice. If we were to relax these assumptions, we would get lower values of $ETFvolume_n(t)$, which would imply even more excess volume on Russell reconstitution days.

2.2 Excess Volume

We now document our main empirical finding about the existence of excess reconstitution-day volume. Stocks that switch between the Russell 1000 and 2000 experience a massive spike in volume on reconstitution day, and most of this volume is not coming directly from trades on behalf of ETFs. 3.1 shares get traded by other investors for each share traded on behalf of ETFs.

Variable Construction

We define the amount of excess volume experienced by index switcher n on Russell's reconstitution day as the stock's total volume minus its average daily volume during the previous year and its expected volume due to ETF rebalancing:

$$\begin{aligned} excessVolume_n(t) = volume_n(t) \\ - avgVolumePastYear_n \\ - ETFvolume_n(t) \end{aligned} \tag{5}$$

On all days except for the Russell reconstitution day, excess volume is the difference between a stock's volume and its average daily volume during the past year.

Here is the logic behind this definition. We subtract off $avgVolumePastYear_n$ to account for a stock's usual amount of volume on a typical trading day—i.e., a trading day that was not the Russell reconstitution day. We subtract off $ETFvolume_n(t)$ to remove the maximum amount of volume that might be due to ETF rebalancing. We call the remaining volume "excess volume" because it cannot be explained by usual sources.

In the analysis below, it will sometimes be helpful to study the dollar value of a stock's excess volume on Russell's reconstitution date:

$$excess\$Volume_n(t) = excessVolume_n(t) \times price_n(t_{ancmt} - 1) \tag{6}$$

We multiply excess volume by $price_n(t_{ancmt} - 1)$ rather than $price_n(t - 1)$ to ensure that our estimates of $excess\$Volume_n(t)$ are not contaminated by the market's reaction to news about a stock's impending switch between Russell indexes.

Investors tend to trade some stocks a lot more than others. So, to compare excess reconstitution-day volume across index switchers, it is necessary to normalize this spike in volume somehow. We will usually do this with a stock's average daily volume during the previous year, $avgVolumePastYear_n$, as in Figures 2 and 3.

$relVolume_n(t)$ is the difference between a stock's volume and its average daily volume last year divided by its average daily volume last year:

$$relVolume_n(t) = \frac{volume_n(t) - avgVolumePastYear_n}{avgVolumePastYear_n} \tag{7a}$$

$$= \frac{volume_n(t)}{avgVolumePastYear_n} - 1 \tag{7b}$$

$relExcessVolume_n(t)$ is a stock's excess volume divided by its average daily volume:

$$relExcessVolume_n(t) = \frac{excessVolume_n(t)}{avgVolumePastYear_n} \tag{8a}$$

$$= \frac{volume_n(t) - ETFvolume_n(t)}{avgVolumePastYear_n} - 1 \tag{8b}$$

$relVolume_n(t) = 0$ if a stock's total volume exactly equals the historical average; $relExcessVolume_n(t) = 0$ if the stock's volume equals its historical average plus the extra trading coming from ETF rebalancing.

Cross-Sectional Averages

Table 2 reports various measures of trading volume on Russell's reconstitution day, $t_{reconst}$, for stocks that switch between the Russell 1000 and 2000 indexes from 2001 to 2020. i.e., there is one observation for each index switcher each year from 2001 through 2020. Numbers in parentheses are standard errors clustered by year.

The first row reports results for the relative volume of index switchers on Russell's reconstitution day, $relVolume_n(t_{reconst})$. In column (1), we see that 5.78 extra shares of the typical index switcher get traded on reconstitution day for every share that is usually traded on a normal trading day. This point estimate corresponds to the height of the black bar at date zero in Figure 2. By comparing columns (2) and (3), we see that this spike in volume has gotten even more pronounced in recent years. Since 2007, the typical index switcher has reconstitution-day volume that is 7.43× higher than its average daily volume during the past year. Columns (4) and (5) confirm that the spike in volume exists for stocks that switch from the Russell 1000 to the 2000 and for those that move from the Russell 2000 to the 1000.

The first row in column (1) of Table 2 says that, for each share of an index switcher that was traded on a typical day during the past year, an extra 5.78 shares get traded on Russell's reconstitution day. The second row in column (1) reveals that, of these 5.78 extra shares that get traded on reconstitution day, 4.39 are traded by investors not directly tied to ETF rebalancing. In other words, ETF rebalancing is not responsible for most of the reconstitution-day bump in volume experienced by index switchers. Columns (2) and (3) show that this is not just a recent phenomenon. Columns (4) and (5) show that it holds for stocks switching in to and out of the Russell 1000.

Column (6) shows analogous results to column (3) but for stocks that almost switched between the Russell 1000 and 2000. We consider a Russell 2000 stock to be an "almost switcher" if is larger than the smallest Russell 1000 stock but not larger than the smallest Russell 1000 stock plus 2.5% of total market capitalization. We call a Russell 1000 stock an "almost switcher" if is smaller than the largest Russell 2000 stock but not smaller than the largest Russell 2000 stock minus 2.5% of total market capitalization.

The first row in column (6) shows that almost switchers experience a much smaller increase in volume on Russell's reconstitution day. Since 2007, the typical almost switcher has seen an extra 0.59 of its shares traded on Russell's reconstitution day for

	All Switchers '01-'20 (1)	All Switchers '01-'06 (2)	All Switchers '07-'20 (3)	1000 to 2000 '01-'20 (4)	2000 to 1000 '01-'20 (5)	Almost Switchers '07-'20 (6)
$relVolume_n$	5.78***	4.45***	7.43***	5.42***	6.08***	0.59***
	(0.60)	(0.72)	(0.36)	(0.67)	(0.63)	(0.05)
$relExcessVolume_n$	4.39***	3.61***	5.36***	4.07***	4.66***	0.47***
	(0.38)	(0.50)	(0.21)	(0.45)	(0.38)	(0.05)
$100 \times \left(\frac{excessVolume_n}{\#shares_n}\right)$	4.08***	2.42***	6.15***	4.18***	4.00***	0.44***
	(0.56)	(0.38)	(0.27)	(0.54)	(0.59)	(0.09)
$10^{-6} \times excessVolume_n$	4.41***	2.18***	7.17***	5.12***	3.83***	0.57***
	(0.75)	(0.47)	(0.42)	(0.89)	(0.72)	(0.15)
$10^{-8} \times excess\$Volume_n$	1.01***	0.39***	1.78***	0.53***	1.41***	0.11***
	(0.20)	(0.10)	(0.14)	(0.09)	(0.31)	(0.03)
#Obs	2,008	1,111	897	909	1,099	931

Table 2. This table reports cross-sectional averages of different measures of trading volume (rows) for different collections of stocks (columns) on Russell's reconstitution day. Column (1) reports results for all stocks that switched between the Russell 1000 and 2000 from 2001 through 2020. Columns (2) and (3) split the sample into subperiods, 2001-2006 and 2007-2020. Columns (4) and (5) split the sample by whether a stock switched from the Russell 1000 to the 2000 or from the 2000 to the 1000. Column (6) reports results for stocks that almost switched between indexes during the period from 2007 through 2020. $relVolume_n$: volume relative to average daily volume during the previous year. $relExcessVolume_n$: excess volume relative to average daily volume during the previous year. $100 \times \left(\frac{excessVolume_n}{\#shares_n}\right)$: excess volume relative to total shares outstanding (units: percent). $10^{-6} \times excessVolume_n$: excess volume without any standardization (units: millions of shares). $10^{-8} \times excess\$Volume_n$: dollar value of excess volume (units: hundreds of millions of dollars). Numbers in parentheses are standard errors clustered by year. *** indicates that a point estimate is statistically distinguishable from zero at the 1% level.

every share traded on a typical trading day during the past year. It makes sense that almost switchers would still see a small spike in volume on Russell's reconstitution day because index membership is not the only thing changing; each stock's weight in the indexes also changes. However, the second row in column (6) reveals that, even for almost switchers, $0.47 / 0.59 \approx 80\%$ of the extra volume experienced on Russell's reconstitution day cannot be attributed to ETF rebalancing.

The remaining rows document that these main findings are robust to using alternative measures of excess reconstitution-day volume. In the third row, we report results where we normalize an index switcher's excess volume by its number of shares outstanding rather than its trading volume during the past year, $100 \times \mathit{excessVolume}_n(t_{reconst}) / \#\mathit{shares}_n$. The results are largely unchanged.

In the fourth row, we show results for excess reconstitution-day volume without doing any standardization, $\mathit{excessVolume}_n(t_{reconst})$. The point estimates in this row have units of millions of shares. In the fifth row, we show results for excess dollar volume on Russell reconstitution days, $\mathit{excess\$Volume}_n(t_{reconst})$. The point estimates in this row have units of millions of dollars. The findings in the fourth and fifth rows should be interpreted with caution since index switchers can have vastly different baseline volume levels. Nevertheless, it is comforting that this analysis still paints the same picture.

Panel Regressions

The most striking thing about Figure 2 is the difference between an index switcher's volume on Russell's reconstitution day and its volume on the days immediately before and after. In the days leading up to reconstitution, the typical index switcher has trading activity that looks no different from the past year. Then, all of the sudden, things change on the reconstitution date. And, just as suddenly, they go back to normal the day after.

This brings up an important distinction: our analysis does not involve comparing the stocks that switch between the Russell 1000 and 2000 to the stocks that do not switch. Excess reconstitution-day volume is a puzzle worthy of future theoretical attention regardless of whether index switchers and almost switchers are similar to one another. Thus, it is not subject to the well-known critiques of this sort of comparison (Appel, Gormley, and Keim, 2020). What is special about excess reconstitution-day volume is a) that there is so much of it and b) that it only occurs on Russell's reconstitution day. There is no excess volume on the days immediately before and after.

To highlight this point, we estimate panel regressions to capture how trading activity on Russell's reconstitution day differs from trading activity on neighboring days. We define event time, $\tau = t - t_{reconst}$, as the number of days since the Russell reconstitution date in a given year. $\tau = 0$ corresponds to Russell's reconstitution day; $\tau < 0$ for the days leading up to reconstitution day; $\tau > 0$ for the days after reconstitution day. Each year, we include data on index switchers starting the day after FTSE Russell's preliminary announcement, $t_{ancmt} + 1$, and ending 7 trading days after Russell's reconstitution date, $t \in \{t_{ancmt} + 1, \ldots, t_{reconst}, \ldots, t_{reconst} + 7\}$.

Table 3 shows results of panel regressions of the form below

$$relExcessVolume_n(\tau) = \hat{\alpha} + \hat{\beta} \cdot isAfter(\tau) + \hat{\gamma} \cdot isEventDate(\tau) + \hat{\varepsilon}_n(\tau) \qquad (9)$$

using data on index switchers and almost switchers each year. Table 2 only uses data on Russell's reconstitution day; by contrast, Table 3 also uses data in the trading days around Russell's reconstitution day.

The intercept, $\hat{\alpha}$, captures how much higher/lower a stock's volume tends to be during the trading days after FTSE Russell's preliminary announcement but prior to reconstitution taking place. If index-switcher volume during this period tends to be the same as during the past year, then we will estimate $\hat{\alpha} = 0$. If volume tends to be higher in the run up to reconstitution, then we will estimate $\hat{\alpha} > 0$.

$isAfter(\tau) = 1_{\{\tau>0\}}$ is an indicator variable that is one if an observation takes place after Russell's reconstitution date. The coefficient on this variable, $\hat{\beta}$, captures the amount of extra volume that index switchers realize in the 7 trading days following the Russell reconstitution date. If $\hat{\beta} > 0$, then index switchers tend to have more volume after reconstitution as they did before reconstitution.

$isEventDate(\tau) = 1_{\{\tau=0\}}$ is an indicator variable that is one if an observation occurs on Russell's reconstitution date. The coefficient $\hat{\gamma}$ captures the amount of excess volume that index switchers realize on Russell's reconstitution date that cannot be explained by trading on behalf of ETFs. If $\hat{\gamma} > 0$, then index switchers have unexplained volume on Russell's reconstitution day. Our null hypothesis is that the extra trading index switchers experience on Russell's reconstitution day is due to ETF rebalancing. If this were true, then we should estimate $\hat{\gamma} \approx 0$.

Dependent Variable: $relExcessVolume_n(\tau)$

	All Switchers '01-'20 (1)	All Switchers '01-'06 (2)	All Switchers '07-'20 (3)	1000 to 2000 '01-'20 (4)	2000 to 1000 '01-'20 (5)	Almost Switchers '07-'20 (6)
intercept	$0.27^{\star\star\star}$	$0.27^{\star\star\star}$	$0.27^{\star\star\star}$	0.10	$0.41^{\star\star\star}$	0.07
	(0.03)	(0.04)	(0.05)	(0.07)	(0.04)	(0.06)
$isAfter(\tau)$	−0.00	−0.04	0.04	$0.08^{\star}$	$-0.07^{\star\star}$	−0.07
	(0.03)	(0.03)	(0.06)	(0.05)	(0.03)	(0.05)
$isEventDate(\tau)$	$4.12^{\star\star\star}$	$3.34^{\star\star\star}$	$5.09^{\star\star\star}$	$4.00^{\star\star\star}$	$4.25^{\star\star\star}$	$0.40^{\star\star\star}$
	(0.36)	(0.48)	(0.20)	(0.42)	(0.39)	(0.06)
Adj. R^2	23.9%	14.6%	40.3%	19.5%	28.8%	1.2%
#Obs	36,061	20,200	15,861	16,268	19,793	16,620

Table 3. Each column reports the results of a different version of the panel regression in Equation (9). The regressions in columns (1)-(5) use data about stocks that switch between the Russell 1000 and 2000 indexes. Column (6) uses data about stocks that almost switched but did not. For each index switcher (or almost switcher) we include data starting the day after FTSE Russell's preliminary announcement, $t_{ancmt} + 1$, and ending 7 trading days after Russell's reconstitution date, $t \in \{t_{ancmt} + 1, \ldots, t_{reconst}, \ldots, t_{reconst} + 7\}$. $\tau = t - t_{reconst}$ denotes event time during this window each year. The dependent variable is relative excess volume, $relExcessVolume_n(\tau)$, at event time τ. $isAfter(\tau) = 1_{\{\tau>0\}}$ is an indicator variable that is one if an observation occurs after Russell's reconstitution date. $isEventDate(\tau) = 1_{\{\tau=0\}}$ is an indicator variable that is one if an observation occurs on Russell's reconstitution date. Numbers in parentheses are standard errors clustered by year. $\star$, $\star\star$, and $\star\star\star$ denote estimates that are statistically distinguishable from zero at the 10%, 5%, and 1% levels.

The first row in Table 3 shows the point estimates for the intercept in Equation (9). Looking across columns, we see that index switchers tend to have daily trading volumes during the lead up to Russell's reconstitution day that are roughly 27% higher than they experienced during the past year. And, the second row shows that this increase in volume carries over to the 7 trading days following reconstitution, $\hat{\beta} \approx 0$.

The third row in Table 3 shows the amount of excess volume on Russell's reconstitution day. The height of the spike in relative excess volume on reconstitution day was 4.39 in Figure 2. And the sum of the intercept and the coefficient on $\mathit{isEventDate}(\tau)$ in third row of column (1) is $0.27 + 4.12 = 4.39$. This finding formalizes our earlier observation that roughly $4.39 / 5.78 \approx 76\%$ of the additional volume that index switchers experience on reconstitution day cannot be explained by trading on behalf of ETFs.

Columns (2) and (3) verify that this result holds in both the 2001-2006 and 2007-2020 subperiods. Columns (4) and (5) confirm that it holds up regardless of whether an index switcher is moving from the Russell 1000 to the 2000 or vice versa. And column (6) shows that, when looking at stocks which almost switch between Russell indexes but do not, the amount of excess reconstitution-day volume is an order of magnitude smaller, mirroring our findings in column (6) of Table 2.

Existence of ETFs

ETF rebalancing cannot account for the excess reconstitution-day volume that index switchers experience. But the existence of ETFs does seem tied to this phenomenon. Table 4 shows that there was little excess volume on Russell reconstitution days before ETFs were first benchmarked to the Russell 1000 and 2000.

Each column reports the results of a different regression. All regressions use yearly data for $y \in \{1991, \ldots, 2020\}$. Let $\mathsf{N}(y)$ denote the set of stocks that switch between the Russell 1000 and 2000 in year y. The dependent variable is $\overline{\mathit{relExcessVolume}}(y)$, the average excess volume on Russell's reconstitution day in year y:

$$\overline{\mathit{relExcessVolume}}(y) = \frac{1}{|\mathsf{N}(y)|} \cdot \sum_{n \in \mathsf{N}(y)} \mathit{relExcessVolume}_n(t_{reconst}) \tag{10}$$

$\overline{\mathit{relExcessVolume}}(y)$ reflects the height of the black bar on day $\tau = 0$ in Figure 3 minus the heights of the green and red bars on that day.

Dependent Variable: $\overline{relExcessVolume}(y)$

	(1)	(2)	(3)
intercept	1.05⋆⋆⋆ (0.14)	1.05⋆⋆⋆ (0.15)	1.05⋆⋆⋆ (0.15)
isAfter2000(y)	3.78⋆⋆⋆ (0.31)	2.88⋆⋆⋆ (0.46)	3.04⋆⋆⋆ (0.40)
$\{\#RussellETFs(y)\,/\,100\} \times isAfter2000(y)$		6.08⋆⋆⋆ (1.96)	
$\{aumRussellETFs(y)\,/\,\$100b\} \times isAfter2000(y)$			1.08⋆⋆⋆ (0.33)
Adj. R^2	73.7%	80.3%	80.8%
#Obs	30	30	30

Table 4. Each column reports the results of a different regression involving data from years $y = 1991, \ldots, 2020$. The dependent variable in all regressions, $\overline{relExcessVolume}(y)$, is the average excess volume experienced by index switchers in year y. $isAfter2000(y) = 1_{\{y>2000\}}$ is an indicator variable for years that Russell-benchmarked ETFs exist. $\#RussellETFs(y)$ is the number of Russell-benchmarked ETFs that exist in year y, and $aumRussellETFs(y)$ is the combined assets under management (AUM) of these ETFs. Numbers in parentheses are standard errors. ⋆⋆⋆ denotes estimates that are statistically distinguishable from zero at the 1% level.

The Russell 1000 and 2000 were created in 1984. However, the first ETFs benchmarked to the Russell 1000 and 2000 were not introduced until May 2000, when iShares listed IWB and IWM, which are benchmarked to the Russell 1000 and 2000 respectively. Prior to 2000, the Russell 1000 and 2000 were still widely used. There just were not any ETFs linked to these indexes.

Let $isAfter2000(y) = 1_{\{y>2000\}}$ be an indicator variable for years that Russell-benchmarked ETFs exist. Column (1) of Table 4 shows that, prior to the advent of Russell-benchmarked ETFs, index switchers still realized a jump in excess volume on reconstitution day. However, this jump in volume was only 105% of a stock's average daily volume during the past year. Once Russell-benchmarked are introduced, though, this spike is nearly five times as large, growing from 105% of average daily volume to 105% + 378% = 483% of average daily volume.

Column (2) shows that the increase in excess reconstitution-day volume following the advent of Russell-benchmarked ETFs got even larger as more and more Russell-benchmarked ETFs were introduced. $\#RussellETFs(y)$ denotes the number of Russell-benchmarked ETFs in year y. The coefficient on this variable in column (2) implies that, when there is one more Russell-benchmarked ETFs, index switchers' excess volume on reconstitution days increases by roughly 6%pt.

Finally, column (3) in Table 4 shows a similar result, only looking at Russell-benchmarked ETFs' assets under management (AUM) rather than their number. $aumRussellETFs(y)$ denotes the combined AUM of all Russell-benchmarked ETFs in year y. The coefficient on this variable in column (3) indicates that, for each additional billions dollars invested in Russell-benchmarked ETFs, index switchers' excess volume on reconstitution days increases by around 1%pt.

2.3 Economic Magnitude

In the last part of this section, we show that the economic scale of excess reconstitution-day volume is absolutely massive. We think about it in two different ways. Our first approach is to compare the excess volume experienced by index switchers on Russell's reconstitution day to the growth of the Russell-benchmarked ETF industry as a whole. The numbers are on the same order of magnitude, and in recent years there has typically been more excess reconstitution-day volume. Our second approach involves asking the question: Would an index switcher still have excess reconstitution-day volume if we assumed that all mutual fund and ETF trading for that stock in Thompson S12 during the entire year took place on Russell's reconstitution day? It turns out that, for most index switchers, the answer is, "Yes."

Relative To Industry-wide Growth

If there is one thing that every financial economist knows about ETFs, it is that the ETF industry has experienced explosive growth in the past decade. "There are now more indexes than stocks."[5] This stylized fact is the main thing that asset-pricing researchers think is worth modeling about the ETF industry. For example, papers such as Bond and García (2021), Buss and Sundaresan (2021), Chabakauri and Rytchkov (2021), Garleanu and Pedersen (2021), Glosten, Nallareddy, and Zou (2021), Haddad, Huebner, and Loualiche (2021), Lee (2021), and Sammon (2021) all want to understand how the rise of indexing has impacted price informativeness. So why not compare the size of excess reconstitution-day volume to the main thing about ETFs that researchers want to explain?

[5] Bernstein, Sanford. "There Are Now More Indexes Than Stocks." *Bloomberg News.* 05/12/2017.

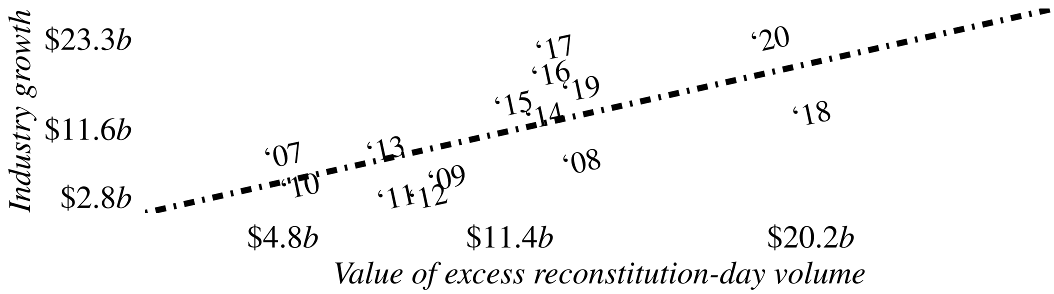

Figure 5. This figure uses data covering the 14 years from 2007 through 2020. y-axis: average growth in the combined assets under management of all Russell-benchmarked ETFs during the five-year period ending in year y, $avgIndustryGrowth(y)$, as defined in Equation (11). x-axis: dollar value of the excess reconstitution-day volume experienced by all index switchers in year y, $totalExcess\$Volume(y)$, as defined in Equation (12). Tick marks denote the min, mean, and max values of each variable. Numbers of the form 'YY denote an observation for that year. Diagonal dashes represent the 45° line.

Let $avgIndustryGrowth(y)$ denote the average growth of the combined assets under management of all Russell-benchmarked ETFs over the five-year period ending in year y:

$$avgIndustryGrowth(y) = \frac{1}{5} \cdot \sum_{\ell=0}^{4} \Delta aumRussellETFs(y - \ell) \qquad (11)$$

We estimate that, from 2007 through 2020, the amount of money invested in Russell-benchmarked ETFs grew by an average of \$11.6b per year as indicated by the middle tick mark on the y-axis in Figure 5. This number represents the combined effect of investor flows and capital appreciation over the past five years. And this statistic jumps to \$17.6b per year if we look only at more recent data from 2014 to 2020.

We compare this rolling average industry-wide growth rate to the combined dollar value of excess reconstitution-day volume experienced by all index switchers in year y:

$$totalExcess\$Volume(y) = \sum_{n \in \mathsf{N}(y)} excess\$Volume_n(t_{reconst}) \qquad (12)$$

While $avgIndustryGrowth(y)$ represents the increase in Russell-benchmarked ETFs' entire holdings over the course of an entire year, $totalExcess\$Volume(y)$ represents the dollar value of the excess volume of only around 66.1 stocks on a single day in late June. Nevertheless, the average value of excess reconstitution-day volume from 2007 through

2020 was \$11.4*b* as indicated by the middle tick mark on the x-axis in Figure 5. As reported in the introduction, if we look only at data starting in 2014, the average value of excess reconstitution-day volume was \$14.5*b*.

Whenever an observation sits below the 45° dashed line in Figure 5, it implies that the dollar value of the excess volume experienced by index switchers on Russell's reconstitution day exceeded the average recent growth of Russell-benchmarked ETFs. This happened in 6 of the 14 years from 2007 through 2020—from 2008 through 2012 and then again in 2018.

The sheer scale of excess reconstitution-day volume changes the narrative about index investing. Right now, when researchers think about index investing, the first thing they think about is the growth of this industry. But if researchers view \$11.6*b* per year industry-wide growth as noteworthy—i.e., something worth modeling—then excess reconstitution-day volume must also warrant their attention. It is on the exact same order of magnitude as the growth of the entire ETF industry. If one flow of cash is economically important, then so is the other.

As A Fraction Of All Rebalancing

Due to data entry errors, ETF rebalancing sometimes gets entered into ETF Global a day or two late. So, when constructing $ETFvolume_n$ in Equation (4), we summed over changes in ETF holdings from 5 trading days before Russell's reconstitution date to 5 trading days after. This approach overstated the amount of ETF volume on Russell's reconstitution date, making it harder for us to find excess volume that cannot be accounted for by ETF rebalancing. Yet, we still found excess reconstitution-day volume.

Our second approach to assessing the economic magnitude of this excess volume involves asking: How much farther would we have to go when constructing $ETFvolume_n$ to eliminate it? For instance, suppose you were worried that we were missing some Russell-benchmarked ETFs when calculating $ETFvolume_n$. No matter. There were 66 stocks that switched between the Russell 1000 and 2000 in June of 2020 which we could match to ETF Global. When we calculate $ETFvolume_n$ by summing over all ETFs in the ETF Global database, $f \in \mathsf{ETFs}$, rather than just the Russell-linked funds, $f \in \mathsf{ETFs}_{Russell}$, every one of these 66 stocks still has excess reconstitution-day volume.

We can even take things a step further. Suppose we use the Thompson S12 data and calculate the change in every fund's holdings from March to September 2020, a 6-month window. This involves summing over all funds, $f \in \mathsf{AllFunds}$, not just ETFs. Even still, when we calculate "$\mathit{ETFvolume}_n$" this way for index switchers in June 2020, it turns out that 2 of these 66 stocks have excess volume!

It is always possible that the explanation for any new and surprising empirical finding is a data error. However, explaining our results this way would require an epic amount of data mislabeling. It would systematically misreporting the amount of index-related trading activity on Russell's reconstitution date by an amount equal to all the trading done by all the investment funds over a 6-month window.

3 Standard Models

There are two broad classes of information-based asset-pricing models that researchers use when studying index funds. Researchers use models built in the spirit of Grossman and Stiglitz (1980) and Admati (1985) to study how the rise of indexing has affected price levels and investor welfare. And they use models in the spirit of Kyle (1985) to study how investors strategically respond to large institutional trades.

Neither framework captures the economics of excess reconstitution-day volume. This phenomenon takes place outside of the basic setup of a Grossman and Stiglitz/Admati-type model. And it directly contradicts the main take-home message of Kyle-type models. Thus, we suspect that a good explanation for excess reconstitution-day volume will require a new class of model emphasizing a novel economic trade off.

We could be wrong. In the future, a theorist may discover a way to twist one of these frameworks two to account for excess reconstitution-day volume. But such a twist will require either altering the basic setup of Grossman and Stiglitz/Admati-type models or finding a reason to reverse the main takeaway of Kyle-type models on one day each year. So, even in this scenario, researchers stand to learn something new by resolving the puzzle of excess reconstitution-day volume.

3.1 Learning From Prices

The modeling framework introduced in Grossman and Stiglitz (1980) and Admati (1985) studies how an asset's equilibrium price aggregates the private information held by different members of a fixed population of traders in a static setting.

Basic Setup

There is a fixed population of infinitesimal investors, $i \in \mathsf{Investors}$. Usually, it is the case that $\mathsf{Investors} = [0, 1]$. There is a single round of trading. The ith investor has $Demand_n^i$ for the nth risky asset. For each share that he holds after trading concludes, the this investor gets $Payout_n$ dollars.

The ith investor chooses his demand for each risky asset, $\{Demand_n^i\}_{n=1}^N$, to maximize his expected utility from terminal wealth, $Wealth^i$, given equilibrium prices, $\{Price_n\}_{n=1}^N$:

$$\max \mathbb{E}^i\Big[- e^{-\rho \cdot Wealth^i} \,\Big|\, Price_1, \ldots, Price_N \Big] \tag{13}$$

ρ captures investors' level of risk aversion. The ith investor's terminal wealth is the sum of his payouts from all risky-asset holdings, $Wealth^i = \sum_{n=1}^N Payout_n \cdot Demand_n^i$.

Different investors have different beliefs about future payout. The ith investor believes that $Payout_n \sim \text{Normal}(\mu_n^i, [\sigma_n^i]^2)$. In Grossman and Stiglitz-type models, these differences in beliefs stem from ex ante identical investors making different information-acquisition choices. In Admati-type models, investors are endowed with different information from the start. The end result is the same either way.

The supply of shares for each risky asset is assumed to be random:

$$Supply_n \sim \text{Normal}(\mu_{Supply}, \sigma_{Supply}^2) \tag{14}$$

When solving Equation (13), investors do not know the exact value of $Supply_n$. But investors do know the distribution of $Supply_n$ in these models.

The ith investor's demand for the nth risky asset is a function of the asset's price, $Demand_n^i(Price_n)$. And, collectively, all investors must hold exactly $Supply_n$ shares of the nth risky asset at the equilibrium price, $Price_n$, for markets to clear:

$$Supply_n = \int_{i \in \mathsf{Investors}} Demand_n^i(Price_n) \cdot \mathrm{d}i \tag{15}$$

Investors' private information and supply shocks both affect market clearing. So when the ith investor sees an unexpectedly high price, he can infer that other investors received good news, $\int_{i' \neq i} \mu_n^{i'} \cdot \mathrm{d}i' > \mu_n^i$, or that there is low supply, $Supply_n < \mu_{Supply}$. Thus, equilibrium prices partially reveal the private information of other investors.

Main Takeaway

This modeling paradigm is useful for studying how changes in the distribution of private information across investors affect price levels and investors' welfare.

For example, if you were to re-run one of these static models over and over again, the realized price of the nth risky asset, $Price_n$, would be different each time. But the asset's average price across simulations

$$\mathbb{E}[\,Price_n\,] \propto \int_{i \in \text{Investors}} \mathbb{V}\text{ar}^i[\,Payout_n\,|\,Price_n\,]^{-1} \cdot \mathrm{d}i \tag{16}$$

should be proportional to the typical investors' residual uncertainty about the asset's future payout after seeing its equilibrium price. Thus, Grossman and Stiglitz/Admati-type models teach us that a risky asset's price will be higher when the typical investor thinks holding it is less risky, either because he has better private information or because he can learn a lot from the asset's equilibrium prices.

When an investor switches from active trading to passive indexing, it stands to reason that he has less incentive to acquire information about the assets he is holding. Why bother if you are just going to hold the index regardless of what you learn? Thus, explosive growth in indexing over the past two decades sure seems like a change in the distribution of private information across investors.

Based on this logic, there is now a booming literature using the Grossman and Stiglitz/Admati framework to analyze the economic consequences of indexing. A list of current working papers and 2021 publications in this mold includes Bond and García (2021), Buss and Sundaresan (2021), Chabakauri and Rytchkov (2021), Garleanu and Pedersen (2021), Glosten, Nallareddy, and Zou (2021), Haddad, Huebner, and Loualiche (2021), Kashyap, Kovrijnykh, Li, and Pavlova (2021), Lee (2021), and Sammon (2021).

Why It Does Not Fit

The rise of index investing in recent years is an important change in financial markets. And there are many aspects of this rise that Grossman and Stiglitz/Admati-type models are useful for studying. However, excess reconstitution-day volume is not one of them. The basic setup of this class of models has nothing to do with the excess volume experienced by index switchers on Russell's reconstitution day. The details of this puzzling phenomenon run contrary to the model primitives outlined above.

Here is a partial list of the ways that the modeling framework fails to fit the facts about what happens in late June each year:

#1) Excess reconstitution-day volume is not about a change in the distribution of private information across investors. FTSE Russell publicly announces which stocks will switch indexes weeks ahead of time.

#2) Excess reconstitution-day volume is not about learning from prices. The relevant information gets revealed in a press release. Investors certainly learn from prices in the days and weeks around Russell's reconstitution event. But this is also true the rest of the year. That is not what makes Russell's reconstitution day special.

#3) Excess reconstitution-day volume implies that the population of investors is not fixed. On one special day in late June each year, the collection of investors who are interested in trading index switchers suddenly changes.

#4) For the same reason, excess reconstitution-day volume is not a static phenomenon. It is about one day a year being different.

Deeper Issue

It could be that that this economically massively phenomenon simply occurs for reasons that are completely unrelated to those captured by Grossman and Stiglitz/Admati-type models. However, we are documenting an economically massive spike in volume that researchers were previously unaware of. So it is a bit worrying that the conclusions drawn from Grossman and Stiglitz/Admati-type models are predicated on the assumptions that investors know the asset-supply distribution (Equation 14) and understand the implications of market clearing (Equation 15). How much can investors be learning from the fact that equilibrium prices must clear the market for risky assets if such a large spike in volume can go unnoticed?

True, excess reconstitution-day volume might only have gone unnoticed by researchers and not investors. As we document in Section 5, investors are clearly aware that index-switcher volume tends to spike on Russell's reconstitution day. But, as far as we can tell, there is no clear story, even among investors, about where this extra trading activity comes from. There does not seem to be broad awareness among investors that most of the extra volume on reconstitution day is not coming directly from index-fund rebalancing. If this is right, then excess reconstitution-day volume represents a much deeper issue for the Grossman and Stiglitz/Admati paradigm.

3.2 Smoothing Out Demand

The modeling framework introduced by Kyle (1985) studies how investors strategically respond to large institutional trades.

Basic Setup

There is a single informed investor who participates in a sequence of $T \geq 1$ auctions for a single risky asset. This risky asset has a terminal payout of $Payout \sim \text{Normal}(0, 1)$ dollars per share, and the informed investor learns *Payout* before trading in the first auction. The equilibrium price in each auction is $Price(t)$ dollars per share.

The informed investor chooses $Demand(t)$ to maximize his total profits:

$$\underbrace{\{\, Payout - Price(t)\,\} \cdot Demand(t)}_{\substack{\text{Profit from trading on informational} \\ \text{advantage in just the } t\text{th auction.}}} + \underbrace{\mathbb{E}_t[\, ContinuationValue(t+1)\,]}_{\substack{\text{Expected profits from trading on informational} \\ \text{advantage in all future auctions } \{\, t+1, \ldots, T\,\}.}} \tag{17}$$

$\mathbb{E}_t[\cdot]$ is the informed investor's beliefs about future outcomes given the pricing rule. $ContinuationValue(t+1) = \sum_{u=t+1}^{T} \{\, Payout - Price(u)\,\} \cdot Demand(u)$ denotes his future profits on all positions acquired after period t given this same pricing rule.

And where does this pricing rule come from? A market maker sets the risky asset's price after observing aggregate demand, $AggDemand(t) = Demand(t) + Noise(t)$:

$$Price(t) = \mathbb{E}\big[\, Payout \,\big|\, \{AggDemand(s)\}_{s=1}^{t}\,\big] \tag{18}$$

The demand noise each period is a random shock, $Noise(t) \overset{\text{IID}}{\sim} \text{Normal}(0, \sigma^2)$, that the informed investor can use to camouflage his demand. The market maker cannot separately observe $Demand(t)$ or $Noise(t)$ on its own. So she cannot immediately jump to the conclusion that the informed investor knows that $Payout > 0$ the moment she sees higher-than-expected aggregate demand.

Still, conditional on observing $AggDemand(t) > 0$, it is more likely that the informed investor knows that $Payout > 0$ and has chosen $Demand(t) > 0$ to take advantage of this knowledge. So the market maker will raise the price a bit whenever she sees higher than expected aggregate demand. The amount by which the price will increase in response to a tiny increase in informed demand is given by $\lambda(t) = \frac{\partial Price(t)}{\partial Demand(t)} > 0$. The informed investor internalizes this price impact when choosing his demand.

Main Takeaway

Suppose the informed investor knows that $Payout > Price(t)$ and so buys an additional share at time t, $Demand(t) \rightarrow Demand(t) + 1$. This increase in demand will cause the market maker to increase the price of the risky asset by $\lambda(t)$ dollars per share.

Such a price impact will reduce the informed investor's profits in two different ways. First, it will reduce his profits on every share purchased at time t. Instead getting $Payout - Price(t)$ on each share he was originally planning on buying at time t, he will only get $Payout - Price(t) - \lambda(t)$ per share. Second, it will reduce the informed investor's profits on all shares he purchases in every future auction at times $\{ t+1, \ldots, T \}$. The $\lambda(t)$ dollar per share price increase in auction t is permanent in this model. So the continuation value will drop to $\sum_{u=t+1}^{T} \{ Payout - Price(u) - \lambda(t) \} \cdot Demand(u)$.

As a result, the informed investor will trade less aggressively to preserve his informational advantage. This is the main takeaway from Kyle-type models. The second line of reasoning implies that the informed investor has a strong incentive to smooth out his demand over time in Kyle-type models. When there are many remaining trading periods, the informed trader should space out his demand over time to take advantage of the noise shocks in each period. The first line of reasoning implies that, even in the final trading period, the informed investor will still throttle his trading. There are no future trading periods to worry about at time T. But the informed investor has to pay $\lambda(T)$ extra dollar for each of the $Demand(T)$ shares he acquires at time T.

Why It Does Not Fit

Grossman and Stiglitz/Admati-type models do not capture the economics of excess reconstitution-day volume because they are built on the wrong assumptions. Kyle-type models do not fit for a different reason: excess reconstitution-day volume violates the main takeaway from these models.

At 6pm on June 5th 2020, authorized participants (APs) of Russell-benchmarked ETFs knew that Lumentum would be switching from the Russell 2000 to the 1000 at the end of trading on June 26th. The logic of a Kyle-type model says that these traders should have gradually resolved their positions over the next few weeks. Yet, when we look at the data, we see that they did all their trading in the last few minutes of June 26th. Kyle-type models also offer no reason why 3.1 extra shares of Lumentum should get traded for every share traded on behalf of ETFs.

4 General Discussion

We just saw that the two standard modeling frameworks for thinking about indexing cannot account for excess reconstitution-day volume. The basic setup of Grossman and Stiglitz/Admati-type models means that these models cannot capture the economics of this phenomenon. And the existence of excess reconstitution-day volume violates the main takeaway from Kyle-type models.

In this section, we now discuss a potpourri of other considerations that might play a role in the excess volume experienced by index switchers on Russell's reconstitution day. While some of these factors can help explain why ETF rebalancing might lead to higher reconstitution-day volume, none can explain why other as-yet-identified traders trade an extra 3.1 shares for every share traded on behalf of ETFs. There are good reasons why an index switcher might experience higher volume on Russell's reconstitution day. We know of no good reasons why they might experience so much excess volume.

4.1 Tracking Error

If an ETF's share price deviates from the value of its holdings, investors can exploit this gap by creating or redeeming shares. This fact can explain why Russell-benchmarked ETFs might wait to do all of their rebalancing on Russell's reconstitution day. If they were to do it gradually beforehand as suggested by a Kyle-type model, then it would create an arbitrage opportunity that other investors could exploit.

It is important to point out that this is not a mark in favor of using a Kyle-type model to analyze ETF rebalancing. It suggests that the model does not capture the relevant economic trade off. If the desire to minimize tracking error swamps the desire to minimize price impact, then researchers should not be looking at this market setting through the lens of a Kyle-type model.

We are not saying that Kyle (1985) is a bad model. We absolutely adore this model. It is just not the right model to apply in this setting. The key economic trade off in Kyle (1985) is not the first-order concern in our setting.

More importantly, our main result is not about ETF rebalancing activity on Russell's reconstitution day. Our main result is about all of the extra reconstitution-day volume on top of the boost in reconstitution-day volume coming from ETF rebalancing. Our paper is called "Excess Reconstitution-Day Volume" not "ETF Reconstitution-Day Volume".

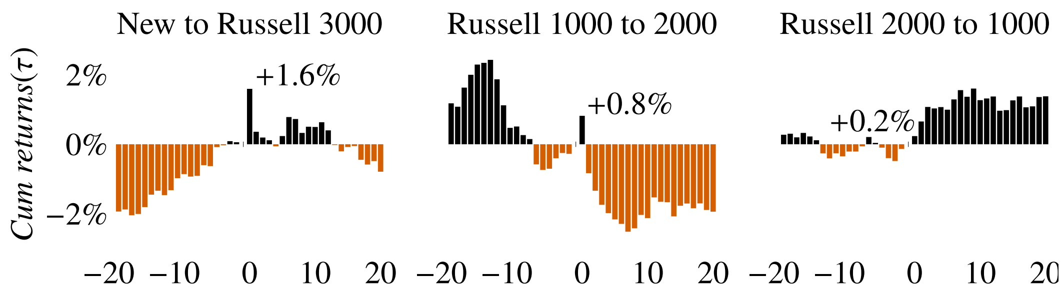

Figure 6. Each panel shows the cumulative abnormal returns of a different portfolio of index switchers. x-axis counts trading days relative to Russell's reconstitution day each year. y-axis reports each portfolio's cumulative return minus the market. We have normalized each portfolio's cumulative returns to be 0% at time $\tau = -1$. "New to Russell 3000" reports results for stocks that are added to the Russell 3000 on reconstitution day. "Russell 1000 to 2000" reports results for stocks that move from the Russell 1000 to the 2000 on reconstitution day. "Russell 2000 to 1000" reports results for stocks that move in the opposite direction (from the Russell 2000 to the 1000) on reconstitution day. Numbers of the form +#.#% represent each portfolio's return minus the market on Russell's reconstitution day.

It is true that ETFs are loath to deviate from their stated benchmarks. And this fact can explain why trading on behalf of Russell-benchmarked ETFs boosts the typical index switcher's volume by 139% on reconstitution day relative to its average daily volume during the previous year. But it cannot explain why other traders collectively exchange an additional 439% of an index switcher's average daily volume.

4.2 Price Changes

So far we have been talking exclusively about the trading volume of stocks that switch between the Russell 1000 and 2000 on Russell's reconstitution day. You might be wondering: "What happens to the prices of these stocks?"

Indeed, using data on Russell reconstitution days from 1996 through 2001, Madhavan (2003) finds that stocks which get added to the Russell 3000 experience roughly a 2% return minus the market on reconstitution day. The left panel of Figure 6 replicates this main finding using our data which runs from 2001 through 2020. We see a 1.6% abnormal return for Russell 3000 additions on reconstitution day.

This close quantitative match is both reassuring and surprising. It is reassuring because we are able to closely replicate the original point estimate. It is surprising because we are able to closely replicate the original point estimate even though we are using more recent data where indexing is much more important.

Things get even weirder when we look at the abnormal returns of stocks that switch between the Russell 1000 and 2000. Table 1 shows that just before reconstitution, the typical stock that moved up from the Russell 2000 to the 1000 made up 0.11% of the Russell 2000 index. By contrast, the typical stock that dropped down from the Russell 1000 to the 2000 made up only 0.02% of the Russell 1000 index. Thus, when a stock moves up from the Russell 2000 to the 1000, ETF rebalancing will tend to result in selling pressure on Russell's reconstitution day. Whereas, when a stock moves down to the Russell 2000, it will experience buying pressure due to ETF rebalancing.

If we were to go by the logic of Kyle-type models, then we should expect stocks which drop down to the Russell 2000 to experience positive returns. That is exactly what we find in the middle panel of Figure 6. Stocks that drop out of the Russell 1000 and into the 2000 have 0.8% abnormal returns relative to the market on reconstitution day.

However, the same logic says that we should expect stocks which move up to the Russell 1000 to experience negative returns. And that is not what we find in the right panel of Figure 6. Stocks that move up to the Russell 1000 from the 2000 have 0.2% abnormal returns on reconstitution day.

4.3 Liquidity Measures

The above evidence is hard to digest using standard asset-pricing logic given the staggering amount of volume transacted on Russell's reconstitution day each year. "In five of the last six years, reconstitution day ranked in the 10 busiest trading sessions."[6] And things turn positively dyspeptic when you realize that traditional liquidity measures all suggest that, if anything, index switchers are less liquid on reconstitution day.

Figure 7 plots two different liquidity measures for stocks that switch between the Russell 1000 and 2000 in the days surrounding Russell's reconstitution date each year. The data come from the WRDS Intraday Indicators suite, which itself is generated using millisecond TAQ data. This data covers the time period from 2004 to 2020.

As before, we define event time, $\tau = t - t_{reconst}$, as the number of days since the Russell reconstitution date in a given year. $\tau = 0$ corresponds to Russell's reconstitution

[6] Wang, Lu. "Stock volume surges as Russell change spurs index trader frenzy." *Bloomberg.* 06/22/2018.

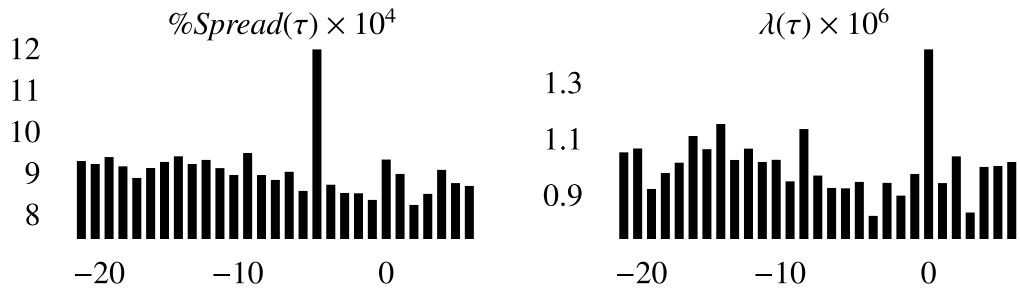

Figure 7. Each panel shows a different measure of liquidity for stocks that switch between the Russell 1000 and 2000 in the days around Russell's reconstitution date. x-axis counts trading days relative to Russell's reconstitution day each year. $\%Spread(\tau) \times 10^4$ reports the median bid-ask spread of index switchers as a percent of the mid-point price on each day. $\lambda(\tau) \times 10^6$ reports the median value of Kyle's λ for index switchers on each day.

day; $\tau < 0$ for the days leading up to reconstitution day; $\tau > 0$ for the days after reconstitution day. Each year, we include data on index switchers starting the day after FTSE Russell's preliminary announcement, $t_{ancmt} + 1$, and ending 7 trading days after Russell's reconstitution date, $t \in \{t_{ancmt} + 1, \ldots, t_{reconst}, \ldots, t_{reconst} + 7\}$.

The left panel in Figure 7 reports the median bid-ask spread of index switchers as a percent of the mid-point price on each day τ, $\%Spread(\tau)$.[7] This panel reveals that index switchers have slightly higher spreads on Russell's reconstitution day as indicated by the red bar. Moreover, this is true even though we already know that the volume of these stocks is going through the roof.

The right panel in Figure 7 then reports the median value of Kyle's λ for index switchers on each day τ. This is the measure of price impact implied by Kyle (1985) described in Section 3.2.[8] On a typical day, buying an extra $1m$ shares of an index switcher will increase its price by roughly \$1.00 per share. But, on reconstitution day, the price impact of buying an extra $1m$ shares is \$1.42 per share.

[7]For each trade k on a given trading day, WRDS first computes

$$2 \cdot \left(\frac{Price(t_k) - MidPoint(t_k)}{MidPoint(t_k)} \right) \times \{ IsBuy(k) - IsSell(k) \}$$

$Price(t_k)$ is the price at which the kth trade is executed. $MidPoint(t_k)$ is an equal-weighted average of the national-best (NB) bid and ask at the time the kth trade took place. $IsBuy(k)$ and $IsSell(k)$ are indicators for whether the kth trade is a buy or sell order. WRDS then averages the intraday results for each stock.

[8]We take the values computed by WRDS using an intercept based on data during market hours.

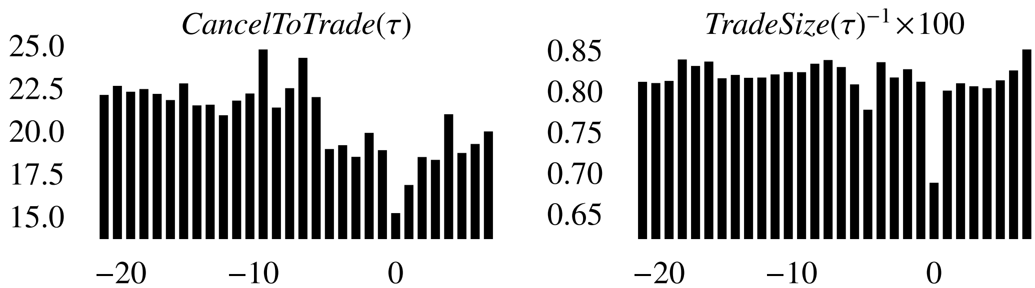

Figure 8. Each panel shows a different proxy for the prevalence of algorithmic trading among stocks that switch between the Russell 1000 and 2000 in the days around Russell's reconstitution date. x-axis counts trading days relative to Russell's reconstitution day each year. *CancelToTrade*(τ) reports the average ratio of canceled orders to completed trades for index switchers on a given day relative to Russell's reconstitution date. *TradeSize*$(\tau)^{-1} \times 100$ reports the average of the ratio of trades to volume for index switchers on a given day.

This pattern of high volume and low liquidity spells trouble for any explanation of excess reconstitution-day volume based on search frictions. It is simply not the case that other traders choose to trade on Russell's reconstitution day because the additional volume from ETF rebalancing makes it is cheaper to trade on this day. Standard liquidity measures paint the opposite picture.

4.4 Algorithmic Trading

If ETFs are constrained to do their rebalancing on Russell's reconstitution day, then one plausible explanation for all of the additional trading that we observe is that algorithmic traders are front running ETF demand. In principle, this explanation could also account for why liquidity dries up on reconstitution days as well.

Unfortunately, the data are not consistent with this explanation. Algorithmic traders tend to cancel lots of orders and make small trades (Weller, 2018). Figure 8 shows that both proxies for algorithmic trading drop on Russell's reconstitution date.

We create this figure using MIDAS data from 2012 through 2020. To calculate *CancelToTrade*(τ), we first divide the number of full or partial order cancellations for an index switcher on day τ by the number of completed trades. Then we average across index switchers on a given date relative to Russell's reconstitution day. To calculate *TradeSize*$(\tau)^{-1}$, we first divide the number of trades for an index switcher on day τ by the number of shares traded. Then we again average across index switchers.

4.5 Mutual Funds

We compute $excessVolume_n$ in Equation (5) by taking an index switcher's total volume and subtracting off both its average daily volume during the previous year and the volume coming from ETF rebalancing. We do not subtract off the volume coming from Russell 1000 and 2000 mutual funds in our main definition. So you might be concerned that trading by mutual funds might be driving out results.

There are two reasons why this is unlikely to be the case. First, the amount of money invested in Russell-benchmarked mutual funds is tiny relative to the amount of money invested in Russell-benchmarked ETFs. At the end of 2020, Russell-benchmarked mutual funds had total assets under management (AUM) of \$14.4*b*; whereas, Russell-benchmarked ETFs had a combined AUM of \$235.9*b*.

Second, unlike an ETF which is benchmarked to the Russell 1000 or 2000, an open-ended mutual fund is not constrained by the threat of creations and redemptions to do all its rebalancing on reconstitution day. Open-ended mutual funds are different from ETFs precisely because they can deviate from their stated benchmark in the short run. If mutual fund trading played an important role, then we would expect to see an increase in their volume during the weeks between FTSE Russell's preliminary announcement and the official reconstitution date.

4.6 Closet Indexing

When computing $ETFvolume_n$, we define the set of Russell-benchmarked ETFs in Equation (3) as all ETFs which are either a) explicitly labeled by ETF Global as benchmarked to the Russell 1000 or 2000 or b) have "Russell 1000" or "Russell 2000" in their name. Perhaps this is too narrow a definition? Perhaps the excess reconstitution-day volume that we document is really coming from ETFs which are closet indexing?

We think this is unlikely for two reasons. First, ETFs are popular because they offer investors a way to hold a particular index together with a guarantee that the per share value of the fund will match the value of the index at the end of each day. This end-of-day no-tracking-error guarantee is what separates ETFs from open-ended mutual funds. If an ETF chose to track a different benchmark than the one it advertised to investors, then these investors could make arbitrage profits via creations and redemptions at the ETF manager's expense.

Thus, the real concern is not closet indexing. The real concern is that we (the researchers; Alex and Marco) are incorrectly classifying Russell-benchmarked ETFs as ETFs which are benchmarked to some other index. After going through the data by hand, we think this is unlikely. But there is always a chance we have missed something.

That being said, even in the worst case, this sort of error cannot account for all the excess reconstitution-day volume that we observe. Consider the thought experiment from Section 2.3. Suppose we assume that the rule we use for deciding whether a fund is benchmarked to either the Russell 1000 or 2000 is as bad as it could possibly be. Suppose we assume that every fund that we said was not benchmarked to the Russell 1000 or 2000 is actually benchmarked to one of these two funds. To get an error this bad, every fund would have to be a Russell-benchmarked ETF.

So, in this extreme case, we could use the Thompson S12 data and calculate the change in every fund's holdings from March to September each year starting in 1991 for every index switcher. This is a 6-month window. And it involves summing over all funds, $f \in \mathsf{AllFunds}$, not just ETFs. Even still, when we calculate "$ETFvolume_n$" this way, it turns out that 179 index switchers still have excess reconstitution-day volume!

4.7 Leveraged ETFs

Finally, our main empirical analysis does not include data on leveraged ETFs when computing $ETFvolume_n$. However, this omission cannot account for our results. Leveraged ETFs manage a tiny amount of money relative to the total size of all Russell-benchmarked ETFs. It is true that, on some otherwise quiet days during the year, trading on behalf of leveraged ETFs can lead to spikes in certain stock's volume at the end of the trading day (Ivanov and Lenkey, 2018).

But we are not studying an otherwise quiet day in financial markets. "'Let's face it, for the New York Stock Exchange, Russell reconstitution, from a trading standpoint, is the greatest show on earth, that's where it all comes down,' said Gordon Charlop, a managing director at Rosenblatt Securities in New York."[9] Our estimates of $ETFvolume_n$ are unchanged to the first decimal place when we redo our analysis including leveraged ETFs.

[9]Mikolajczak, Chuck. "Investors brace for annual Russell index rebalancing with pandemic imprint." *Reuters.* 06/18/2021.

5 What Investors Say

There is lots of money invested in Russell-benchmarked ETFs. So it makes sense that, when a stock switches between the Russell 1000 and 2000, it will realize a spike in volume on Russell's reconstitution date due to ETF rebalancing. Our main finding is that these index switchers actually experience a spike in volume on Russell's reconstitution date that is much bigger than can be explained by ETF rebalancing. For every share traded by ETFs, an extra 3.1 shares get traded by somebody else.

Even though the scale of this excess reconstitution-day volume is economically massive—it is on par with the annual growth of Russell-benchmarked ETFs—this phenomenon is a complete mystery from the viewpoint of existing information-based asset-pricing models. For example, in spite of its size, excess reconstitution-day volume is only associated with small reconstitution-day price movements. This is true even though standard measures of liquidity appear to dry up on reconstitution day.

However, "Where does excess reconstitution-day volume come from?" is only a mystery in the way that "Why am I sitting in traffic at 2pm on a Thursday afternoon?" is a mystery. You may not have expected to see a wall of brake lights snaking off into the distance on your early commute home. But every driver involved in the traffic jam made a decision to be on the road. So the reason why you are currently sitting in bumper-to-bumper traffic is in principle knowable by talking to the other drivers.

So what do the investors who made a decision to trade on Russell's reconstitution day say about all of the excess volume we observe? This section documents how market participants and journalists talk about Russell's reconstitution day in print. We have also had numerous conversations with industry participants on background. We have tried to choose quotes that are broadly consistent with all that we have learned.[10]

There is no guarantee that investors' reasoning will be economically sound or even properly justify their actions. Nevertheless, knowing how investors are reasoning about a situation is invaluable for guiding model construction (Chinco, Hartzmark, and Sussman, 2021). And that is exactly why we are interested in excess reconstitution-day volume. The scale of this phenomenon suggests that our standard models are missing something important about the economic implications of indexing.

[10] If you are an investor with information about where excess reconstitution-day volume is coming from, please let us know. Our contact details are on the title page. We would love to hear your thoughts.

Investors are aware that the ETF industry has exploded in recent years. "Everyone knows that exchange-traded funds have soared in size over the past decade, with the biggest now valued at $\$258b$. But where ETFs have really excelled is in grabbing trading, with far more trading now in several of the big ETFs than in the largest members of the indexes to which they are linked."[11]

Investors are aware that stocks which switch between popular indexes, such as the Russell 1000 and 2000, will realize a massive spike in volume on reconstitution day. A June 2015 article titled "Why Friday could be the year's biggest trading day." starts with "Stock-trading volume is expected to surge in Friday's last moments of trading, as the Russell indexes go through their annual routine of adding and removing stocks."[12]

Investors are aware that ETF rebalancing is different from other kinds of trading. "There are nine dates each year when index funds do most of their trading. Data suggests that index funds are around 40% of all trading on index rebalance dates."[13] And most of this action occurs at market close. "KBW analyst Melissa Roberts expects the bulk of passive fund trading related to the reconstitution will occur in the last 15 minutes or so of the session."[14]

FTSE Russell is even starting to worry that its decision to rebalance the Russell 1000 and 2000 indexes once a year is distorting the market and hurting investors. "FTSE Russell has launched an internal review into the annual rebalancing frequency of its family of US stock indices followed by about $\$9t$ of investors' money, on concerns that the rejigs cause unhealthy concentrations of trading."[15]

But, as far as we can tell, investors are not generally aware that much of the spike in volume that index switchers experience on Russell's reconstitution day is not coming directly from ETF rebalancing. We cannot find any media outlets talking about the gap between expected ETF rebalancing activity and realized volume on Russell's reconstitution day. And the investors that have talked to thus far were not previously aware of the unexplained gap between the two numbers.

---

[11] Mackintosh, James. "How ETFs Swallowed the Stock Market." *The Wall Street Journal.* 08/19/2019.
[12] Reklaitis, Victor. "Why Friday could be the year's biggest trading day." *MarketWatch.* 06/26/2015.
[13] Mackintosh, Phil. "Who is trading on US markets?" *Nasdaq News + Insights.* 01/28/2021.
[14] Mikolajczak, Chuck. "Investors brace for annual Russell index rebalancing with pandemic imprint." *Reuters.* 06/18/2021.
[15] Johnson, Steve. "FTSE Russell to review $\$9t$ US index shake-up." *Financial Times.* 07/19/2021.

We may not have talked to the right investors yet, but we have talked to many investors. So we can conclude, at the very least, that the existence of excess reconstitution-day volume is not widely appreciated by market participants. And this is useful information for any theorists trying to model this phenomenon. An investor cannot be strategically reacting to a phenomenon he is wholly unaware of. Given the scale of excess reconstitution-day volume, the decisions of many investors must turn out to be related to this phenomenon. But it is not obvious that investors are pricing in the nature of this relationship. Indeed, this would be consistent with the small price movements that we document on Russell's reconstitution day in Figure 6.

6 Conclusion

This paper documents that stocks which switch between the Russell 1000 and 2000 experience economically massive amounts of unexplained trading volume on Russell's reconstitution day. When a stock switches between these two indexes, any ETF benchmarked to either index must rebalance. And this ETF rebalancing activity typically increases an index switcher's reconstitution-day volume by 139% relative to its daily average over the previous year. However, for each share of an index switcher that gets traded on behalf of ETFs, an extra 3.1 shares get traded by someone else on Russell's reconstitution day.

The dollar value of this excess volume on one trading day is on the same order of magnitude as the annual growth in AUM for all Russell-benchmarked ETFs each year. The sheer scale of this phenomenon should change researchers' views about what needs to be modeled about the rise of index investing. And, what is more exciting is that modeling excess reconstitution-day volume will require a new kind of model. The standard information-based asset-pricing frameworks of Grossman and Stiglitz (1980)/Admati (1985) and Kyle (1985) cannot speak to our main empirical finding. So, regardless of what the explanation turns out to be, researchers stand to learn something important about how financial markets work by trying to better understand excess reconstitution-day volume.

.

References

Admati, A. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica*.

Appel, I., T. Gormley, and D. Keim (2020). Identification using Russell 1000/2000 index assignments: A discussion of methodologies. *Critical Finance Review*.

Baker, M., B. Bradley, and J. Wurgler (2011). Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*.

Barberis, N., A. Shleifer, and J. Wurgler (2005). Comovement. *Journal of Financial Economics*.

Ben-David, I., F. Franzoni, and R. Moussawi (2018). Do ETFs increase volatility? *Journal of Finance*.

Beneish, M. and J. Gardner (1995). Information costs and liquidity effects from changes in the Dow Jones Industrial Average list. *Journal of Financial and Quantitative Analysis*.

Beneish, M. and R. Whaley (1996). An anatomy of the “S&P Game”: The effects of changing the rules. *Journal of Finance*.

Bond, P. and D. García (2021). The equilibrium consequences of indexing. *Review of Financial Studies*.

Brogaard, J., M. Ringgenberg, and D. Sovich (2019). The economic impact of index investing. *Review of Financial Studies*.

Burnham, T., H. Gakidis, and J. Wurgler (2018). Investing in the presence of massive flows: The case of MSCI country reclassifications. *Financial Analysts Journal*.

Buss, A. and S. Sundaresan (2021). More risk, more information: How passive ownership can improve informational efficiency. *Working Paper*.

Chabakauri, G. and O. Rytchkov (2021). Asset pricing with index investing. *Journal of Financial Economics*.

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

Chinco, A. and V. Fos (2021). The sound of many funds rebalancing. *Review of Asset Pricing Studies*.

Chinco, A., S. Hartzmark, and A. Sussman (2021). A new test of risk factor relevance. *Journal of Finance*.

Da, Z. and S. Shive (2018). Exchange traded funds and asset return correlations. *European Financial Management*.

Easley, D., D. Michayluk, M. O’Hara, and P. Tālis (2021). The active world of passive investing. *Review of Finance*.

Garleanu, N. and L. Pedersen (2021). Active and passive investing: Understanding Samuelson's dictum. *Working Paper*.

Glosten, L., S. Nallareddy, and Y. Zou (2021). ETF activity and informational efficiency of underlying securities. *Management Science*.

Greenwood, R. (2008). Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. *Review of Financial Studies*.

Grossman, S. and J. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economics Review*.

Haddad, V., P. Huebner, and E. Loualiche (2021). How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. *Working Paper*.

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. *Journal of Finance*.

Ivanov, I. and S. Lenkey (2018). Do leveraged ETFs really amplify late-day returns and volatility? *Journal of Financial Markets*.

Kashyap, A., N. Kovrijnykh, J. Li, and A. Pavlova (2021). The benchmark inclusion subsidy. *Journal of Financial Economics*.

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

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*.

Lee, J. (2021). Passive investing and price efficiency. *Working paper*.

Lettau, M. and A. Madhavan (2018). Exchange-traded funds 101 for economists. *Journal of Economic Perspectives*.

Madhavan, A. (2003). The Russell reconstitution effect. *Financial Analysts Journal*.

Madhavan, A. (2016). *Exchange-traded funds and the new dynamics of investing*. Oxford University Press.

Novick, B., S. Cohen, A. Madhavan, T. Bunzel, J. Sethi, and S. Matthews (2017). Index investing supports vibrant capital markets. *BlackRock Whitepaper*.

Sammon, M. (2021). Passive ownership and price informativeness. *Working Paper*.

Stambaugh, R. (2014). Presidential address: Investment noise and trends. *Journal of Finance*.

Weller, B. (2018). Does algorithmic trading reduce information acquisition? *Review of Financial Studies*.

Wurgler, J. (2011). On the economic consequences of index-linked investing. In *Challenges to Business in the Twenty-First Century*.

Wurgler, J. and E. Zhuravskaya (2002). Does arbitrage flatten demand curves for stocks? *Journal of Business*.