Excess Reconstitution-Day Volume*

Alex Chinco[†] and Marco Sammon[‡]

February 9, 2022

Abstract

FTSE/Russell reconstitutes its Russell 1000 and 2000 indexes on the last Friday of June each year, and it announces which stocks will switch indexes roughly two weeks in advance. Exchange-traded funds (ETFs) must do all of their rebalancing on Russell's reconstitution day, but other kinds of funds (mutual funds, pension funds, endowments, etc) could use this two-week period to gradually rebalance. Textbook models suggest that some non-ETFs should take advantage of this opportunity. Yet index-switcher volume does not increase at all during the period prior to Russell's reconstitution day. It suddenly spikes on reconstitution day with $3.15 \times$ more volume than can be explained by ETF rebalancing alone. The existence of so much excess reconstitution-day volume suggests that there are important economic forces missing from existing models of index investing. It is as if all Russell-linked funds are trading like ETFs even though most of these funds do not face the same end-of-day portfolio constraints that ETFs do.

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

*We would like to thank Malcom Baker, Hank Bessembinder, David Brown, Rodney Comegys, Zhi Da, Slava Fos, Diego Garcia, Robin Greenwood, Valentin Haddad, Jiacui Li, Sida Li, Anna Pavlova, Matthew Pritsker, Nick Roussanov, Andrei Shleifer, Laura Veldkamp, Luis Viceira, and Jeff Wurgler as well as seminar participants at CUNY Graduate College 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 announced that Lumentum Holdings would move from the Russell 2000 to the 1000. Friday June 26th would be its last trading day as part of the Russell 2000, so a sizable chunk of Lumentum's 75*m* shares outstanding needed to change hands by then. Exchange-traded funds (ETFs) benchmarked to the Russell 2000 needed to sell roughly 2.2*m* shares of Lumentum, and ETFs benchmarked to the Russell 1000 needed to acquire 300*k* shares. Mutual funds, pension funds, endowments, etc that tracked the Russell 1000 or 2000 also had to rebalance.

ETFs benchmarked to the Russell 1000 or 2000 had to do all their rebalancing on Russell's reconstitution day. So you would expect Lumentum's volume on June 26th to be 2.2m + 300k = 2.5m shares higher than usual. Mutual funds, pension funds, endowments, etc linked to the Russell 1000 or 2000 should do some of their trading on June 26th as well. But these non-ETFs did not have to do all their rebalancing on this one day. And textbook models suggest that, given the opportunity, some of these funds should have gradually rebalanced their positions in the days leading up to June 26th.

But this is not what we observe. As shown in Figure 1, there was no gradual increase in the company's volume prior to Russell's reconstitution. Instead, Lumentum's volume spiked to 12.6*m* shares on Friday June 26th. This is $8.6m/2.5m \approx 3.4$ times more than required by ETF rebalancing alone:

$$Excess Volume = 8.6m = 12.6m - 1.5m - 2.5m,$$
(1)
$$\begin{pmatrix} Total \\ volume \end{pmatrix} \begin{pmatrix} Avg \ volume \\ last \ year \end{pmatrix} \begin{pmatrix} ETF \\ volume \end{pmatrix}$$

For every share traded by ETFs, an extra 3.4 shares got 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. Stock's average daily volume during the previous year. The past year 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. By contrast, these index switchers did not experience any run-up in trading volume prior to reconstitution day as textbook models would suggest.

The bar outlined in green shows that, relative to average daily volume during the previous year, trading on behalf of Russell-benchmarked ETFs boosts the typical index switcher's volume by 139% on reconstitution day. However, on top of this predictable increase in volume coming from ETF rebalancing, other traders collectively exchange $4.39 \times$ an index switcher's average daily volume on reconstitution day. For every share traded by ETFs on reconstitution day, $4.39/1.39 \approx 3.15$ 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

Many other kinds of funds are benchmarked to the Russell 1000/2000. However, if rebalancing by these non-ETFs is responsible for excess reconstitution-day volume, then

¹Rolf 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.

it would only deepen the puzzle. Mutual funds, pension funds, endowments, etc do not face the same portfolio constraints as ETFs. They do not have to do all of their rebalancing on Russell's reconstitution day. At least some of them should be able to reduce their trading costs by rebalancing prior to reconstitution. Yet index-switcher volume does not gradually ramp up as reconstitution day approaches. It suddenly spikes on reconstitution day with most of this volume occurring at market close.

For non-ETFs to explain excess reconstitution-day volume, an astonishingly large number of mutual funds, pension funds, endowments, etc would have to be mechanically matching the Russell 1000/2000 weights at market close each day. In June 2020, 11.5% of the total market value of all Russell 3000 stocks (\$3.7t of all \$32.5t) would need to be held by strict end-of-day Russell indexers to account for excess reconstitution-day volume. This is $21\times$ the size of all Russell-benchmarked ETFs. Many funds track the Russell 1000 or 2000 in some way. But is it really the case that this many funds are mechanically pegged to one of these indexes on a daily basis?

By construction, the $4.39 \times$ increase in reconstitution-day volume that we observe is not coming 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. And this changes the narrative about the rise of index investing.

Right now, when a researcher thinks about index investing, the first thing he usually thinks about is this industry's explosive growth. This was the topic of Robert Stambaugh's 2014 AFA address (Stambaugh, 2014). 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, 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 of the excess reconstitution-day volume for just these 66 stocks is \$14.5*b*. Thus, the excess reconstitution-day volume experienced by 66 stocks on one trading day is on the same order of magnitude as the annual growth of all Russell-benchmarked ETFs.

Neither of the two main information-based asset-pricing frameworks can account for both the existence of so much excess reconstitution-day volume and the absence of



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.

any additional volume prior to Russell's reconstitution day. Models in the spirit of Grossman and Stiglitz (1980) and Admati (1985) 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 reconstitution day. FTSE/Russell publicly announces which stocks will switch indexes; all traders learn the same information from a press release not prices. The existence of excess volume on Russell's reconstitution day also implies that the population of traders is not fixed. It is different on one day a year.

Conversations with industry practitioners suggest that, while everyone expects index-switcher volume to spike on Russell's reconstitution day, most think this spike is due to ETF rebalancing. In Grossman and Stiglitz/Admati-type models, investors draw inferences from noisy prices by filtering out the effects of supply shocks. Investors clearly draw inferences from prices. But, if they are not widely aware of excess reconstitution-day volume, then it is a stretch to model their inference problem as a process of filtering out the effects of supply shocks on prices.

Models in the spirit of Kyle (1985) study an investor who strategically trades on his long-lived private information. Unfortunately, this setup also does not describe what happens on Russell reconstitution day. FTSE/Russell publicly announces which stocks will switch indexes. Everyone learns this information at the same time. And even if this were not the case, Kyle-type models predict that an investor with inside information would strategically smooth out his demand to limit price impact. And the traders responsible for excess reconstitution-day volume 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

All of this suggests that existing models for studying the rise of index investing are missing something important. The dollar value of excess reconstitution-day volume is on par with the annual growth in index investments. So a model that only explains the rise of index investing and does not account for the existence of reconstitution day volume would be *missing half the story!* Without a more complete understanding of who else trades on Russell's reconstitution day and why, researchers and regulators should hesitate before drawing strong policy conclusions from these frameworks.

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

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. Regardless of which you prefer, it is massive.

There are two broad classes of information-based asset-pricing models that researchers typically use to analyze indexing. In Section 3, we describe how it is hard for both kinds of models to explain excess reconstitution-day volume.

In Section 4 we discuss tracking error, price changes, intermediated trades, liquidity, algorithmic trading, rebalancing by non-ETFs, closet indexing, classification errors, leveraged ETFs, derivatives usage, and heartbeat trades.

Finally, in Section 5, we look at what investors say about excess reconstitution-day volume. As pointed out in Chinco, Hartzmark, and Sussman (2021), it helps to know what investors are aware of when writing down a model about what they are trying to do.

Related Literature

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 (Wurgler, 2011).

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 of index additions/deletions on prices. 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 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.

Pavlova and Sikorskaya (2021) give evidence that benchmarking leads to relatively inelastic demand for index switchers. We find that index switchers have more reconstitution-day volume than implied by perfectly inelastic benchmarking demand, which is related to Gabaix and Koijen (2021)'s inelastic market hypothesis.

ETFs are now the main way that investors index (Madhavan, 2016; Lettau and Madhavan, 2018). 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 quite a lot (Easley, Michayluk, O'Hara, and Tālis, 2021) and do most of their trading when their benchmark index reconstitutes (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 of ETF rebalancing on one of these days, 3.15 additional shares get traded by other investors.

Finally, this paper connects to the broader literature on why investors trade so much (Milgrom and Stokey, 1982; Grinblatt and Keloharju, 2001). Common explanations are based on differences of opinion (Harris and Raviv, 1993; Wang, 1994; Kandel and Pearson, 1995; Bessembinder, Chan, and Seguin, 1996; Foster and Viswanathan, 1996; Hong and Stein, 2007; Banerjee and Kremer, 2010) and/or overconfidence (Odean, 1998; Barber and Odean, 2001; Scheinkman and Xiong, 2003; Statman, Thorley, and Vorkink, 2006). We document an important setting with lots of unexplained trading activity even though there is no new information to disagree over or be overconfident about.

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 3.15 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

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. FTSE Russell does not release the market cap data it uses to rank stocks in May of each year. But this does not make some traders substantially better informed. We follow the standard procedure used in the literature (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.

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

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



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.

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.

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, 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 *n*th stock's volume on day *t*. Likewise, let $price_n(t)$, $return_n(t)$, and $\#shares_n(t)$ denote the *n*th 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

	All		1000		2000		
	Switchers		to 2000		to 1000		
Announcement Date	Avg	Sd	Avg	Sd	Avg	Sd	
Market Cap (\$ <i>b</i>)	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 (<i>bp</i>)	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 (\$ <i>m</i>)	42.0	36.6	38.4	20.2	45.7	47.8	
Average Weight in ETF Portfolio (<i>bp</i>)	6.2	4.1	5.0	2.4	7.4	5.0	
#Obs	634		323		3	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, $t_{reconst}$.

by the average daily trading volume of each stock during the 252 trading days leading up to FTSE Russell's preliminary announcement, t_{ancmt} :

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

We only include stocks with 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 day 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.

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:

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

Let $holdings_{n,f}(t)$ denote the number of shares that fund f holds of the *n*th stock on day t. Ideally, we could compute ETF volume for the *n*th stock on the Russell reconstitution day as $\sum_{f \in \mathsf{ETFS}_{Russell}} |holdings_{n,f}(t_{reconst} + 1) - holdings_{n,f}(t_{reconst} - 1)|$. 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

$$ETFvolume_n(t) = \begin{cases} \sum_{f \in \mathsf{ETFs}_{Russell}} \left| \Delta_{\pm 5}[holdings_{n,f}(t)] \right| & \text{if } t = t_{reconst} \\ 0 & \text{otherwise} \end{cases}$$
(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. We also ignore changes in ETF holdings arising from creations/redemptions.

Vanguard ETFs only report holdings at the end of each month. These funds are actually share classes of mutual funds, giving them different disclosure requirements. We compute the contribution of Vanguard ETFs to $ETFvolume_n(t)$ as the difference between these funds' holding at the end of May and their holdings 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 correlation between the two approaches to calculating $ETFvolume_n(t)$ is north of 0.99.

The bottom two panels of Table 1 describe ETF holdings of index switchers immediately before and after reconstitution. The right-most columns 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.

There are at least two reasons why our approach to calculating $ETFvolume_n(t)$ likely understates the amount of unexplained trading activity on Russell's reconstitution date, $t_{reconst}$. First, we assume that all ETF trading in a multi-day window around Russell's reconstitution date occurs on the $t_{reconst}$ itself. 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. If we had access to more granular data that pinned down the exact timing of ETF trades, we would get lower values of $ETFvolume_n(t)$, which would imply even more unexplained volume on Russell reconstitution days. Second, our calculations assume that every Russell-benchmarked ETF will buy every stock added to the Russell 1000/2000 and sell every stock dropped from that index. However, ETFs engage in optimized sampling. When an ETF does not rebalance its position in an index addition or deletion, its tracking error relative to the Russell 1000/2000 will increase; however, an ETF incurs no trading costs on the trades it does not execute. So, if a stock enters into or drops out of the Russell 1000/2000 with a sufficiently small weight, some Russell-benchmarked ETFs may prefer to do nothing. As a result, some of the index-switcher volume we attribute to ETF rebalancing on Russell's reconstitution day, ETF volume_n(t), is likely coming from elsewhere.

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. 3.15 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:

$$excessVolume_{n}(t) = volume_{n}(t)$$

$$- avgVolumePastYear_{n}$$

$$- ETFvolume_{n}(t)$$
(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:

$$excess \$Volume_n(t) = excess Volume_n(t) \times price_n(t_{ancmt} - 1)$$
(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. $relVolume_n(t)$ is the difference between a stock's volume and its average daily volume last year divided by its average daily volume last year:

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

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

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

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

$$= \frac{volume_n(t) - ETFvolume_n(t)}{avgVolumePastYear_n} - 1$$
(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 index switchers from 2001 to 2020—i.e., one observation per index switcher each year. 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

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

15

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 without any standardization (units: millions of shares). $10^{-8} \times excessVolume_n$: 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.

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 $7.43 \times$ 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 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, $100 \times excessVolume_n(t_{reconst})/#shares_n$. The results are largely unchanged.

In the fourth row, we show results for excess reconstitution-day volume without doing any standardization, $excessVolume_n(t_{reconst})$. 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, $excess$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 sudden spike in volume on Russell's reconstitution day. In the days before reconstitution, the typical index switcher has trading activity that looks no different from the past year. Then, all of the sudden, things change on reconstitution day. And, just as suddenly, they go back to normal.

This brings up an important distinction: the focus of our analysis is not 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. $\tau = 0$ is reconstitution day and $\tau < 0$ for the days leading up to 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, $t \in \{t_{ancmt} + 1, \dots, t_{reconst}, \dots, t_{reconst} + 7\}$. Table 3 shows results of panel regressions of the form below

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

using data on switchers and almost switchers each year. Table 2 only uses data on Russell's reconstitution day; Table 3 also uses data from the surrounding days.

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 $\hat{\gamma} \approx 0$.

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 *isEventDate*(τ) in third row of column (1) is 0.27 + 4.12 = 4.39. This finding formalizes our earlier

	All Switchers	All Switchers	All Switchers	1000 to 2000	2000 to 1000	Almost Switchers
	·01-·20	'01-'06	'07-'20	·01-·20	·01-·20	'07-'20
	(1)	(2)	(3)	(4)	(5)	(6)
intercept	0.27*** (0.03)	0.27*** (0.04)	0.27*** (0.05)	0.10 (0.07)	0.41*** (0.04)	0.07 (0.06)
$isAfter(\tau)$	-0.00 (0.03)	-0.04 (0.03)	$\underset{(0.06)}{0.04}$	$0.08^{\star}_{(0.05)}$	$-0.07^{\star\star}_{(0.03)}$	$\underset{(0.05)}{-0.07}$
$isEventDate(\tau)$	4.12*** (0.36)	3.34*** (0.48)	5.09*** (0.20)	4.00*** (0.42)	4.25*** (0.39)	$\underset{(0.06)}{0.40}^{\star\star\star}$
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

Dependent Variable: $relExcessVolume_n(\tau)$

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 on Russell's reconstitution date. Numbers in parentheses are standard errors clustered by year. *, **, and *** denote estimates that are statistically distinguishable from zero at the 10%, 5%, and 1% levels.

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 $y \in \{1991, ..., 2020\}$. Let N(y) denote the set of stocks that switch between the Russell 1000 and 2000 in year y. The dependent variable is $\overline{relExcessVolume}(y)$, the average excess volume on Russell's reconstitution day in year y:

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

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.

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 Russellbenchmarked 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 Dependent Variable: *relExcessVolume*(y)

	(1)	(2)	(3)
intercept	1.05*** (0.14)	1.05*** (0.15)	1.05^{***}
isAfter2000(y)	3.78*** (0.31)	$2.88^{\star\star\star}_{(0.46)}$	$3.04^{***}_{(0.40)}$
$\{ \# Russell ETFs(y) / 100 \} \times isAfter 2000(y)$		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, ..., 2020. The dependent variable in all regressions, *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.

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 Russellbenchmarked 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 excess reconstitution-day volume is absolutely massive 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: How much trading volume would have to be reclassified as occurring on Russell's reconstitution day to explain our results? We would have to assume that all mutual-fund and ETF trading during the six months surrounding the Russell's reconstitution day took place on the reconstitution day.

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?

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*:

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

We estimate that, from 2007 through 2020, the amount of money invested in Russellbenchmarked ETFs grew by an average of \$11.6*b* 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 \$17.6*b* per year if we look only at more recent data from 2014 to 2020.

⁵Sanford Bernstein. "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*, totalExcessVolume(*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.

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:

$$totalExcess\$Volume(y) = \sum_{n \in \mathsf{N}(y)} excess\$Volume_n(t_{reconst})$$
(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 \$11.4b 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 \$14.5b.

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.

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, so is the other. Any model of the economic consequences of indexing that ignores excess reconstitution-day volume is ignoring half the story!

It is true that ETFs like the iShares Russell 1000 ETF (IWB) are some of the most actively traded assets in modern financial markets. The dollar value of IWB's average daily volume in 2020 was \$123.6*m*. However, this does not mean that we should not be comparing the dollar value of excess reconstitution-day volume to industry-wide growth.

On a typical trading day, investors trade shares of IWB more than you would expect based on the logic of textbook models. We document that, on Russell reconstitution day, investors trade shares of index switchers much more than you would expect based on textbook logic. The first observation does not rationalize or explain away the second. When one investor buys a share of IWB from another, the ETF does not need to rebalance. On Russell reconstitution day, investors in IWB do not need to do anything.

The scale of excess reconstitution-day volume for index switchers suggests that existing models of indexing are missing something important. Pointing to inexplicably large amounts of trading volume in other contexts is not an argument to the contrary. The usual explanations for high trading volume are things like disagreement and overconfidence. We are documenting excess volume in a setting where there is no new information to disagree over or be overconfident about.

Finally, most volume on a typical trading day comes from trades made by highfrequency traders (HFTs). It would not make sense to compare the dollar value of HFT volume to the annual growth of the ETF industry. However, in Section 4.5 below, we specifically show that the excess reconstitution-day volume is not driven by HFTs.

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 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, $f \in ETFs$, rather than just the Russell-linked funds, $f \in ETFs_{Russell}$, every one of these 66 stocks still has excess reconstitution-day volume.

We can take things a step further. Suppose we calculate changes in every fund's holdings from March to September 2020 using Thompson S12. This involves summing over all funds, $f \in AllFunds$, not just ETFs over a 6-month window. When we calculate "*ETFvolume*_n" this way in June 2020, 2 of 66 index switchers still have excess volume.

It is always possible that the explanation for any new and surprising empirical finding is a data error. However, the calculation above indicates that, to explain our results this way, there would have to be an epic amount of data mislabeling. Both our primary data sources would have to be systematically misreporting the amount of index-related trading activity on Russell's reconstitution date by an amount equal to all the trading done by every investment fund over an entire 6-month window. It is far more plausible that there is lots of extra trading activity on Russell's reconstitution day.

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 can account for the existence of so much excess volume on Russell's reconstitution day. At best, these two modeling paradigms are merely incomplete. There is an entire ecosystem of trading activity that takes place on Russell's reconstitution day which is roughly equal in size to the annual growth of index investing. And this ecosystem exists for reasons that are completely outside the scope of the standard Grossman and Stiglitz/Admati- and Kyle-type models.

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

Grossman and Stiglitz (1980) and Admati (1985) study how an asset's price aggregates the private signals of a fixed population of traders in a static setting.

Basic Setup

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

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

$$\max \mathbb{E}^{i} \Big[-e^{-\rho \cdot Wealth^{i}} \Big| Price_{1}, \dots, Price_{N} \Big]$$
(13)

 ρ captures investors' level of risk aversion. The *i*th 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 *i*th investor believes $Payout_n \sim Normal(\mu_n^i, [\sigma_n^i]^2)$. 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 prior beliefs from the beginning.

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

$$Supply_n \sim Normal(\mu_{Supply}, \sigma^2_{Supply})$$
 (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 *i*th investor's demand for the *n*th 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 *n*th risky asset at the equilibrium price, $Price_n$, for markets to clear:

$$Supply_{n} = \int_{i \in \text{Investors}} Demand_{n}^{i}(Price_{n}) \cdot di$$
(15)

Investors' private information and supply shocks both affect market clearing. So when the *i*th investor sees an unexpectedly high price, he can infer that other investors received good news, $\int_{i'\neq i} \mu_n^{i'} \cdot di' > \mu_n^i$, 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 *n*th risky asset, $Price_n$, would be different each time. But the asset's average price across simulations

$$\mathbb{E}[Price_n] \propto \int_{i \in \mathsf{Investors}} \mathbb{V}\mathrm{ar}^i [Payout_n | Price_n]^{-1} \cdot \mathrm{d}i$$
(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 this economically massively phenomenon simply occurs for reasons that are completely unrelated to those captured by Grossman and Stiglitz/Admatitype 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 $T \ge 1$ auctions for a single risky asset. This risky asset has a terminal payout of *Payout* ~ 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:

$$\{ Payout - Price(t) \} \cdot Demand(t) + \mathbb{E}_t [ContinuationValue(t+1)]$$

$$Profit from trading on informational advantage in just the tth auction.$$

$$Expected profits from trading on informational advantage in all future auctions {t+1,...,T}.$$

$$(17)$$

 $\mathbb{E}_t[\cdot]$ 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):

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

The demand noise each period is a random shock, $Noise(t) \stackrel{\text{IID}}{\sim} Normal(0, \sigma^2)$, 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 {t + 1, ..., 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. This critique also applies to Kyle-type models. Excess reconstitution-day volume is not about an investor strategically trading on long-lived private information. FTSE Russell publicly announces which stocks will switch indexes.

The timing of excess reconstitution-day volume also violates the main takeaway from Kyle-type models. At 6pm on June 5th 2020, Russell-benchmarked ETFs knew that Lumentum would be switching from the Russell 2000 to the 1000 at the end of trading on June 26th. Kyle (1985) says that they should have gradually resolved their positions over the next few weeks. Yet, when we look at the data, they did all their trading at market close on June 26th.

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 each modeling paradigm does not match the facts about what happens on Russell's reconstitution day. The nature of excess reconstitution-day volume also poses deeper problems for both Grossman and Stiglitz/Admati- and 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. There are good reasons why an index switcher might experience higher volume on Russell's reconstitution day. We can find no good reason why an index switcher should experience so much excess volume. None of these other considerations can predict anything close to 3.15 extra shares traded for each share of ETF rebalancing.

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. As we document in Subsection 4.4 below, 75% of all trades on Russell's reconstitution day in June 2020 got executed as part of the closing auction or via market-on-close orders.

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. This is why we entitled our paper "Excess Reconstitution-Day Volume" rather than "ETF Reconstitution-Day Volume".

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. It is reassuring that we can closely replicate the original estimate. But it is also surprising that our estimate is so similar when using updated data where indexing is more prevalent.



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.

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 $t_{reconst}$.

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

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 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, $t \in \{t_{ancmt} + 1, \dots, t_{reconst}, \dots, t_{reconst} + 7\}$.

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... even though 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 1*m* 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 1*m* shares is \$1.42 per share.

$$2 \cdot \left(\frac{Price_n(t_k) - MidPoint_n(t_k)}{MidPoint_n(t_k)}\right) \times \left\{IsBuy_n(k) - IsSell_n(k)\right\}$$

⁶Lu Wang. "Stock volume surges as Russell change spurs index trader frenzy." *Bloomberg.* 06/22/2018. ⁷For each trade k on day τ for stock n, WRDS first computes

*Price*_n(t_k) is the price at which the *k*th trade is executed for stock *n*. *MidPoint*_n(t_k) is an equal-weighted average of the national-best (NB) bid and ask for stock *n* at the time the *k*th trade took place. *IsBuy*_n(k) and *IsSell*_n(k) are indicators for whether the *k*th trade for stock *n* was a buy or a sell order. *%Spread*_n(τ), is the dollar-weighted average of this measure across trades for stock *n* on day τ .

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



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.

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. Other markets may work differently. For instance, Bessembinder, Carrion, Tuttle, and Venkataraman (2016) documents that liquidity improves in the crude-oil futures market on days when a large crude-oil ETF (the United States Oil Fund; USO) places predictable trades. However, the same is not true for equity markets on Russell reconstitution day.

4.4 Intermediated Trading

In Section 3, we discuss how the two standard information-based asset-pricing frameworks cannot account for excess reconstitution-day volume. However, both these frameworks focus on buy-and-hold returns. It is unlikely that a long-term investor took the other side of every ETF rebalancing trade. If many of the trades involved intermediaries, then the additional trading could substantially increase volume.

In the Brunnermeier and Pedersen (2005) model, the extra intermediated trading is "predatory" in nature, and market quality suffers as a result. In Bessembinder, Carrion, Tuttle, and Venkataraman (2016), the intermediated trading is more benign and gets the label "strategic trading". In the Admati and Pfleiderer (1991) "sunshine trading" model, an investor who plans to place a large order will announce this fact to the market to



Figure 8. Fraction of all trades executed at the end of trading on a given trading day for index switchers in June 2020 using TAQ data. Pr[*executed in final minute*]: fraction of all trades each day for an index switcher that were executed after 3:59pm. Pr[*executed at market close*]: fraction of all trades each day for an index switcher that were executed either at the closing auction or via market-on-close orders. The middle bar in each panel depicts results using data on Russell's reconstitution day, $t_{reconst}$ = June 26th, 2020. Left bars depict results using data from the previous 20 trading days, and right bars depict results using data from the 7 trading days after reconstitution.

attract additional liquidity in the form of short-term intermediated trades. And Lyons (1997) posits that "hot potato" inventory trading is one of the main reasons for the large trading volumes observed in foreign-exchange markets.

Intermediated trading clearly plays an important role in asset markets. However, for this story to explain excess reconstitution-day volume, we should be able to find evidence of lots of short-term trading by intermediaries. We have already seen in Figure 2 that there is no evidence of increased trading volume for index switchers in the days prior to Russell's reconstitution. So, if intermediated trading serves to boost trading volume around Russell's annual reconstitution event, then the effect has to operate intraday.

Yet, when we examine TAQ data on the 44 stocks which are primarily listed on the NYSE that switched between the Russell 1000 and 2000 in June 2020, the vast majority of their trading activity occurred at market close on Russell's reconstitution day. The left panel in Figure 8 shows that 80% of all shares traded were placed during the final minute of trading on Russell's reconstitution day. The right panel reveals that 75% of all reconstitution-day trades were executed through the closing auction or via market-on-close orders. This is consistent with Li (2022) who documents that ETFs do the vast majority of their rebalancing right at market close, which leaves no room for any "hot potato" effect.



Figure 9. 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*(τ)⁻¹ × 100 reports the average of the ratio of trades to volume for index switchers on a given day.

4.5 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 trading is one particular kind of intermediated trading, and theories based on intermediated trading have a hard time explaining excess reconstitution-day volume. Most of the excess volume gets executed all at once at the end of trading on Russell's reconstitution day.

We proxy for algorithmic trading activity by leveraging the observation that algorithmic traders tend to cancel lots of orders and make small trades (Weller, 2018). Figure 9 shows that both proxies for algorithmic trading drop on Russell's reconstitution date. This finding is exactly what you would expect given the results shown in Figure 8.

We create this figure using MIDAS data from 2012 through 2020. To calculate $CancelToTrade(\tau)$, 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.

4.6 Other Kinds of Funds

We compute *excessVolume*_n(*t*) 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 expected volume coming from ETF rebalancing. However, many other kinds of funds (mutual funds⁹, pension funds, sovereign wealth funds, endowments, etc...) are either explicitly benchmarked to the Russell 1000/2000 or have their performance evaluated against one of these indexes. And we do not subtract off the volume coming from these other kinds of funds when computing *excessVolume*_n(*t*).

However, these other kinds of funds are unlikely to be the source of all the unexplained volume we observe on Russell reconstitution day. Unlike a Russellbenchmarked ETF, non-ETFs like mutual funds, pension funds, sovereign wealth funds, and endowments are not constrained to do all their rebalancing on reconstitution day. The net asset value of their holdings can deviate from the value of the Russell 1000 or 2000 in the short run.

Any of these non-ETFs could reduce their trading costs by doing some of their rebalancing prior to Russell's reconstitution day. For example, if mutual-fund trading played an important role, we would expect to see an increase in volume during the weeks between FTSE Russell's preliminary announcement and the official reconstitution date. We do not. Index-switcher volume does not gradually ramp up as reconstitution day approaches in Figure 2. It suddenly spikes on Russell's reconstitution day and then disappears the day after.

When FTSE Russell reconstitutes the Russell 1000 and 2000 in late June each year, any fund benchmarked to or evaluated against either index has to rebalance. This statement applies to ETFs as well as to mutual funds, pension funds, sovereign wealth funds, and endowments. The collective rebalancing volume coming from all these other kinds of funds is substantial.

The question is why it would all get executed at market close on Russell's reconstitution day alongside the rebalancing trades coming from ETFs. If non-ETFs are the culprit, it would just deepen the puzzle.

⁹The 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*.

The middle panel of Figure 6 tells us that, by purchasing a day before Russell's reconstitution date, a Russell-2000 benchmarked fund could pocket an 80bp one-day return on each Russell 2000 addition. The right panel tells us that, by doing the same thing, a Russell-1000 benchmarked fund could pocket a 20bp one-day return on each Russell 1000 addition. Such foregone profits represent a meaningful drag on the performance of a low-cost index fund.

If non-ETFs were responsible for excess reconstitution-day volume, it would imply that a truly massive amount of capital was strictly indexed to the Russell 1000/2000 at market close each day. The following calculation helps illustrate this point. Suppose Russell-benchmarked non-ETFs were entirely responsible for index switchers' additional reconstitution-day volume, $volume_n(t_{reconst}) - avgVolumePastYear_n$.

For example, suppose every share of additional volume experienced by Lumentum on June 26th 2020 was sold by a fund pegged to the Russell 2000 and purchased by a fund pegged to the Russell 1000. If this were the case, then strict end-of-day Russell-1000 indexers would hold an additional 11.5*m* shares of Lumentum at market close. Given that Lumentum has a minuscule weight in the Russell 1000, the need to buy so many shares would imply that there is a truly massive amount of money strictly indexed to the Russell 1000 at market close each day.

If we perform the same exercise for all index switchers on June 26th 2020 and average across the results, we find that strict end-of-day Russell indexers would need to have \$3.7*t* in assets under management to explain all the additional volume that index switchers observe. This is $21 \times$ larger than the combined assets of all Russell-benchmarked ETFs in June 2020, \$177.1*b*.¹⁰ In fact, the total market value of all Russell 3000 stocks was \$32.5*t* in June 2020. So strict end-of-day Russell indexers would have to hold 11.5% of the entire market, $11.5\% \times $32.5t \approx $3.7t$.

There are lots of funds linked to the Russell 1000/2000 in some way. But it seems implausible that so many funds would be strict end-of-day indexers. Why would so many mutual funds, pension funds, sovereign wealth funds, and endowments be trading like ETFs when they do not face the same constraints?

¹⁰Our primary data source for ETF holdings is ETF Global. We thank Anna Pavlova and Taisiya Sikorskaya for cross-checking this number using their data from Morningstar.

4.7 Closet Indexing

When computing $ETFvolume_n$, we defined the set of Russell-benchmarked ETFs in Equation (3) as all ETFs which were either a) explicitly labeled by ETF Global as benchmarked to the Russell 1000 or 2000 or b) had "Russell 1000" or "Russell 2000" in their fund name. Perhaps this was 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. ETFs are popular because they offer investors a guarantee that the per share value of the fund will match the value of the index at the end of each day. This promise of zero end-of-day tracking error 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.

Brown, Cederburg, and Towner (2021) gives evidence that ETFs sometimes advertise themselves as something different from what their reported benchmark would suggest. For example, when you inspect their reported benchmark, an ESG Russell 1000 ETF is often just a Russell 1000 ETF minus one or two gun stocks. This kind of closet indexing cannot explain our results, though. In spite of the difference between the fund's advertising and benchmark, the ESG Russell 1000 ETF would still be captured as a Russell-benchmarked ETF in our data.

4.8 Classification Errors

Thus, the real concern is not closet indexing. The real concern is that we 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 we observe. Consider the thought experiment from Section 2.3. Suppose the rule we use to decide whether a fund is Russell benchmarked is as bad as it could possibly be. Suppose every fund that we said was not benchmarked to the Russell 1000/2000 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 means computing changes in holdings over a 6-month rather than a 5-day window. And it involves summing over all funds, $f \in AllFunds$, not just ETFs. Even still, when we calculate something analogous to "*ETFvolume_n*" in this way, it turns out that 179 index switchers still have excess reconstitution-day volume.

The calculation above suggests that, to explain our results by pointing to classification errors, every single fund in Thompson S12 would have to be a Russell-benchmarked ETF. Moreover, all the trading done by these funds over an entire 6-month window would have to take place on Russell's reconstitution day. And, even then, it would not always be possible to account for the total amount of excess reconstitution-day volume. It is far more plausible that there is lots of extra trading activity on Russell's reconstitution day.

4.9 Leveraged ETFs

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. On some otherwise quiet days, leveraged ETF trading can cause spikes in volume at the end of the trading day (Ivanov and Lenkey, 2018).

But we are not studying an otherwise quiet day in financial markets. "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."¹¹ Moreover, our estimates of *ETFvolume_n* are unchanged to the first decimal place when we redo our analysis including leveraged ETFs.

4.10 Use of Derivatives

It could be the case that ETFs are doing some of their rebalancing by entering into futures positions. This is possible in principle. However, in practice, several important market participants have told us that the market for Russell futures contracts is not large enough to accommodate all of the rebalancing demand coming from Russell-benchmarked ETFs on reconstitution day.

¹¹Chuck Mikolajczak. "Investors brace for annual Russell index rebalancing with pandemic imprint." *Reuters.* 06/18/2021.

As a result, rebalancing is done almost exclusively in the equity market. We can verify this observation in our data from ETF global. If a Russell-benchmarked ETF was taking sizable derivatives positions, then there would be a gap between this fund's net asset value and the value of its equity holdings. We find no such gap in our data, confirming that Russell-benchmarked ETFs are rebalancing by trading stocks.

Derivatives market makers might also be adjusting their hedge positions in response to changes in the Russell 1000/2000. This would result in additional trading on reconstitution day. However, when one of these market makers adjusts a hedge position, it requires buying/selling ≤ 1 share of the underlying stock. So it would be hard to explain a sizable chunk of the 3.15 extra shares traded on Russell reconstitution day.

4.11 Heartbeat Trades

Finally, ETFs often execute a trade using an in-kind transfer to a partner investment bank for tax reasons. The process, which goes by the name of "heartbeat trades", has received a lot of attention in the financial media. For example, a March 2019 Bloomberg article calls this "ETF tax dodge... Wall Street's dirty little secret."¹² And Moussawi, Shen, and Velthuis (2021) argues that the tax savings from heartbeat trades are one of the major drivers of the rise of ETF investing.

Heartbeat trades are an important feature of the ETF market. But they do not account for our main finding. If anything, the widespread use of heartbeat trades would cause us to understate the amount of excess reconstitution-day volume. We measure *ETFvolume_n* as the change in ETF holdings regardless of how these trades get executed. Trading volume in CRSP is based on market transactions. So, if some ETF rebalancing takes place via in-kind transfers, then the amount of reconstitution-day volume in CRSP coming from ETF rebalancing will be less than *ETFvolume_n*. Hence, the amount of excess reconstitution-day volume will be even larger than what we calculate.

It is also worth pointing out that, when executing a heartbeat trade, the investment bank that partners with the ETF still has to trade shares of the index switcher at some point. And we do not see evidence of a run-up in trading activity in the days prior to Russell's reconstitution day. We also do not see a burst of trading activity in the days immediately following reconstitution.

¹²Zachary Mider, Rachel Evans, Carolina Wilson, and Christopher Cannon. "The ETF Tax Dodge Is Wall Street's 'Dirty Little Secret'." *Bloomberg*. 03/29/2019.

5 What Investors Say

There is a lot 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 3.15 shares get traded by somebody else.

Even though the scale of excess reconstitution-day volume is on par with the annual growth of Russell-benchmarked ETFs, this phenomenon is hard to explain using existing information-based asset-pricing models. It is outside the scope of these 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 investors and journalists talk about Russell 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.¹³

There is no guarantee that investors' reasoning will be economically sound. However, the scale of excess reconstitution-day volume suggests that standard models are missing something important about the economics of indexing. And knowing how investors reason about a situation still helpful to any researcher who is trying to figure out what that something is (Chinco, Hartzmark, and Sussman, 2021). It is easier to investors' optimization problem when you know what they are trying to optimize.

¹³If 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 \$258*b*. 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."¹⁴

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."¹⁵

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."¹⁶ "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."¹¹

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 \$9*t* of investors' money, on concerns that the rejigs cause unhealthy concentrations of trading."¹⁷

But, as far as we can tell, investors are not generally aware that most of the spike in index-switcher volume 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 reconstitution-day volume. We have had discussions with major ETF providers, such as Vanguard. We have chatted with authorized participants (APs) for these ETFs. We have shown our results to proprietary traders who have executed trades based on index reconstitution events.

Everyone we talked to was aware that index switchers realize large spikes in volume on reconstitution day. Some of the people we talked to had never thought about where

¹⁴James Mackintosh. "How ETFs Swallowed the Stock Market." *The Wall Street Journal*. 08/19/2019.
¹⁵Victor Reklaitis. "Why Friday could be the year's biggest trading day." *MarketWatch*. 06/26/2015.
¹⁶Phil Mackintosh. "Who is trading on US markets?" *Nasdaq News* + *Insights*. 01/28/2021.

¹⁷Steve Johnson. "FTSE Russell to review \$9t US index shake-up." Financial Times. 07/19/2021.

all this reconstitution-day volume was coming from. Obviously, some of it is due to ETF rebalancing. But it is not obvious why a proprietary trader would care about exactly how much was coming from ETF rebalancing vs. other sources.

Of the people who had thought about the composition of reconstitution-day volume, most were under the impression that it was largely the result of ETF rebalancing. People in this group were surprised by our findings, but no one raised any serious methodological objections.

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 reconstitutionday volume is not widely appreciated by market participants. And this is useful information for any theorists trying to model this phenomenon.

Given the scale of excess reconstitution-day volume, the decisions of many investors must be related to this phenomenon somehow. But investors cannot be strategically reacting to a phenomenon they are wholly unaware of. Thus, it seems unlikely that investors are actively pricing in this relationship, which would be consistent with the small price movements that we document on Russell's reconstitution day in Figure 6. Moreover, the broad unawareness of excess reconstitution-day volume suggests that investors are not trying to decipher the effects of such supply shocks as in a Grossman and Stiglitz/Admati-type model.

6 Conclusion

Stocks that switch between the Russell 1000 and 2000 experience massive amounts of unexplained volume on Russell's reconstitution day. When a stock switches between these two indexes, ETFs benchmarked to either index rebalance. This ETF rebalancing activity increases an index switcher's reconstitution-day volume by 139% relative to its daily average the previous year. However, for each share of ETF rebalancing activity, an extra 3.15 shares get traded by someone else.

The standard information-based asset-pricing frameworks of Grossman and Stiglitz (1980)/Admati (1985) and Kyle (1985) cannot speak to the existence of excess reconstitution-day volume. At best, our findings suggest that these models are extremely incomplete. The dollar value of this excess volume on one trading day is on the same order of magnitude as the annual growth all Russell-benchmarked ETFs each year. Thus,

the existing models which are typically used to study the rise of index investing are *missing half the story!*

However, the existence of so much unexplained volume on Russell reconstitution days also poses some deep conceptual problems for anyone trying to understand the rise of indexing through the lens of a Grossman and Stiglitz/Admati- or Kyle-type model. Excess reconstitution-day volume pokes holes in the core economic logic in each modeling paradigm. Thus, 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*.
- Admati, A. and P. Pfleiderer (1991). Sunshine trading and financial market equilibrium. *Review of Financial Studies*.
- 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*.
- Banerjee, S. and I. Kremer (2010). Disagreement and learning: Dynamic patterns of trade. *Journal of Finance*.
- Barber, B. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common-stock investment. *Quarterly Journal of Economics*.
- 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*.
- Bessembinder, H., A. Carrion, L. Tuttle, and K. Venkataraman (2016). Liquidity, resiliency and market quality around predictable trades: Theory and evidence. *Journal of Financial Economics*.

- Bessembinder, H., K. Chan, and P. Seguin (1996). An empirical examination of information, differences of opinion, and trading activity. *Journal of Financial Economics*.
- 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*.
- Brown, D., S. Cederburg, and M. Towner (2021). Dominated ETFs. Working Paper.
- Brunnermeier, M. and L. Pedersen (2005). Predatory trading. Journal of Finance.
- 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*.
- Foster, D. and S. Viswanathan (1996). Strategic trading when agents forecast the forecasts of others. *Journal of Finance*.
- Gabaix, X. and R. Koijen (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. *Working Paper*.
- 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 crosssectional variation in Nikkei 225 weights. *Review of Financial Studies*.
- Grinblatt, M. and M. Keloharju (2001). What makes investors trade? Journal of Finance.

- 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*.
- Harris, M. and A. Raviv (1993). Differences of opinion make a horse race. *Review of Financial Studies*.
- Hong, H. and J. Stein (2007). Disagreement and the stock market. *Journal of Economic Perspectives*.
- Ivanov, I. and S. Lenkey (2018). Do leveraged ETFs really amplify late-day returns and volatility? *Journal of Financial Markets*.
- Kandel, E. and N. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*.
- 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*.
- Li, S. (2022). Should passive investors actively manage their trades? Working paper.
- Lyons, R. (1997). A simultaneous trade model of the foreign-exchange hot potato. *Journal of International Economics*.
- 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.
- Milgrom, P. and N. Stokey (1982). Information, trade, and common knowledge. *Journal* of *Economic Theory*.
- Moussawi, R., K. Shen, and R. Velthuis (2021). ETF heartbeat trades, tax efficiencies, and clienteles. *Working Paper*.
- Novick, B., S. Cohen, A. Madhavan, T. Bunzel, J. Sethi, and S. Matthews (2017). Index investing supports vibrant capital markets. *BlackRock Whitepaper*.
- Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *Journal of Finance*.

Pavlova, A. and T. Sikorskaya (2021). Benchmarking intensity. Working Paper.

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

- Scheinkman, J. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal* of Political Economy.
- Stambaugh, R. (2014). Presidential address: Investment noise and trends. *Journal of Finance*.
- Statman, M., S. Thorley, and K. Vorkink (2006). Investor overconfidence and trading volume. *Review of Financial Studies*.
- Wang, J. (1994). A model of competitive stock trading volume. *Journal of Political Economy*.
- 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*.