The Passive-Ownership Share Is Double What You Think It Is*

Alex Chinco[†] and Marco Sammon[‡]

July 19, 2022
[Latest version]

Abstract

We estimate that passive investors held at least 37.8% of the US stock market in 2020. This estimate is based on the closing volumes of index additions and deletions on reconstitution days. 37.8% is more than double the widely accepted previous value of 15%, which represents the combined holdings of all index funds. What's more, 37.8% is a lower bound. The true passive-ownership share for the US stock market must be higher. This result suggests that index membership is the single most important consideration when modeling investors' portfolio choice. In addition, existing models studying the rise of passive investing give no hint that prior estimates for the passive-ownership share were 50% too small. The size of this oversight restricts how useful these models can be for policymakers.

Keywords: Passive ownership, Index funds, Reconstitution day, Information-based asset-pricing models

*We would like to thank Malcom Baker, Markus Baldauf, Hank Bessembinder, David Brown, Rodney Comegys, Zhi Da, Tom Ernst, Slava Fos, Diego Garcia, Robin Greenwood, Valentin Haddad, Jiacui Li, Sida Li, Josh Mollner, Anna Pavlova, Julien Penasse, Matthew Pritsker, Nick Roussanov, Andrei Shleifer, Laura Veldkamp, Luis Viceira, Yajun Wang, and Jeff Wurgler as well as seminar participants at Baruch, the University of Amsterdam, and the Democratize Quant conference for extremely helpful comments and suggestions. A prior version of this paper was circulated under the title *Excess Reconstitution-Day Volume*.

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

[‡]Harvard Business School; mcsammon@gmail.com.

1 Introduction

How much of the US stock market do passive investors own? We answer this question using data about the closing volumes of index additions and deletions on reconstitution day. Our approach allows us to estimate the ownership share of strict end-of-day indexers—i.e., a subset of passive investors who mirror their chosen benchmark at the end of each trading day. We find that strict end-of-day indexers who are benchmarked to either the S&P 500 or the Russell 1000/2000 held 37.8% of the US stock market in 2020 as shown by the right-most black dot in Figure 1.

According to the Investment Company Institute (ICI), index mutual funds and exchange-traded funds (ETFs) owned 15% of the US stock market at year-end 2020 and 16% of the market at year-end 2021. Thus, our findings suggest that index funds account for less than half of all passively invested money. The combined AUM of all investment funds, both active and passive, amounts to 30% of the US stock market. "Other investors—including hedge funds, pension funds, life insurance companies, and individuals—[hold] the majority (70%). (ICI Fact Book, 2022)" Most passive investing is directly indexed by this remaining 70%.

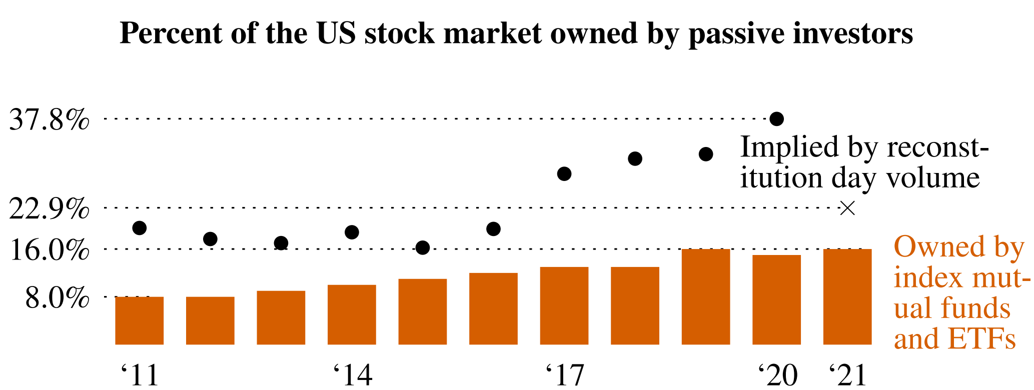

Figure 1. Two methods for estimating the percent of the US stock market owned by passive investors. Orange bars depict the ownership share of index funds as reported by the Investment Company Institute each year from 2011 through 2021. Black dots depict the implied ownership share of strict end-of-day indexers benchmarked to either the S&P 500 or the Russell 1000/2000 from 2011 through 2020. × mark depicts the implied ownership share of just the subset of strict end-of-day indexers who are benchmarked to the S&P 500 in 2021.

Imagine a passive investor who religiously matches his portfolio weights to his chosen benchmark (e.g., the S&P 500 or the Russell 1000) at market close each day. On a day where this index adds one stock (ticker:ADD) and drops another (ticker:DROP), our strict end-of-day indexer must rebalance his holdings immediately prior to close. Suppose he manages a \$1b portfolio. If ADD has an index weight of 0.05% after reconstitution, this strict indexer will have to buy \$500k worth of ADD shares at 4pm.

If there are also lots of other strict end-of-day indexers who behave just like this guy, ADD will experience a spike in closing volume on reconstitution day. And that is exactly what we observe in the data for index changers. Russell reconstitution day is now "generally considered [to be] the single-biggest trading day in US markets."[1] And > 90% of index-changer volume on reconstitution days gets executed at market close.

We convert the spike in volume at market close on reconstitution days into an estimate for the total amount of money benchmarked to a particular index. Suppose ADD's closing price was \$50 per share. In that case, the strict end-of-day indexer with \$1b in assets under management needed to buy (0.05% · \$1b) / (\$50 per share) = 10k shares to match his benchmark's weight. So, if ADD's total realized volume at the close was 10^4-times higher, $10^4 \times 10k = 100m$ shares, then it stands to reason that there is 10^4-times more money strictly indexed, which amounts to $10^4 \times \$1b = \$10t$.

Because we start from trading-volume data rather than data on fund holdings, our approach to estimating the passive-ownership share reflects both index-fund investors and direct indexers. Put differently, 37.8% reflects the money held in the iShares Russell 1000 ETF (ticker:IWB) and the money held in state pension funds that are directly indexed to the Russell 1000. This is how our 37.8% estimate for the passive-ownership share in 2020 can be more than double the previous estimate.

Yet, 37.8% is almost certainly too low. To ensure that the trading volume we analyze comes from index rebalancing, we narrowly focus on just the trading volume experienced by adds and drops right at market close on reconstitution days. But not all passive investors are strict end-of-day indexers. In principle, some passive investors could rebalance more slowly. And our approach does not reflect the holdings of these more relaxed passive investors. This is one reason why 37.8% is a lower bound.

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

Another reason is that it only reflects the holdings of strict end-of-day indexers who are benchmarked to either the S&P 500, the Russell 1000, or the Russell 2000. While these are important indexes, they are not the only indexes. The holdings of a strict end-of-day indexer who is benchmarked to the Nasdaq 100, for example, is not captured by our 37.8% headline number.[2]

There are two important takeaways from our analysis. The first is the magnitude of our main result. We estimate that passive investors own at least 37.8% of the US stock market. This is a massive number. It is more than double the widely accepted previous value of 15% at year-end 2020. The $\times$ mark in Figure 1 indicates that the subset of strict end-of-day indexers benchmarked to the S&P 500 owned more of the US stock market in 2021 (22.9%) than all index funds combined (16%).

These numbers imply that index inclusion is the single most important consideration when modeling portfolio holdings. As a point of comparison, note that Balasubramaniam, Campbell, Ramadorai, and Ranish (2022) find that all other investor- and stock-level characteristics combined only explain 4.0% of the variation in households' portfolio holdings. In other words, when trying to explain who holds what and why, index inclusion has an order of magnitude more explanatory power than all the other things that usually show up in our favorite asset-pricing models.

Passive investors are an important part of modern financial markets. Financial economists should care about the percent of the US stock market that is owned by passive investors for the same reasons that rain forest ecologists care about the relative biomass of insects in the canopy. These aggregate numbers matter. We need to get them right. That is just good science. Think about how different the macro-finance literature would look today if Mehra and Prescott (1985) had estimated an equity premium that was half as large—i.e., 4% rather than 8% per year.

The second important takeaway is methodological. The rise of passive investing has been one of the most talked about developments in financial markets. And this public discussion has triggered a tsunami of theoretical research.[3] These models make clear

[2] We would love to include passive investors benchmarked to other indexes in our calculations as well. However, the reconstitution timing of the Nasdaq 100 and similar indexes makes this difficult. The Nasdaq 100 reconstitutes once a year at market close on the third Friday in December. This timing coincides with the quadruple witching date. See Section 3 for further details.

[3] Here is a partial list of publications and working papers from just the past two years: Baruch and Zhang

predictions about how a doubling of the passive-ownership share should transform financial markets in economically meaningful ways. Yet, when researchers took these models to the data, they were using numbers for the US passive-ownership share that were 50% too small. We think it is noteworthy that no one noticed the problem.

All the papers in this literature are built on the same underlying logic. Suppose Alice starts out as an active investor who searches out and trades on news about firm fundamentals. If Alice becomes a passive investor, then she will stop doing these things. As a result, any information that Alice alone would have uncovered will no longer get incorporated into prices. Hence, prices will become slightly less informative about firm fundamentals after Alice's active-to-passive conversion.

One possible interpretation of our results is that this is the wrong way to think about passive investing. Perhaps researchers should not associate passive investing with uninformed traders in a Grossman and Stiglitz (1980)-type model. After all, State Street pays S&P Dow Jones billions of dollars for information about index membership (An, Benetton, and Song, 2022). Perhaps passive investors are just paying for and trading on a different kind of information.

Another possible interpretation is that existing theoretical models are correct but incomplete. Holding all else equal, the logic outlined above could hold for each new passive investor. But this is under the assumption that everything else is equal. And all else is never equal. So perhaps researchers should search for a new economic consideration that matters to market participants but is missing from existing models.[4] We view the fact that most passive investing is done via direct indexing as a natural starting point for this search.

Regardless of which interpretation you prefer, new methods are clearly needed when it comes to modeling the rise of passive investing. Existing models are not precise enough to recognize that the US passive-ownership share was off by a factor of two. The size of this blind spot poses a real problem for anyone trying to use these models to make policy decisions.[5]

(2021), Bond and García (2022), Buss and Sundaresan (2021), Chabakauri and Rytchkov (2021), Coles, Heath, and Ringgenberg (2022), Glosten, Nallareddy, and Zou (2021), and Lee (2021).

[4] An effective new drug can appear ineffective in clinical trials if doctors break protocol and systematically give it to their sickest patients. For financial theorists, the challenge would be to find the analog to the omitted variable, "severity of illness", when studying the rise of passive investing.

[5] Climate models make clear predictions about the sorts of bad future outcomes we can expect if global

1.1 Paper Outline

We begin our analysis in Section 2 by outlining our methodology. Then, in Section 3, we describe the data we use when applying this methodology to real-world financial markets. Section 4 contains our estimates for the share of the US stock market owned by passive investors. In Section 5, we examine how prices and liquidity change on reconstitution days. Section 6 discusses the broader implications of our main finding.

1.2 Related Literature

This paper borrows from and builds on several strands of related literature. First, there is the theoretical literature mentioned above, which studies the equilibrium effects of passive ownership on price informativeness and investor welfare. We find that previous estimates for the passive-ownership share in the US stock market were at least 50% too low. Yet the models in this literature gave predictions that were too imprecise to recognize it. This suggests that something is missing from our current theoretical understanding of how passive investing affects financial markets. It also underscores the utility of being able to estimate the effect of passive ownership on price informativeness in a relatively model-independent way (Sammon, 2022).

In addition to equilibrium effects, there are numerous direct effects of index-linked investing (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/deletions on return attributes 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). Our analysis suggests that, when analyzing these effects, researchers should consider trading by both index funds and direct indexers.

average temperatures were to rise by $x°$. "It is likely that the percentage of species at high risk of extinction will be 9% at 1.5°C, 10% at 2°C, 12% at 3.0°C, 13% at 4°C, and 15% at 5°C (Pörtner et al., 2022)" So, if we were to find out tomorrow that scientists had been accidentally underestimating the global average temperature changes by, say, 6°C for the past few years, then we would need to reevaluate these models. One out of every six species has not recently gone missing.

The ETF industry has experienced explosive growth over the past decade (Madhavan, 2016; Lettau and Madhavan, 2018). Many people believe that ETFs can distort the pricing of the assets that it holds. And several recent papers have provided evidence consistent with this hypothesis (Ben-David, Franzoni, and Moussawi, 2018; Da and Shive, 2018; Chinco and Fos, 2021). The effect sizes in these papers are large given that ETFs hold a relative small fraction of the US stock market (ICI Fact Book, 2022). We reconcile this inconsistency by documenting the existence of a large group of strict end-of-day indexers who trade like ETFs during reconstitution events.

Our work is connected to a related literature studying the distinction between active and passive investing. Sometimes active funds can trade like passive investors. Pavlova and Sikorskaya (2022) gives evidence that active funds trade more than you would expect in response to index reconstitutions. We show that direct indexers hold at least as much money as index funds. The combined effect of these two forces can produce extremely inelastic demand for index switchers on reconstitution day. This observation is related to Gabaix and Koijen (2022)'s inelastic market hypothesis.

Other researchers have pointed out that there is a minimum level of trading activity that even "passive" investors must engage in. And this minimum level is surprising high (Pedersen, 2018; Easley, Michayluk, O'Hara, and Tālis, 2021). It is also true that ETFs are some of the most actively traded assets in modern financial markets, and much of this trading activity comes from institutional investors (Robertson, 2019; Huang, O'Hara, and Zhong, 2021). Industry surveys regularly find that institutional investors are replacing index-futures positions with analogous positions in ETFs (Greenwich Associates, 2016). We view these findings as interesting but orthogonal to our analysis. We are estimating the ownership share of all passive investors. How much do these investors trade in aggregate? Who buys shares in index funds? These are interesting questions, but they are outside the scope of this paper.

Finally, this paper adds to the forensic-economics literature (Zitzewitz, 2012; Ljungqvist, Malloy, and Marston, 2009; Piskorski, Seru, and Witkin, 2015). When confronted with past errors, researchers usually respond by pushing for better reporting standards and data transparency. And we strongly support this prescription (Chen and Zimmermann, 2021). However, in this paper, we also build on the insight in Chinco (2022): unrecognized errors can inform model development going forward.

2 Methodology

This section outlines how we estimate the passive-ownership share for the US stock market by studying the volume experienced by index changers at market close on reconstitution day. We start by describing how we would approach the problem in an idealized setting where we could directly observe the rebalancing volume coming from passive investors (Subsection 2.1). We then discuss practical concerns that arise when implementing this thought experiment using real-world data (Subsection 2.2). Finally, we outline how we quantify the size of the resulting estimation errors (Subsection 2.3).

2.1 Thought Experiment

We are interested in determining the amount of money benchmarked to a specific value-weighted index, such as the S&P 500, the Russell 1000, or the Russell 2000. Once we can do this for one index, we can do the same thing for several of them and add up the results. Suppose that we could directly observe the rebalancing volume coming from passive investors, *RebalVolume*, whenever this index reconstitutes. In this subsection, we describe how it would then be possible to accomplish our goal.

Index Construction

In popular value-weighted benchmarks, such as the S&P 500 and the Russell, the index weights are float adjusted. Suppose that there are $N \geq 2$ stocks in an index, $n \in \{1, 2, \ldots, N\}$. Let $\mathit{FracCloselyHeld}_n$ denote the fraction of the nth stock's shares that are held by officers and directors, another publicly traded company, employee benefits programs, government entities including sovereign wealth funds, or individuals with a $\geq 5\%$ stake in the company. The number of tradable shares for stock n is thus:

$$\mathit{\#Tradable}_n = \mathit{\#Shares}_n \times (1 - \mathit{FracCloselyHeld}_n) \tag{1}$$

The market capitalization of an index is the total value of all tradable shares for the constituent stocks:

$$\mathit{IndexCap} = \sum_{n=1}^{N} \mathit{Price}_n \cdot \mathit{\#Tradable}_n \tag{2}$$

$Price_n$ denotes the nth firm's price per share at market close. The weight of the nth stock in the index corresponds to its share of the index's overall market capitalization:

$$IndexWeight_n = \left(\frac{Price_n \cdot \#Tradable_n}{IndexCap} \right) \tag{3}$$

Note that the level of the S&P 500 and the Russell 1000/2000 is not the raw market capitalization. It is a scaled version of this number:

$$IndexLevel = \frac{IndexCap}{Divisor} \tag{4}$$

The scaling factor, *Divisor*, is chosen each period to adjust for changes in index capitalization due to reconstitutions. If the scaling factor did not adjust following reconstitution, then an index's level would mechanically change following reconstitution if the added stock had a larger market cap.

Total AUM Indexed

Suppose that the benchmark index reconstitutes. At the end of trading today, one stock will get added to the index (ticker:ADD) and another will be dropped (ticker:DROP). Let $RebalVolume_{\text{ADD}}$ denote the amount of rebalancing activity experienced by ADD coming from passive investors who are benchmarked to the index as a result of reconstitution. For the thought experiment in this subsection, we are imagining that we can directly observe this quantity.

Reconstitution involves a change. So we need to consider two different time periods. We will use *Before* to indicate calculations made immediately prior to reconstitution when DROP is still part of the index. Think about these calculations as taking place at 3:59pm on reconstitution day. Likewise, we will use *After* to indicate calculations made immediately following reconstitution when ADD has been added to the index. Think about these as taking place at 4:01pm on reconstitution day.

Immediately prior to reconstitution, ADD had an index weight of 0%:

$$IndexWeight_{\text{ADD}}^{Before} = 0 \tag{5a}$$

$$IndexWeight_{\text{ADD}}^{After} = \left(\frac{Price_{\text{ADD}} \cdot \#Tradable_{\text{ADD}}}{IndexCap^{After}} \right) \tag{5b}$$

Let $\mathit{IndexWeight}^{\mathit{After}}_{\mathrm{ADD}}$ denote ADD's weight in the index immediately after being added with $|\Delta \mathit{IndexWeight}_{\mathrm{ADD}}| = \left|\mathit{IndexWeight}^{\mathit{After}}_{\mathrm{ADD}} - \mathit{IndexWeight}^{\mathit{Before}}_{\mathrm{ADD}}\right| = \mathit{IndexWeight}^{\mathit{After}}_{\mathrm{ADD}}$.

The total dollar value of the reconstitution-day rebalancing volume experienced by ADD is $\mathit{RebalVolume}_{\mathrm{ADD}} \times \mathit{Price}_{\mathrm{ADD}}$. So, to account for all the rebalancing activity coming from passive investors, we must have:

$$\underbrace{\overset{\text{known in thought exp}}{\mathit{RebalVolume}_{\mathrm{ADD}}} \times \overset{\text{observed}}{\mathit{Price}_{\mathrm{ADD}}}}_{\substack{\text{total dollar value of obs-}\\ \text{erved rebalancing volume}}} = \underbrace{\overset{\text{observed}}{|\Delta \mathit{IndexWeight}_{\mathrm{ADD}}|}}_{\substack{\text{required dollar value of}\\ \text{rebalancing per \$1 AUM}}} \times \overset{\text{unknown}}{\mathit{AUMindexed}} \tag{6}$$

This is an accounting identity. It states that the observed rebalancing volume must be equal to the required amount of rebalancing.

The key observation is that three of the four variables in this accounting identity are observable in this thought experiment. And that makes it possible to estimate the fourth unobservable object: the total amount of money benchmarked to the index, *AUMindexed*. Solving for this variable, we get:

$$\widetilde{\mathit{AUMindexed}}_{\mathrm{ADD}} \leftarrow \mathit{IndexCap}^{\mathit{After}} \times \left(\frac{\mathit{RebalVolume}_{\mathrm{ADD}}}{\#\mathit{Tradable}_{\mathrm{ADD}}}\right) \tag{7}$$

We write $\widetilde{\mathit{AUMindexed}}_{\mathrm{ADD}}$ rather than $\mathit{AUMindexed}_{\mathrm{ADD}}$ to indicate that we are working with an implied value based on this accounting identity. We write $\widetilde{\mathit{AUMindexed}}_{\mathrm{ADD}}$ rather than $\widetilde{\mathit{AUMindexed}}$ to emphasize that this implied value is based on the rebalancing volume observed for a single stock. The value implied by DROP should be the same.

2.2 Real-World Concerns

Now that we have an idea of how our estimation strategy would work in an ideal setting, we can discuss some of the problems that might arise when implementing it using real-world data. *IndexCap* and $\#\mathit{Tradable}_n$ are directly observable. Most of the difficulties lie in computing $\mathit{RebalVolume}_n$. Not all trading activity on reconstitution day comes from passive investors, and passive investors sometimes trade before or after reconstitution day.

To address these issues, we exploit an underappreciated fact about rebalancing in response to reconstitution events: most of this trading activity is prepositioned well in advance and occurs right at market close (Li, 2022; Madhavan, Ribando, and Udevbulu, 2022). Initially, financial intermediaries set up these prepositioned trades for ETFs due to tax considerations. But now lots of passive investors do this. Thus, by focusing our attention on closing volumes, we can capture the rebalancing volume coming from a subset of passive investors, which we call “strict end-of-day indexers”.

There are other reasons why trading volume might be high at the close. For example, the third Friday of March, June, September, and December are triple witching days when single-stock options, index options, and index futures all expire at the same time. On these days, there is high closing volume for reasons unrelated to any reconstitution events. This limits the number of benchmark indexes we can use in our analysis. Many indexes, such as the Nasdaq 100, always reconstitute on witching days. The Russell never does, and the S&P 500 frequently reconstitutes on non-witching days.

We restrict our analysis to only consider a subset of passive investors (strict end-of-day indexers) who are benchmarked to a subset of all indexes (the ones which reconstitute on non-witching days). These two restrictions imply that any estimates produced by our approach will represent lower bounds for the true passive-ownership share. Moreover, it is entirely plausible that these restrictions could lead to estimates that are lower than the ones produced using index-fund holdings.

So far, we have covered reasons why some of the trading on reconstitution days might not come from passive investors. We also need to consider reasons why all passive investors might not rebalance on reconstitution day. This can occur whenever a stock is added or dropped from an index for reasons other than market-cap changes. For example, if a company gets dropped from the S&P 500 because it was acquired, the official date of the drop may occur after the acquisition is finalized. In which case, some passive investors may be forced to divest in the company early. For this reason, we will focus our attention on index additions and deletions that are solely due to market-cap changes.

Last but not least, the number of tradable shares may be very different from total shares outstanding. Think about a stock where the founder retains a large stake (e.g., Tesla). Since $\#Tradable_n$ shows up in the denominator, mistakenly using $\#Shares_n$ for these sorts of firms will result in large errors. This is just one reason why replicating index weights is a non-trivial task. We will return to this observation in Section 6 when discussing the broader implications of our results.

2.3 Error Assessment

When we take our estimation strategy to the data, we take steps to ameliorate the various concerns outlined above. But how can we know if these steps were successful? We exploit the fact that additions and deletions should yield the same result

$$AUMindexed = \widetilde{AUMindexed}_{\text{ADD}} = \widetilde{AUMindexed}_{\text{DROP}} \tag{8}$$

to assess the size of our remaining errors.

Let $IndexWeight^{Before}_{\text{DROP}}$ denote DROP's weight in the benchmark index immediately before reconstitution:

$$IndexWeight^{Before}_{\text{DROP}} = \left(\frac{Price_{\text{DROP}} \cdot \#Tradable_{\text{DROP}}}{IndexCap^{Before}} \right) \tag{9a}$$

$$IndexWeight^{After}_{\text{DROP}} = 0 \tag{9b}$$

Right after reconstitution, the index allocates 0% of its portfolio to DROP.

Using the same accounting identity as before, we get:

$$\underbrace{\overset{\text{known in thought exp}}{RebalVolume_{\text{DROP}}} \times \overset{\text{observed}}{Price_{\text{DROP}}}}_{\substack{\text{total dollar value of obs-} \\ \text{erved rebalancing volume}}} = \underbrace{\overset{\text{observed}}{|\Delta IndexWeight_{\text{DROP}}|}}_{\substack{\text{required dollar value of} \\ \text{rebalancing per \$1 AUM}}} \times \overset{\text{unknown}}{AUMindexed} \tag{10}$$

Setting $|\Delta IndexWeight_{\text{DROP}}| = IndexWeight^{Before}_{\text{DROP}}$ and solving for $AUMindexed$ yields:

$$\widetilde{AUMindexed}_{\text{DROP}} \leftarrow IndexCap^{Before} \times \left(\frac{RebalVolume_{\text{DROP}}}{\#Tradable_{\text{DROP}}} \right) \tag{11}$$

Equations (7) and (11) both give the same basic recipe. To estimate the total amount of money indexed, first compute the fraction of tradable shares that were exchanged by passive investors. Then multiply the total index capitalization by this fraction.

If we could estimate $IndexCap$, $RebalVolume_n$, and $\#Tradable_n$ perfectly for every index switcher, Equation (8) would hold exactly. But errors in any of these empirical objects will generate slight discrepancies. By examining how large these discrepancies are, we can gauge the amount of noise in our estimation procedure.

3 Data

The previous section outlined our estimation strategy. In this section, we give an overview of the data we use as inputs to this procedure. Subsection 3.1 gives details about why we focus on S&P 500 and Russell reconstitutions. Subsection 3.2 describes what rebalancing looks like in the days and weeks around reconstitution events.

3.1 Reconstitutions

Our analysis focuses on reconstitution events associated with the S&P 500 as well as with the Russell 1000 and the Russell 2000 (the "Russell" for short). We now describe these two reconstitution events in detail. We also discuss why we limit our analysis to this pair of benchmark indexes.

S&P 500

The S&P 500 is a float-adjusted value-weighted index comprised of 500 large US companies (S&P Dow Jones Indices, 2022). The index is maintained by S&P Dow Jones Indices. All constituents of the S&P 500 are US companies that meet a minimum market-cap requirement and have a float-adjusted market cap that is at least 50% of this threshold. For example, in June 2022, the minimum market-cap threshold was \$14.6*b*, which means that a company must have \$7.3*b* in publicly available shares.

In addition, at least 10% of each S&P 500 company's shares must be publicly available at all times. Companies are also required to maintain adequate levels of liquidity and reasonable price levels. Prior to being added to the S&P 500, a company must also have positive earnings in the most recent quarter as well as positive earnings over the past four quarters when summed together.

Officially, the S&P 500 rebalances once a quarter on the third Friday in March, June, September, and December. These dates exactly line up with the quarterly triple-witching days. However, S&P 500 stocks can also be added and dropped from the index throughout the quarter on an ad hoc basis at the discretion of the S&P Dow Jones Indices committee. For example, on May 11 2020, Domino's Pizza (ticker:DPZ) replaced Capri Holdings (ticker:CPRI) in the S&P 500. We only use these sorts of intra-quarter reconstitution events which are due to changes in market cap.

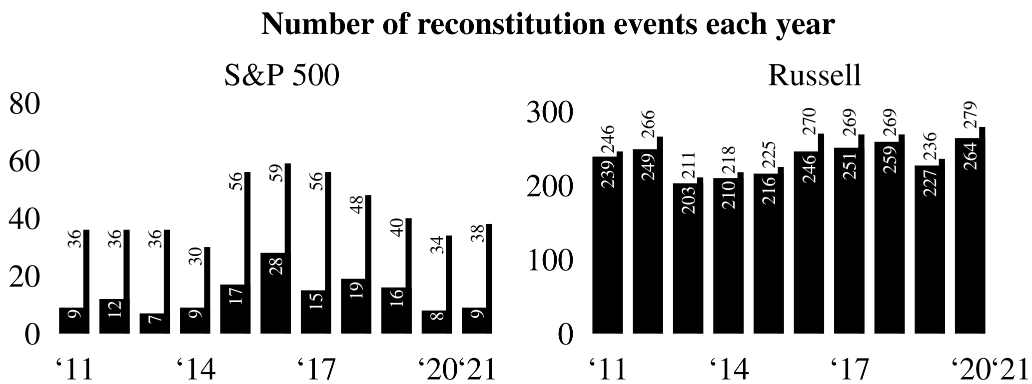

Figure 2. Number of reconstitution events each year for the S&P 500 (left panel) and the Russell 1000/2000 (right panel). Black numbers/thin bars denote all events. White numbers/thick bars denote events used to estimate the passive-ownership share for the US stock market. For the S&P 500, our data cover January 2011 through December 2021. For the Russell 1000/2000, our data cover January 2011 through December 2020.

S&P Dow Jones made 465 changes to the S&P 500 from January 2011 through December 2021. These changes involved 235 additions to the index and 230 deletions. If an existing S&P 500 firm adds a share class, no firm needs to be dropped from the index. This is why there are more additions than deletions. S&P Dow Jones released this information in 184 different announcements. Because $465 > 2 \times 184$, some announcements included news about more than one stock being replaced.

Panel (a) in Table 1 provides a breakdown of the reasons given by S&P Dow Jones for these changes. 185 of the 235 index additions were due to increases in market cap; whereas, 101 of the 230 index deletions were due to decreases in market cap. We exclude all other kinds of changes to the S&P 500 from our analysis. We cannot be sure that the reconstitution date will be the relevant rebalancing date for passive investors in these other cases. We further exclude any market-cap events that coincided with a triple-witching day. This yields 149 usable S&P 500 events: 101 adds; 48 drops.

We give summary statistics for stocks that get added to or dropped from the S&P 500 in Panel (b) of Table 1. We report the number of shares and market cap of each stock as of market close on reconstitution day. For index additions, we report the weight in the S&P 500 immediately after being added; for deletions, we report the weight immediately prior to being dropped. Average daily volume and cumulative returns over the past year are computed using trading days $\{t_{recon} - 22, \ldots, t_{recon} - 250\}$.

S&P 500 reconstitution events

Panel a)	Add		Drop
Market cap increase	185	Acquired by other firm	104
New firm created by spin off	21	Market cap decrease	101
Name change	15	Name change	14
New firm created via merger/acquisition	10	Too small after spin off	9
New share class	6	Other	4
Total	235	Total	230

Panel b)	All		Add		Drop	
	Avg	Sd	Avg	Sd	Avg	Sd
$\log_{10}(\#Shares)$	8.1	0.4	8.1	0.4	8.2	0.3
$\log_{10}(MarketCap)$	9.9	0.4	10.2	0.3	9.5	0.3
$IndexWeight$ [bps]	6.5	6.5	8.6	6.6	2.3	3.5
$\log_{10}(AvgVolume)$	6.2	0.5	6.0	0.5	6.4	0.4
$PastReturn$ [%]	27.8	62.5	52.3	59.2	−23.7	29.3
	149		101		48	

Table 1. Panel (a) classifies the reasons given by S&P Dow Jones for each of the 465 additions to and deletions from the S&P 500 index over the period from January 2011 through December 2021. When estimating the passive-ownership share for the US stock market, we only consider the 185 additions and 101 deletions due to market-cap changes. And, among these 296 changes that we made due to market-cap reasons, we further restrict our sample to the subset of 101 additions and 48 deletions that did not occur on a triple-witching day, leaving 149 S&P 500 events in total. Panel (b) reports summary statistics for the companies involved in these 149 events. Market caps and index weights are computed after reconstitution for additions and before reconstitution for deletions. Average daily volume and returns over the past year are computed using trading days $\{t_{recon} - 22, \ldots, t_{recon} - 250\}$.

S&P Dow Jones strategically chooses which companies to add to the index based on a variety of factors including industry composition and firm characteristics. Our analysis does not rely on the assumption that the firm being added is similar to the firm being dropped. We are studying how the trading volume of each of these two firms behaves in the time period around reconstitution.

Russell

The Russell 1000 and the Russell 2000 are both float-adjusted value-weighted indexes like the S&P 500. Both indexes are maintained by FTSE Russell. The Russell 1000 is comprised of 1000 largest stocks in the Russell 3000E universe, and the Russell 2000 contains the next 2000 largest stocks, which have rankings 1001:3000 in the Russell 3000E (FTSE Russell, 2022). Unlike the S&P 500, whose constituents are chosen by committee, membership in the Russell 1000 and 2000 is largely rule based.

The entire Russell family of US-equity indexes reconstitutes on the last Friday in June each year. Russell reconstitution day is the biggest trading day of the year. "Let's face it, for the New York Stock Exchange, Russell reconstitution, from a trading standpoint, is the greatest show on earth, that's where it all comes down."[6] FTSE Russell announces roughly two weeks in advance which stocks will be added directly to the Russell 1000 and 2000 as well as which stocks will switch between these two indexes. FTSE Russell makes its ranking decisions in late May.

For the same reasons as discussed above with the S&P 500, we restrict our attention to stocks that either a) get added directly to the Russell 1000/2000 or b) switch between these two indexes. We do not use stocks that drop out of the Russell 1000/2000 entirely. This group of stocks represents a group of firms which ceased to exist at some point during the previous quarter. So there is no guarantee that passive investors will rebalance their positions in these firms on reconstitution day. Many passive investors must divest their positions in these stocks at the time when the company ceases to be.

FTSE Russell made 2,364 changes to the Russell 1000/2000 from 2011 through 2020. Our data for the Russell does not include the June 2021 reconstitution day. We report summary statistics for these observations in Table 2. Number of shares and market

[6]Gordon Charlop, a managing director at Rosenblatt Securities, as quoted by Chuck Mikolajczak in "Investors brace for annual Russell index rebalancing with pandemic imprint." *Reuters.* Jun 18, 2021.

Russell reconstitution events

	All		2000 to 1000		Direct to 1000		1000 to 2000		Direct to 2000	
	Avg	Sd	Avg	Sd	Avg	Sd	Avg	Sd	Avg	Sd
$\log_{10}(\mathit{\#Shares})$	7.6	0.5	7.9	0.3	8.2	0.4	8.0	0.3	7.4	0.4
$\log_{10}(\mathit{MarketCap})$	8.7	0.6	9.7	0.1	9.8	0.4	9.1	0.2	8.4	0.3
IndexWeight [*bps*]	2.9	3.1	2.3	0.6	4.0	4.7	7.5	2.6	2.1	2.5
$\log_{10}(\mathit{AvgVolume})$	5.3	0.9	5.9	0.5	6.2	0.5	6.2	0.5	4.9	0.7
PastReturn [%]	58.5	121.7	68.5	71.2	25.4	47.9	−28.1	31.4	75.9	136.4
	2364		295		107		285		1677	

Table 2. Summary statistics describing changes to the Russell 1000 and Russell 2000 from June 2011 through June 2020 (10 reconstitution days in total). To estimate the passive-ownership share, we use data on stocks that a) switch from the Russell 2000 to the 1000, b) get added directly to the Russell 1000, c) switch from the Russell 1000 to the 2000, or d) get added directly to the Russell 2000. We do not use the information from stocks that drop out of the Russell 3000E. Shares outstanding and market cap are computed at market close on reconstitution day. Index weight reflects the stock's weight immediately following reconstitution. Average daily volume and returns over the past year are computed using trading days $\{t_{recon} - 22, \ldots, t_{recon} - 250\}$.

cap are computed at market close on reconstitution day. For stocks that enter the Russell 1000, we report the weight in the Russell 1000 immediately after they are added; for stocks that enter the Russell 2000, we report the weight in the Russell 2000 immediately after they are added to that index. Just like before, we calculate average daily volume and cumulative returns using trading days $\{t_{recon} - 22, \ldots, t_{recon} - 250\}$.

Even though the FTSE Russell uses a rules-based approach to determining membership in the Russell 1000 and the Russell 2000, there is no guarantee that a stock which just barely gets added to, say, the Russell 1000 will be similar to the stock that it replaces in ways other than size (Appel, Gormley, and Keim, 2020). We want to emphasize that our analysis in this paper does not rely on such an assumption. We study how the volume of each stock behaves in the time period around reconstitution.

Other Indexes

The S&P 500 and the Russell are popular benchmark indexes in the US equity space. But they are not the only indexes. To give a sense of the market landscape, we list the 15 largest US equity ETFs as of June 2022 as well as their benchmark indexes in Table 3. In addition to the S&P 500 and the Russell, these ETFs use various CRSP indexes as well as the Nasdaq 100.

In an ideal world, we would be able to study the rebalancing volumes of passive investors benchmarked to these indexes as well. However, for all the other popular benchmark indexes that we have looked at, the reconstitution timing seems to line up with the quarterly triple-witching date. For example, consider the CRSP US Total Market Index, which is used by the Vanguard Total Stock Market ETF (ticker:VTI). For this index, "reconstitution occurs quarterly after the market close on the third Friday of March, June, September, and December. (CRSP Indexes, 2022)" Likewise, the Nasdaq 100, which is used by the Nasdaq QQQ ETF (ticker:QQQ), reconstitutes "once annually…after the close of trading on the third Friday in December. (Nasdaq, 2021)"

We also note that, while Russell reconstitution day is now the biggest trading day of the year in US equity markets, only three of the largest 15 US equity ETFs are benchmarked to the Russell 1000/2000. This is consistent with our observation that most passive investing is not done through index funds. It is done via direct indexing.

Benchmark indexes for largest US equity ETFs

	Fund name		Benchmark	AUM
1.	SPDR S&P 500	SPY	S&P 500	$345*b*
2.	iShares Core S&P 500	IVV	S&P 500	$280*b*
3.	Vanguard S&P 500	VOO	S&P 500	$246*b*
4.	Vanguard Total Stock Market	VTI	CRSP US Total Market	$245*b*
5.	Invesco QQQ	QQQ	Nasdaq 100	$156*b*
6.	Vanguard Value	VTV	CRSP US Large Value	$93*b*
7.	Vanguard Growth	VUG	CRSP US Large Growth	$68*b*
8.	Vanguard Dividend Appreciation	VIG	S&P US Dividend Growers	$59*b*
9.	iShares Core S&P Small-Cap	IJR	S&P SmallCap 600	$59*b*
10.	iShares Russell 1000 Growth	IWF	Russell 1000 Growth	$58*b*
11.	iShares Core S&P Mid-Cap	IJH	S&P MidCap 400	$55*b*
12.	iShares Russell 2000	IWM	Russell 2000	$52*b*
13.	iShares Russell 1000 Value	IWD	Russell 1000 Value	$51*b*
14.	Vanguard Mid-Cap	VO	CRSP US Mid Cap	$47*b*
15.	Vanguard High Dividend Yield	VYM	FTSE High Dividend Yield	$43*b*

Table 3. Largest US equity ETFs as sorted by assets under management in June 2022. The data come from https://www.etf.com/channels/us-etfs.

3.2 Rebalancing

The conventional wisdom is that investors face a trade off between liquidity and immediacy. Based on this logic, you might expect that passive investors would try to gradually rebalance their holdings in anticipation of a scheduled reconstitution event. FTSE/Russell reconstitutes its Russell family of indexes on the last Friday of June each year. And they make an announcement weeks ahead of time to ensure that "stock-market trading [is] very orderly"[7] Yet we see no evidence that passive investors are gradually rebalancing their holdings. Instead, stocks that get added or dropped experience a huge spike in volume right at market close on reconstitution days.

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

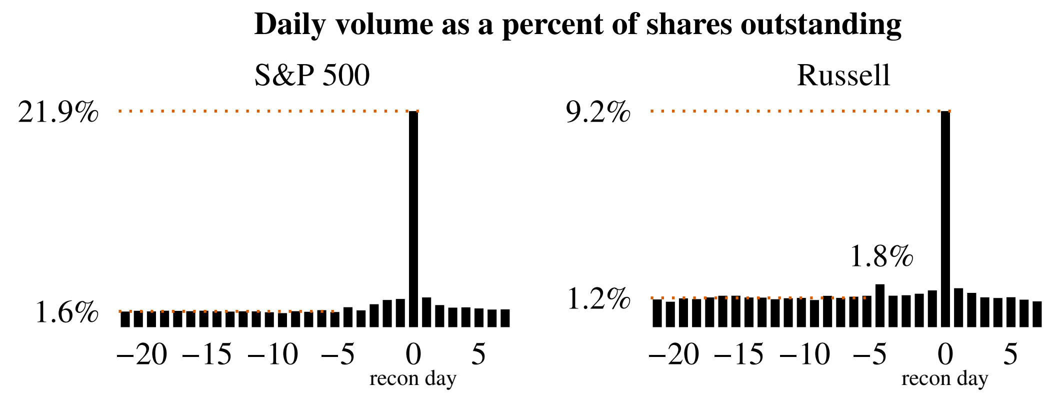

Figure 3. Volume relative to shares outstanding in the 30 trading days around a reconstitution event, $\{t_{recon} - 22, \ldots, t_{recon} + 7\}$. Left panel reports results for S&P 500 reconstitution events (sample period: January 2011 through December 2021). Right panel reports results for Russell reconstitution events (sample period: June 2011 through June 2020). There are two percentages listed on each y-axis. The smaller one is the average percent of shares outstanding that got traded on days $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$. The larger one is the percent of shares outstanding that got traded on reconstitution day (t_{recon}). The 1.8% value in the right panel is the percent of shares outstanding that got traded on day ($t_{recon} - 5$) for Russell reconstitution events. This date corresponds to the triple-witching day on the third Friday in June.

Daily Volume

Let's start with data on the daily trading volumes experienced by stocks that move between indexes. These data come from CRSP. Some passive indexers, such as index ETFs, have to rebalance their holdings on the reconstitution day itself. However, other kinds of index-linked investors, like a pension fund that is directly indexing, do not have to be so strict. You might imagine that passive investors who are not constrained to be strict end-of-day indexers would gradually rebalance their holdings in advance of reconstitution events. You might expect the volumes of stocks that get added to or dropped from the S&P 500 or the Russell to increase in the days prior to reconstitution.

This is not what the data show as indicated in Figure 3. The left panel shows that 1.6% of all shares outstanding get traded on an average trading day for the typical stock that will soon be added to or dropped from the S&P 500. Here, an "average trading day" means the average taken over days $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$. Likewise, for Russell stocks, 1.2% of all shares outstanding change hands on an average day.

Trading volume on reconstitution days

	$\frac{Volume_n}{AvgVolume_n}$		$100 \times \left(\frac{Volume_n}{\#Shares_n}\right)$		$100 \times \left(\frac{ETFvolume_n}{Volume_n}\right)$	
	Avg	Sd	Avg	Sd	Avg	Sd
All	26.4	34.9	9.9	8.7	27.6	208.3
S&P 500						
All	22.9	13.7	21.9	9.5	16.9	8.7
Add	26.8	14.1	21.3	10.7	18.1	8.2
Drop	14.6	8.1	23.2	6.3	14.3	9.1
Russell						
All	26.7	36.0	9.2	8.0	28.4	215.5
From 2000 to 1000	10.0	6.3	9.5	3.0	31.1	7.0
Direct to 1000	7.9	5.1	5.4	4.6	10.7	4.0
From 1000 to 2000	7.3	4.4	10.4	3.5	30.3	7.5
Direct to 2000	35.4	41.0	9.1	9.2	28.7	255.7

Table 4. This table reports cross-sectional statistics for different measures of trading volume (columns) for different kinds of reconstitution events (rows). These statistics are computed on reconstitution day itself. $\frac{Volume_n}{AvgVolume_n}$: Ratio of the volume on reconstitution day to the volume on an average trading day during the past year. How many days worth of volume was traded on reconstitution day? $\frac{Volume_n}{\#Shares_n}$: Ratio of reconstitution-day volume to shares outstanding. How much of the total supply of shares was traded on reconstitution day? $\frac{ETFvolume_n}{Volume_n}$: Ratio of ETF-rebalancing volume to total reconstitution-day volume. How many times more trading was there on reconstitution day than could be explained by ETFs alone? S&P 500 data cover the period from January 2011 through December 2021. Russell data cover the 10 reconstitution days from June 2011 through June 2020. ETF data comes from ETF Global. All other data comes from CRSP.

Figure 3 shows that there is very little change right up until the reconstitution day itself. Then, all of the sudden, trading activity explodes. 21.9% of shares outstanding get traded on reconstitution day for stocks that get added to or dropped from the S&P 500. 9.2% of shares outstanding get traded on reconstitution day for stocks that are involved in the Russell's annual event. The only minor uptick in trading volume prior to reconstitution day comes on day ($t_{recon} - 5$) for Russell reconstitution events. This date coincides with the third Friday in June, which is a triple witching day.

Table 4 reports summary statistics describing the amount of reconstitution-day volume experienced by stocks that get added to the S&P 500 or the Russell. Because trading volumes can differ across stocks by orders of magnitude, it is important to normalize these volume statistics. One way to do this is to look at the ratio of reconstitution-day volume to shares outstanding, $\frac{Volume_n}{\#Shares_n}$, like we did in Figure 3. i.e., how much of the total supply of shares was traded on reconstitution day?

Another common approach is to compute the ratio of the volume on reconstitution day to the volume on an average trading day during the past year, $\frac{Volume_n}{AvgVolume_n}$. We compute this average daily volume over the past year using data from trading days $\{t_{recon} - 22, \ldots, t_{recon} - 250\}$. This approach answers the question: How many "normal days" worth of volume got traded on reconstitution day? Table 4 reveals that, for stocks the move in and out of the S&P 500, 22.9 normal-days worth of volume occurs on reconstitution day. For Russell movers, 26.7 normal-days worth of volume changes hands on the last Friday in June.

An index ETFs benchmarked to the S&P 500 or the Russell 1000/2000 would have to rebalance on reconstitution day. And ETFs have become an important part of modern financial markets. The amount of money managed by this kind of fund as grown by over 500% during the past decade according the ICI Fact Book (2022). So you might wonder whether trading by this particular kind of index fund might be enough to explain the spike in reconstitution-day volume that we observe.

The final two columns in Table 4 show that this is not the case. These columns report the mean and standard deviation of the ratio of ETF-rebalancing volume to total reconstitution-day volume, $\frac{ETFvolume_n}{Volume_n}$. We compute $ETFvolume_n$ for a given stock using data from ETF Global. These data contain the end-of-day holdings of US ETFs. From these holdings, we impute the required rebalancing volume of all ETFs tracking a

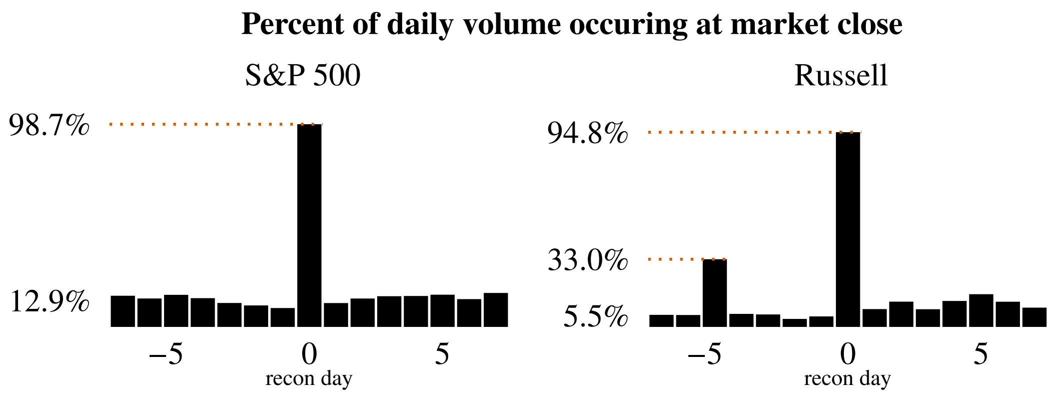

Figure 4. Volume at close as a percent of total volume during the 30 trading days around a reconstitution event, $\{t_{recon} - 7, \ldots, t_{recon} + 7\}$. Left panel reports results for S&P 500 reconstitution events (sample period: January 2011 through December 2021). Right panel reports results for Russell reconstitution events (sample period: June 2011 through June 2020). The two percentages listed on each y-axis in the left panel are the average percent of volume at the close on trading days $\{t_{recon} - 7, \ldots, t_{recon} - 1\}$ and the percent of volume at the close on reconstitution day (t_{recon}). The three percentages listed on each y-axis in the right panel are the average percent of volume at the close on trading days $\{t_{recon} - 7, \ldots, t_{recon} - 1\}$, the percent of volume at the close on trading day ($t_{recon} - 5$), and the percent of volume at the close on reconstitution day (t_{recon}).

particular index. There is roughly 6× as much volume on S&P 500 reconstitution days as can be explained by ETF rebalancing alone. On Russell reconstitution day, there is a little less than 4× more volume than required by ETF rebalancing.

Volume at Close

Of course, not all trading activity on reconstitution days has to come from passive investors who are rebalancing their holdings. You might worry that the sudden burst in reconstitution-day volume that we document in Figure 3 and Table 4 is being driven by other kinds of traders. For example, perhaps short-term arbitrageurs are entering the market to exploit the predictable demand coming from passive investors? If they are trading a lot among themselves during the course of a reconstitution day, then this would artificially inflate the daily volume statistics.

To address this sort of concern, we exploit the fact that most passive investors preposition their rebalancing trades to get executed right at market close (Li, 2022; Madhavan, Ribando, and Udevbulu, 2022). We use intraday data from TAQ to compute the fraction of a stock's trading volume during a given day with occurs either at the closing auction or in the minutes immediately after. Due to high volumes at close, some prescheduled trades do not officially hit the tape until a few minutes after market close. We define $VolumeAtClose_n$ as the trading volume from 4:00pm-4:20pm.[8]

Figure 4 shows that, on an average trading day, a stock which will move into or out of the S&P 500 will usually have 12.9% of its volume get executed at the close. However, on reconstitution day itself, 98.7% of the total daily volume listed in CRSP gets executed from 4:00pm-4:20pm. Note that this number can larger than 100%. When computing $VolumeAtClose_n$ using TAQ data, we are including trades that occur from 4:01pm-4:20pm which are classified as after hours by CRSP.

The analogous numbers for Russell reconstitution day are also stunning. On a typical trading day, a Russell index changer will have 5.5% of its volume at the close. On Russell reconstitution day, 94.8% of trading volume will take place from 4:00pm-4:20pm. We report a more detailed breakdown of these numbers in Table 5. For instance, we find that, for stocks that move from the Russell 2000 to the 1000, 8.2 "normal days" worth of volume occurs at market close on reconstitution day.

Figure 4 again reveals a smaller spike in closing volume on day ($t_{recon} - 5$). This is the triple-witching day on the third Friday in June. The jump in end-of-day volume on this earlier date is due to rebalancing associated with the expiration of stock options, index options, and index futures. These derivatives expire right at market close.

If a reconstitution event were to occur on a triple-witching day, we would not be able to distinguish end-of-day volume due to index reconstitutions from end-of-day volume due to derivatives trading. This is why we do not apply our approach to estimate the passive-ownership share of investors who are benchmarked to indexes, such as the Nasdaq 100, which reconstitute on triple-witching days.

There is obviously a lot of money benchmarked to other popular indexes. However, if we were to include these indexes in our analysis, we would tend to overstate the passive-ownership share. And we are trying to be as conservative as possible. In spite of these precautions, in the following section, we show that the passive investors own twice as much of the US stock market as previously thought.

[8]That being said, we estimate that 94% of the closing volume for index changers on reconstitution day occurs right at 4:00pm. i.e., for every $1m$ traded at 4:00pm, $60k$ shares get traded from 4:01-4:20pm.

Volume at close on reconstitution days

	$100 \times \left(\frac{VolumeAtClose_n}{Volume_n}\right)$		$\frac{VolumeAtClose_n}{AvgVolume_n}$		$100 \times \left(\frac{VolumeAtClose_n}{\#Shares_n}\right)$	
	Avg	Sd	Avg	Sd	Avg	Sd
All	95.0	40.7	27.2	38.3	9.3	6.3
S&P 500						
All	98.7	33.8	24.0	18.3	21.4	9.0
Add	103.2	32.3	28.7	19.4	21.7	8.9
Drop	89.0	35.1	14.2	10.2	20.9	9.4
Russell						
All	94.8	41.1	27.5	39.5	8.5	5.1
From 2000 to 1000	75.3	29.2	8.2	6.9	7.5	4.2
Direct to 1000	72.6	36.7	8.4	6.9	3.7	3.7
From 1000 to 2000	70.8	33.9	5.4	4.7	7.7	5.6
Direct to 2000	104.4	40.9	38.2	45.0	9.1	5.0

Table 5. This table reports cross-sectional statistics for different measures of volume at close (columns) for different kinds of reconstitution events (rows). These statistics are computed on reconstitution day itself. $\frac{VolumeAtClose_n}{Volume_n}$: Ratio of the volume at market close on reconstitution day to the total volume on reconstitution day. $\frac{VolumeAtClose_n}{AvgVolume_n}$: Ratio of the volume at close on reconstitution day to the volume on an average trading day during the past year. How many days worth of volume was traded at market close on reconstitution day? $\frac{VolumeAtClose_n}{\#Shares_n}$: Ratio of the volume at close on reconstitution day to the total number of shares outstanding. How much of the total supply of shares was traded right at market close on reconstitution day? S&P 500 data cover the period from January 2011 through December 2021. Russell data cover the 10 reconstitution days from June 2011 through June 2020. Data on closing volumes comes from TAQ. All other data comes from CRSP.

4 Main Result

In this section, we plug real-world data into our methodological approach to produce estimates for the passive-ownership share in the US stock market each year from 2011 through 2021. Subsection 4.1 provides these annual estimates. Then, in Subsection 4.2, we characterize the precision of this finding.

4.1 Passive-Ownership Share

Every time an index replaces one stock with another, we get two estimated values for the amount of money strictly benchmarked to that index. We start with all stocks that get added to or dropped from an index. We then compute the float-adjusted number of shares outstanding for each of these stocks, $\#Tradable_n$. After that, we estimate the amount of rebalancing volume experienced by each of these stocks, $RebalVolume_n$. Finally, we multiply the index's total capitalization, $IndexCap$, by the rebalancing volume experienced by a single index changer as a fraction of float-adjusted shares outstanding.

We can directly observe the total market capitalization of the S&P 500 as well as the Russell 1000 and Russell 2000 on a daily basis. We can also calculate the float-adjusted shares outstanding for each stock in each of these indexes. This gives us $\#Tradable_n$ for the nth stock and $IndexCap$ for the index that it belongs to. For the $RebalVolume_n$ term in Equations (7) and (11), we use the nth stock's closing volume on reconstitution day. These three inputs yield a point estimate for the AUM strictly indexed to a particular benchmark at market close each day, $\widetilde{AUMindexed}_n$.

\$1t in AUM means something different today than it did 10 years ago. The US stock market has grown over this time period. So we scale our each estimate for $\widetilde{AUMindexed}_n$ by the market cap of the Russell 3000 at the time that particular estimate was made:

$$\widetilde{\%Indexed}_n = \frac{\widetilde{AUMindexed}_n}{Russell3000cap} \tag{12}$$

$\widetilde{\%Indexed}_n$ represents a point estimate for the US passive-ownership share based on the closing reconstitution-day volume experienced by the nth stock. The first three columns of Table 6 report the average $\widetilde{\%Indexed}_n$ associated with the S&P 500, the Russell 1000, and the Russell 2000 each year from 2011 through 2021.

Percent of the US stock market owned by passive investors

	S&P 500	Russell 1000	Russell 2000	Total	ICI Fact Book	Difference
2011	14.63	4.15	0.75	19.52	8	11.52
	(1.70)	(0.34)	(0.03)	(2.06)		[5.59]
2012	13.73	3.40	0.54	17.67	8	9.67
	(1.62)	(0.32)	(0.02)	(1.96)		[4.93]
2013	12.75	3.67	0.57	16.99	9	7.99
	(1.67)	(0.36)	(0.02)	(2.05)		[3.89]
2014	13.92	4.33	0.55	18.81	10	8.81
	(1.54)	(0.45)	(0.03)	(2.02)		[4.35]
2015	12.45	3.27	0.51	16.23	11	5.23
	(1.11)	(0.29)	(0.02)	(1.42)		[3.67]
2016	13.68	5.10	0.58	19.36	12	7.36
	(1.03)	(0.45)	(0.02)	(1.49)		[4.92]
2017	21.60	6.22	0.76	28.58	13	15.58
	(0.81)	(0.48)	(0.02)	(1.31)		[11.93]
2018	22.29	7.84	0.98	31.11	13	18.11
	(1.04)	(0.42)	(0.03)	(1.49)		[12.16]
2019	23.27	7.84	0.76	31.87	16	15.87
	(1.24)	(0.57)	(0.03)	(1.84)		[8.62]
2020	25.53	11.36	0.87	37.77	15	22.77
	(2.51)	(0.67)	(0.03)	(3.22)		[7.08]
2021	22.87				16	
	(1.69)					

Table 6. The first three columns in this table report the percent of the US stock market owned by strict end-of-day indexers benchmarked to the S&P 500, the Russell 1000, and the Russell 2000. Each row reports results for a different year during the period from 2011 through 2021. The column labeled "Total" reports the sum of these three percentages. These values correspond to the black dots in Figure 1. The "ICI Fact Book" column depicts the percent of the US stock market owned by index mutual funds and ETFs according to the Investment Company Institute (ICI Fact Book, 2022). These values correspond to the orange bars in Figure 1. "Difference" represents the percentage-point difference between the total ownership share of strict end-of-day indexers, which is based on closing volumes for index changers on reconstitution days, and the ICI Fact Book estimate, which is based on index-fund holdings. Numbers in parentheses are standard errors clustered by announcement. Numbers in square brackets are one-sided t-statistics for the null hypothesis that the difference is zero.

For example, Figure 2 shows that there were 17 S&P 500 additions and deletions that we could use to impute the passive-ownership share in 2015. In the first column of Table 6, we see that these 17 events implied that 12.45% of the US stock market was held by strict end-of-day S&P 500 indexers. The column labeled "Total" reports the total passive-ownership share associated with all three benchmarks we analyze.

The column labeled "ICI Fact Book" depicts the percent of the US stock market owned by index mutual funds and ETFs according to the Investment Company Institute (ICI Fact Book, 2022). "Difference" represents the percentage-point difference between the total ownership share of strict end-of-day indexers, which is based on closing volumes for index changers on reconstitution days, and the ICI Fact Book estimate, which is based on index-fund holdings.

For example, our approach implies that 16.23% of the US stock market was held by strict end-of-day indexers who were benchmarked to the S&P 500, the Russell 1000, or the Russell 2000 in 2015. By contrast, if you only looked at index-fund holdings, you would guess that passive investors only held 11% of the US stock market in 2015. This is a difference of 5.23%pt. This is the smallest difference in our sample period; yet, even this difference is highly statistically significant. The numbers in square brackets are one-sided t-statistics for the null hypothesis that the difference is zero.

Our 37.8% estimate for the passive-ownership share in 2020 is more than twice the previously accepted value of 15% in 2020. Table 6 shows that 2020 was not a fluke. Our estimate was a shade less than double the value given in ICI Fact Book (2022) for 2019. And it was more than double this value in 2018, 2017, 2012, and 2011. This is a very consistent pattern in the data. And it shows every sign of continuing into the future. In 2021, if we were to consider the subset of strict end-of-day indexers who were benchmarked to the S&P 500, then this group alone would have owned more of the US stock market (22.9%) than all index funds combined (16%).

4.2 Estimation Error

In the first four columns of Table 6, we report standard errors that are clustered by announcement. Based on these standard errors, the difference between our new estimate for the US passive-ownership share and the prior estimate based on index-fund holdings is highly statistically significant. There are two years with double-digit t-statistics.

In addition to looking at standard errors, we also compute the within-announcement volatility of our estimates of the share of the US stock market owned by strict end-of-day S&P 500 indexers. Table 7 gives the average and median of $\mathrm{Sd}(\widetilde{\%Indexed}_n \,|\, n \in \text{same announcement})$ across changes to the S&P 500. This analysis is motivated by Section 2.3. We should have $\widetilde{\%Indexed}_{\mathrm{ADD}} = \widetilde{\%Indexed}_{\mathrm{DROP}}$ for stocks in the same announcement. Any differences between these values represent estimation error.

Table 7 says that, if you selected an S&P 500 announcement at random, you should expect the range of passive-ownership shares implied by the stocks involved in this announcement to be ±4.42% on average. The median range is 3.61%. These results are based on announcements for which we have usable data about more than one of the stocks added or dropped. This means two or more stocks must move in or out for market-cap reasons on a day that did not coincide with a triple-witching day.

There are 184 announcements about changes to the S&P 500 constituent list from January 2011 through December 2021. The first number in Table 7 reports that 123 of these announcements involve at least one addition or deletion we can use to impute the passive-ownership share. The second number in Table 7 reports that 18 of these 123 events involve multiple changes announced at the same time.

In is important to emphasize that the numbers in Table 7 are standard deviations not standard errors. They represent the volatility of the noise in each observation. Our estimate for the US passive-ownership share each year is based on multiple observations. So the volatility of these estimates will be lower (precision will be higher). Table 6 shows that, in 2020, we estimate that the US passive-ownership share is 22.77%-points higher than previous estimates. In 2019, we estimate that it is 15.87%-points higher. In 2018 and 2017, we estimate that it is 18.11% and 15.58%-points higher, respectively. These differences are large when based on multiple observations of a noisy signal with noise volatility of 4.42%.

The numbers in Table 7 reflect the economic restriction that all adds and drops in a single reconstitution event should imply the same passive-ownership share. Any discrepancies are the result of measurement error. Our data are not perfect. Rebalancing around reconstitution events is subtle. We are glossing over details, such as tax considerations (Poterba and Shoven, 2002; Moussawi, Shen, and Velthuis, 2021). Table 7 shows that the resulting noise is too small to materially affect our main result.

Within-announcement volatility of estimates for S&P 500

#Announcements		$\mathrm{Sd}(\widetilde{\%Indexed}_n \,\vert\, n \in \text{same ancmt})$	
≥ 1 stock	≥ 2 stocks	Average	Median
123	18	4.42	3.61

Table 7. Within-announcement volatility of estimates for the passive-ownership share associated with the S&P 500. These results are based on the 18 announcements about changes to the S&P 500 for which we have usable data about more than one of the stocks involved. This means that two or more stocks were moved in or out of the index for market-cap reasons on a day that did not coincide with a triple-witching day. ≥ 1 stock: number of S&P 500 announcements where we have at least one usable data point over the period from 2011 through 2021. ≥ 2 stocks: number of announcements where we have at least two usable data points. "Average" is the average within-announcement volatility. "Median" is the median within-announcement volatility.

5 Prices and Liquidity

Given how much passive investors are trading right at market close on reconstitution days, a natural question is: "What happens to prices?" In Subsection 5.1, we show that the prices of index changers do not tend to move much on reconstitution day itself. So you might guess that liquidity must skyrocket for these stocks at the time of this event. However, in Subsection 5.2, we show how this does not at first seem to be the case. If anything, it looks like liquidity dries up a bit on reconstitution day. We then demonstrate how to resolve this puzzle (small price movements with massive volume) by recognizing a quirk in how standard liquidity measures are calculated.

5.1 Prices

In Subsection 3.2, we show that index changers tend to experience a huge spike in volume on reconstitution day. In this subsection, we document that there is no corresponding huge change in prices on reconstitution day. Prices are relatively stable on reconstitution day for stocks that move into and out of the S&P 500, the Russell 1000, and the Russell 2000.

Figure 5 shows the cumulative abnormal returns of stocks that get added to or dropped from the S&P 500 during our sample period. The left panel reports results for

Cumulative abnormal returns around S&P 500 reconstitution events

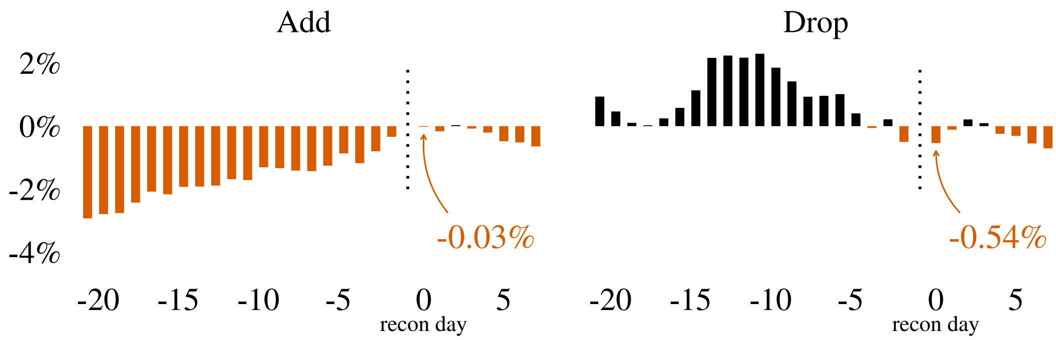

Figure 5. Cumulative abnormal returns in the 30 trading days around S&P 500 reconstitution events, $\{t_{recon} - 22, \ldots, t_{recon} + 7\}$. Left panel reports results for stocks that get added to the S&P 500. Right panel reports results for stocks that get dropped from the index. Sample period covers January 2011 through December 2021. Cumulative abnormal returns are defined relative to market returns and are set to zero at event time $\tau = -1$. Black bars denote positive values; red bars denote negative values. The percentage in the lower right-hand corner of each panel denotes the realized return in excess of the market on reconstitution day itself.

S&P 500 additions; the right panel reports analogous results for S&P 500 deletions. Abnormal returns are defined as realized returns minus the return on the market. To make it easier to compare results across the two panels, we normalize the cumulative abnormal returns to be zero on the day prior to reconstitution, $\tau = -1$.

The left panel of Figure 5 shows that additions to the S&P 500 experience positive abnormal returns in the month prior to reconstitution. However, on reconstitution day itself when all the trading takes place, S&P 500 additions only experience a −0.03% abnormal return as shown in the lower right-hand corner of the panel.[9] Stocks that get dropped from the S&P 500 experience a similarly small price decline on reconstitution day. These stocks have single day abnormal returns of −0.32% relative to the market.

Figure 6 shows the cumulative abnormal returns of stocks that move into or between Russell indexes. The top two panels report results for stocks that become a part of the Russell 1000 at the end of reconstitution day. The left two panels describe the cumulative abnormal returns of stocks that switch from one index to the other. The right

[9] Suppose that, immediately prior to reconstitution, you bought stocks that would be added to the S&P 500 and sold stocks that would be dropped. This trading strategy used to be quite profitable. However, its profitability has waned in recent years (Bennett, Stulz, and Wang, 2022).

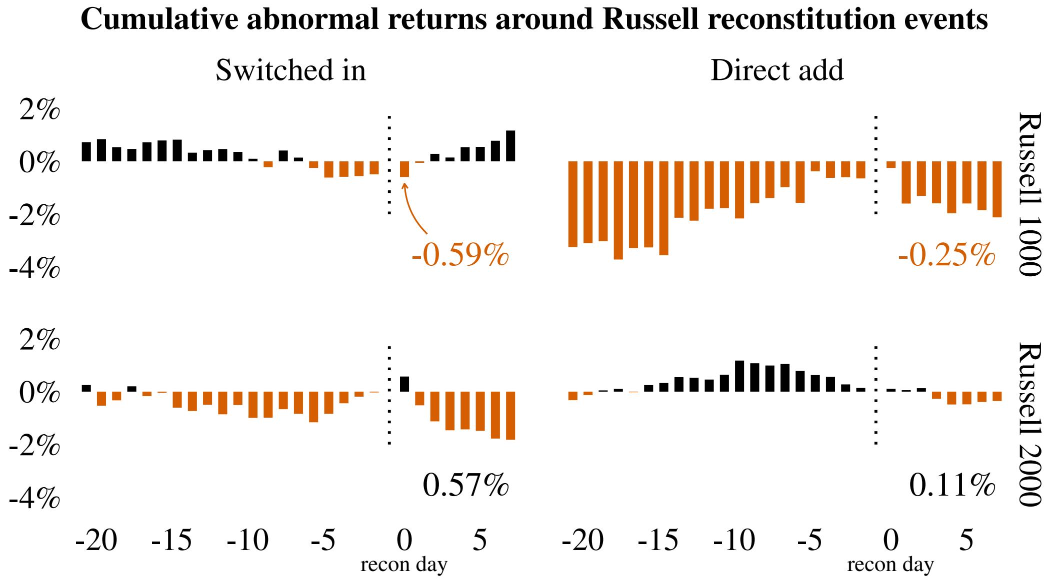

Figure 6. Cumulative abnormal returns in the 30 trading days around Russell reconstitution events, $\{t_{recon} - 22, \ldots, t_{recon} + 7\}$. Top panels report results for stocks that got added to the Russell 1000; bottom panels report analogous results for the Russell 2000. Top left panel reports results for stocks that moved from the Russell 2000 up to the Russell 1000; bottom left panel reports analogous results for stocks going from the Russell 1000 to the Russell 2000. Right panels report results for stocks that moved directly into one of the two indexes. Sample period covers June 2011 through June 2020. Abnormal returns are defined relative to the market. Cumulative abnormal returns are normalized to be zero at event time $\tau = -1$. Black bars denote positive values; red bars denote negative values. The percentage in the lower right-hand corner of each panel denotes the realized abnormal return on reconstitution day itself.

two panels report results for stocks that move directly into either the Russell 1000 or the Russell 2000.

Just like in Figure 5, we normalize the cumulative abnormal returns for each group of stocks to be zero on the day before reconstitution, $\tau = -1$. For stocks that get added directly to the Russell 1000 or to the Russell 2000, we see a steady run-up in prices over the month leading up to reconstitution (right two panels). We see an increase in prices for stocks that move from the Russell 2000 to the Russell 1000 (top left panel). And we see the opposite effect for stocks that move in the opposite direction. All of these longer-term trends are consistent with economic intuition. However, we again see muted price effects on reconstitution day itself.

5.2 Liquidity

While trading volume for index changers is through the roof on reconstitution day, their prices hardly move. So you might expect that standard measures of liquidity for these stocks would show that they have become much more liquid. Not so. In this subsection, we first outline what the data seem to show. Then, we describe the problem with this initial naïve analysis. Index changers are very liquid on reconstitution day.

Paradox

Figure 7 reports the median effective spread (top panels) and Kyle (1985)'s λ (bottom panels) in the 30 trading days around S&P 500 (left panels) and Russell (right panels) reconstitution events. The estimates come from the WRDS Intraday Indicators database, which is based on the millisecond TAQ database.

The effective spread is the difference between the executed price and the midpoint divided by the midpoint price. For each trade k on a given trading day for the nth stock, WRDS computes

$$2 \cdot \left(\frac{Price_n(k) - MidPoint_n(k)}{MidPoint_n(k)} \right) \times \{ IsBuy_n(k) - IsSell_n(k) \} \tag{13}$$

$Price_n(k)$ is the price at which the kth trade is executed for stock n. $MidPoint_n(k)$ is the national-best (NB) bid and ask for stock n at the time the kth trade took place. $IsBuy_n(k)$ and $IsSell_n(k)$ are indicators for whether the kth trade for stock n was a buy or a sell order. The effective spread is the dollar-weighted average of this measure across trades for stock n on day τ. We report this value in basis points in Figure 7.

Kyle's λ is the change in price per million shares traded. We take the values computed by WRDS using an intercept based on data during market hours (does not include 4:01pm-4:20pm). Using the observed collection of trades $k = 1, 2, \ldots, K$ on a single trading day for the nth stock, WRDS estimates the regression below:

$$Return_n(k) = \hat{\alpha}_n + \hat{\lambda}_n \cdot \log \$Volume_n(k) \times \{ IsBuy_n(k) - IsSell_n(k) \} + \hat{\varepsilon}_n(k) \tag{14}$$

The value of Kyle's λ on a particular trading day for the nth stock is the estimated value of $\hat{\lambda}_n$ from this regression.

Liquidity in days around reconstitution events

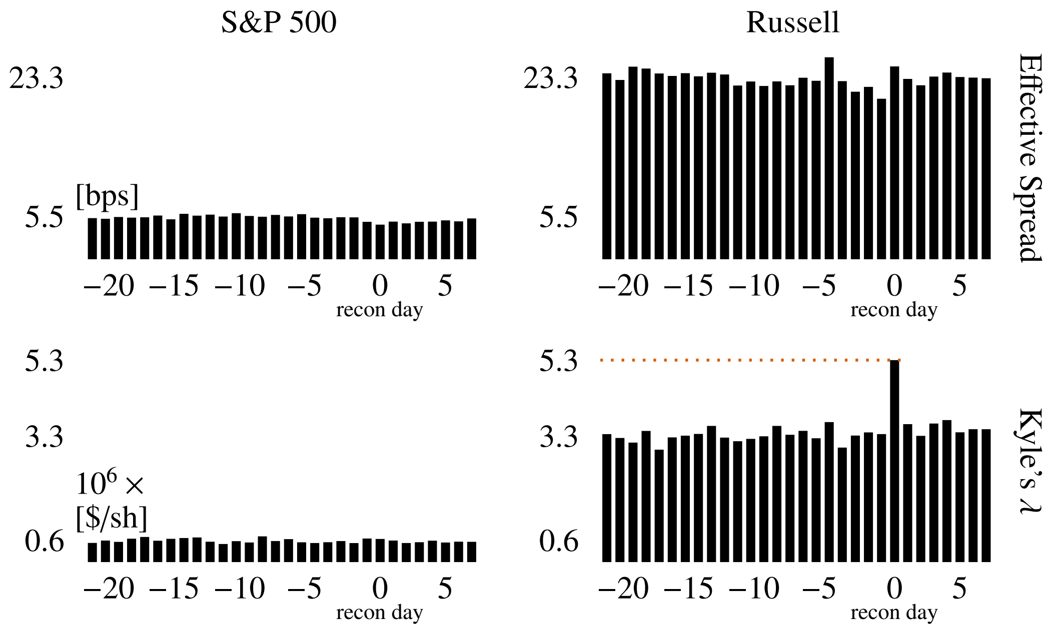

Figure 7. Two measures of liquidity in the 30 trading days around a reconstitution event, $\{t_{recon} - 22, \ldots, t_{recon} + 7\}$. The top two panels report the effective spread; the bottom two panels report Kyle's λ. The left panels report these values for S&P 500 reconstitution events; the right panels report them for Russell reconstitution events. The effective spread is the difference between the executed price and the midpoint divided by the midpoint price. We report this value in basis points. Kyle's λ is calculated as the change in price per million shares traded. Both statistics come from the WRDS Intraday Indicators database. Each bar reports the median value for all index changers on a given day relative to reconstitution. The average effective spread for additions and deletions to the S&P 500 on days $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$ was 5.5bps. For Russell index changers, it was 23.3bps over the same time period. The average value of Kyle's λ for additions and deletions to the S&P 500 on days $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$ was \$0.6 per million shares. For Russell index changers over the same time period, it was \$3.3 per million shares traded. In addition, Kyle's λ jumps to \$5.3 per million shares traded on Russell reconstitution day for affected stocks.

The average effective spread for additions and deletions to the S&P 500 on days $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$ was 5.5bps. For Russell index changers, it was 23.3bps over the same time period. The top two panels in Figure 7 seem to indicate that these average effective spreads do not change much on reconstitution day. The average value of Kyle's λ for additions and deletions to the S&P 500 on days $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$ was \$0.6 per million shares. For Russell index changers over the same time period, it was \$3.3 per million shares traded. The bottom left panel in Figure 7 indicates that Kyle's λ does not change on S&P 500 reconstitution days. Whereas, the bottom right panel shows that Kyle's λ jumps to \$5.3 per million shares on Russell reconstitution days.

Resolution

In Figure 7, we see that standard measures of liquidity on reconstitution days suggest that index changers become less liquid on reconstitution days. This is paradoxical. If they are less liquid, then where are all the large price changes given the huge spikes in reconstitution-day volume experienced by these stocks.

However, the results in Figure 7 are misleading. To compute effective spreads and Kyle's λ, researchers need to be able to sign trades. They need to be able to tell whether a particular trade was initiated as a buy order or as a sell order. The factor of $\{\, IsBuy_n(k) - IsSell_n(k) \,\}$ shows up in both Equation (13) and Equation (14). The standard way to do this is using the Lee and Ready (1991) algorithm. According to this rule, any trade that gets executed at a price much higher than the prevailing midpoint is classified as a buy order. Any trade that gets executed below the midpoint is a sell.

Here is the issue: if a trade gets executed either at the midpoint or at the closing auction, the Lee and Ready algorithm will not be able to classify it one way or the other. Such trades do not get used by WRDS Intraday Indicators when computing the effective spread and Kyle's λ. Thus, the standard liquidity measures that are widely used in the literature do not reflect closing-auction trading.

Figure 8 shows that, on an average day during the period $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$ prior to reconstitution, the Lee and Ready algorithm is able to sign 86.9% of all trades for additions and deletions to the S&P 500 and 90.6% of all trades for Russell index changers. However, on reconstitution day itself, the algorithm can only classify 20.6% and 23.7% of trades for these two groups.

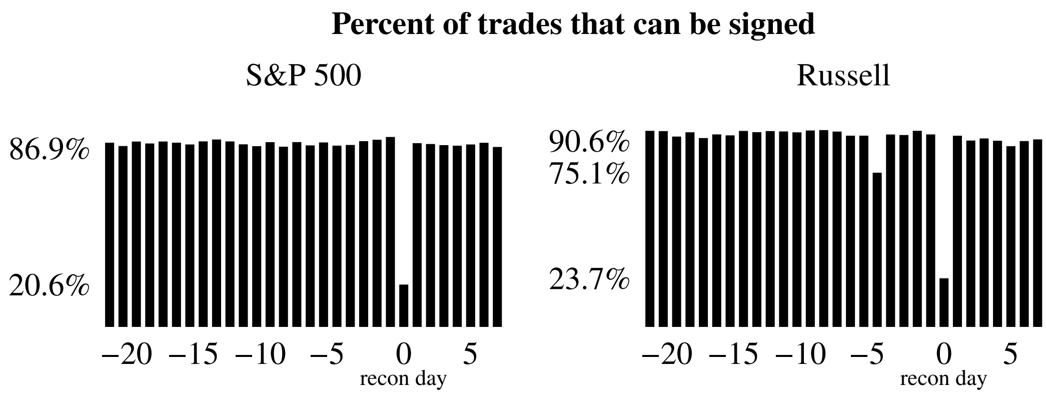

Figure 8. Percent of trades that can be signed using the Lee and Ready (1991) algorithm in the 30 trading days around a reconstitution event, $\{t_{recon} - 22, \ldots, t_{recon} + 7\}$. Left panel reports results for S&P 500 reconstitution events (sample period: January 2011 through December 2021). Right panel reports results for Russell reconstitution events (sample period: June 2011 through June 2020). On an average day during the period $\{t_{recon} - 22, \ldots, t_{recon} - 6\}$ prior to reconstitution, 86.9% of all trades can be signed for additions and deletions to the S&P 500; 90.6% of trades can be signed for Russell index changers during the same interval. However, on reconstitution day itself, only 20.6% of trades can be signed for S&P 500 additions and deletions. And only 23.7% of trades can be signed for Russell index changers on Russell reconstitution day.

Thus, while the measures on reconstitution days that we report in Figure 7 seem to suggest that index changers get slightly less liquidity on reconstitution days, these calculations do not reflect the fact that roughly 80% of the trades on reconstitution day cannot be signed. This is because these orders arrive as prepositioned trades made on behalf of strict end-of-day indexers. They are structured to be executed at the price which is determined by the closing auction on reconstitution day.[10]

The fact that the Lee and Ready algorithm fails to sign more than 80% of trades is consistent with the idea that it is incredibly cheap to trade large quantities of index additions and deletions right at market close on reconstitution day. This observation has important implications for the findings in this paper. It also suggests researchers must be careful when estimating liquidity on reconstitution days in other applications.

[10]Consistent with our earlier discussion, we note that only 75.1% of trades for Russell index changers can be signed on day ($t_{recon} - 5$), which corresponds to the triple-witching day on the third Friday in June. These trading days also involve prepositioned orders that are scheduled to be executed at market close. But, instead of coming from passive investors who are rebalancing in response to an index reconstitution, the prepositioned trades on day ($t_{recon} - 5$) stem from the expiration of derivatives contracts.

6 Discussion

We estimate that passive investors own at least 37.8% of the US stock market. This value is more than double the widely accepted previous estimate for the US passive-ownership share at year-end 2020, 15%. We conclude our analysis by discussing some of the broader implications of this main finding.

6.1 Who Holds What And Why

Consider a "simple" static portfolio-choice problem. The ith investor chooses the fraction of his initial wealth to invest in $N \geq 2$ risky assets so as to maximize his expected utility from consuming his terminal wealth:

$$\text{maximize} \quad \mathrm{E}[\,\mathrm{U}(\mathit{Wealth}^i \,|\, \boldsymbol{\theta}^i) \,\big|\, \mathit{InfoSet}^i\,] \tag{15a}$$

$$\text{subject to} \quad \mathit{Wealth}^i = \omega^i \times \{\,(1+\delta) + \textstyle\sum_{n=1}^{N} \mathit{Weight}^i_n \cdot (\mathit{Return}_n - \delta)\,\} + \cdots \tag{15b}$$

$$\mathit{InfoSet}^i = \{\,\{\delta, \omega^i, \boldsymbol{\theta}^i, \ldots\},\ \boldsymbol{Returns} \sim \mathrm{cdf}(\cdot),\ \ldots\,\} \tag{15c}$$

$\mathrm{E}[\,\cdot \,|\, \mathit{InfoSet}^i\,]$ denotes the ith investor's conditional expectation given his $\mathit{InfoSet}^i$. $\mathrm{U}(\mathit{Wealth}^i \,|\, \boldsymbol{\theta}^i)$ denotes this investor's utility from consuming Wealth^i after trading concludes. $\boldsymbol{\theta}^i$ represents the collection of parameters that define his utility function, such as time preference, risk aversion, etc. δ denotes the risk-free rate, and $\boldsymbol{Returns} = \{\mathit{Return}_n\}_{n=1}^{N} \sim \mathrm{cdf}(\cdot)$ denotes the joint distribution of returns.

Bach, Calvet, and Sodini (2020) argues that rich people (i.e., people with high ω^i values) tend to have different risk tolerances (i.e., different values of $\boldsymbol{\theta}^i$). Barber and Odean (2001) points to differences in holdings driven by innate differences in $\boldsymbol{\theta}^i$ across investors, such as gender. Malmendier and Nagel (2011) gives evidence that past experience influences $\boldsymbol{\theta}^i$. And Hong, Kubik, and Stein (2004) shows how social interactions can influence portfolio holdings via investors' beliefs about the distribution of returns, $\boldsymbol{Returns} \sim \mathrm{cdf}(\cdot)$.

All of these effects are interesting and important in their own right. But they each explain only a small fraction of investors' decisions. Our 37.8% estimate for the fraction of the US stock market owned by passive investors implies that that index inclusion is the single most important consideration when modeling portfolio choice.

By comparison, Balasubramaniam, Campbell, Ramadorai, and Ranish (2022) find that all other investor- and stock-level characteristics combined only explain 4.0% of the variation in households' portfolio holdings. Thus, when trying to explain who holds what and why, index inclusion has an order of magnitude more explanatory power than all the other things that financial economists have previously studied, $37.8\% \approx 10 \times 4.0\%$. If these other effects are important to include in realistic models of portfolio choice (and we think they are!), then index inclusion should be the first feature added to such models going forward.

6.2 Modeling the Rise of Passive Investing

Financial economists currently equate passive investors with the uninformed traders in a Grossman and Stiglitz (1980)-type model.[11] We think it is noteworthy that this line of reasoning gave no hint that previous estimates for the passive-ownership share were off by a factor of two. The size of this blind spot suggests that, at the very least, there is an important ingredient missing from these models.

It is easy to imagine a world where things were different. Look at the literature on the equity-premium puzzle (Mehra and Prescott, 1985). "Stocks have a substantially higher average return than bonds. Typical estimates put the equity premium between 4% and 8%. Even 4% is puzzling. Why do people not try to hold more stocks, given the power of compound returns to increase wealth dramatically over long horizons? (Cochrane, 2017)" Say what you want about the textbook consumption asset-pricing model (Merton, 1973), it clearly makes predictions that are precise enough to be in disagreement with the observed data. This fact has sparked a decades long search to figure out why.

By contrast, existing models of the rise in passive investing are not precise enough to recognize that previous estimates for the US passive-ownership share were 50% too low. This is a problem for researchers and policymakers who want to use these models to reason about financial markets. We hope the findings in this paper push theorists in this literature to tighten the range of predictions compatible with their models.

[11] This modeling approach starts with the optimization problem outlined in Equations (15a)-(15c). Then, it introduces a cost of information acquisition in place of the "$+\cdots$" in Equation (15b). It also assumes that the "$\ldots$" term in investors' information set in Equation (15c) includes the equilibrium price of each risky asset as well as the share of investors who acquired the costly signal.

6.3 Excess Trading Volume

Here is a slightly different way of looking at the problem at our main result poses for the existing way of thinking about passive investing. Suppose our interpretation of the closing volumes for index additions and deletions on reconstitution day is wrong. Imagine that these trades are not coming from passive investors and that the true passive-ownership share for the US stock market is on par with the numbers reported by the Investment Company Institute. Even in that world, the sudden spike in index-changer volume on reconstitution day would be problematic.

There is more trading activity in real-world financial markets than most models would predict (Milgrom and Stokey, 1982; Odean, 1999). But the usual explanations for this excess volume are based on differences of opinion (e.g., Harris and Raviv, 1993; Hong and Stein, 2007) and/or overconfidence (e.g., Barber and Odean, 2001; Scheinkman and Xiong, 2003).

In Figure 3, we see a massive spike in trading volume on reconstitution day itself and almost no increase in volume during the days leading up to it. Yet, all relevant information about which stocks will be added or dropped from an index is known well in advance of reconstitution day. There is precious little for traders to disagree over or be overconfident about by the time reconstitution day finally arrives.

6.4 The True Cost of "Free"

It could be that existing models for the rise of passive investing are correct but incomplete. But there is also another interpretation. Perhaps financial economists should not be thinking about passive investors as uninformed traders in a Grossman and Stiglitz (1980)-type model?

Looking around at how the index-fund industry operates, it seems like there is some merit to this interpretation. As Pedersen (2018) and Easley, Michayluk, O'Hara, and Tālis (2021) point out, passively tracking the market portfolio requires a substantial amount of trading. "New shares are issued, others are repurchased, and indices are reconstituted. (Pedersen, 2018)" And important market participants in the passive investing industry, such as Vanguard, BlackRock, and State Street, spend substantial sums of money managing this turnover in an intelligent way. There are sell-side analysts specializing in reconstitution events.

There are good reasons to think that holding the market portfolio is not "free". It is easy to find examples where investors have suffered major losses because they failed to appropriately account for index rebalancing.[12] Moreover, index providers are able to charge large fees for the privilege of tracking their benchmark (An, Benetton, and Song, 2022). So perhaps passive investors are still paying for information; it is just a different kind of information. Instead of paying for a signal about an asset's future payout, they might be paying for a signal about index inclusion.

6.5 Index Funds vs. Indexing

At the heart of our paper is an observation about how indexing is not the same thing as buying an index fund. Figure 1 shows that, in 2020, there was more money passively invested by direct indexers than was held by all index funds. This important distinction is currently being overlooked in the literature. And we hope it receives more attention going forward.

We study the closing volumes on reconstitution day of the stocks that get added or dropped from an index. These are the assets that ETFs hold. But the ETFs themselves also have surprisingly high trading volumes. The SPDR S&P 500 ETF (ticker:SPY) is one of the most actively traded securities in US markets.

One reason that ETFs get traded so much is that they provide institutional investors with a new way to build trades. For instance, "industry ETFs [can] help informed investors hedge industry risk. (Huang, O'Hara, and Zhong, 2021)" And surveys regularly find that institutional investors are replacing index-futures positions with analogous positions in ETFs (Greenwich Associates, 2016).

ICI Fact Book (2022) documents the explosive growth of the size of the ETF industry in recent years. We think this is a super important trend. However, we see it as important for different reasons. This trend does not reflect the rise of passive investing, although this too was going on in the background. Instead, it reflects a change in how institutional investors are building portfolios.

[12] Nathan Vardi. "Hedge Funds Suffered Losses As Index Rebalancing Trade Went Awry." *Forbes.* Mar 27, 2020.

7 Conclusion

We show how to estimate the share of the US stock market held by passive investors using data about the closing volumes of index additions and deletions on reconstitution day. We find that strict end-of-day indexers held 37.8% of the US stock market in 2020.

There are two important takeaways from our analysis. The first is quantitative. Our 37.8% estimate is more than double the widely accepted previous value of 15%, which represents the combined holdings of all index funds. What's more, 37.8% is a lower bound. The true passive-ownership share for the US stock market must be higher. All this implies that index inclusion is the single most important consideration when modeling portfolio holdings.

The second important takeaway is methodological. The rise of passive investing has been one of the most talked about developments in financial markets. There is now an entire theoretical literature which aims to understand the consequences for price informativeness and investor welfare. We think it is noteworthy that, when calibrated to the data, none of these models recognized that the assumed values for the US passive-ownership share were off by a factor of two. The size of this blind spot poses a real problem for anyone trying to use these models to make policy decisions.

References

An, Y., M. Benetton, and Y. Song (2022). Index providers: Whales behind the scenes of ETFs. *Working Paper*.

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

Bach, L., L. Calvet, and P. Sodini (2020). Rich pickings? risk, return, and skill in household wealth. *American Economic Review*.

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

Balasubramaniam, V., J. Campbell, T. Ramadorai, and B. Ranish (2022). Who owns what? a factor model for direct stockholding.

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*.

Baruch, S. and X. Zhang (2021). The distortion in prices due to passive investing. *Management Science*.

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*.

Bennett, B., R. Stulz, and Z. Wang (2022). Does joining the S&P 500 index hurt firms? *Working Paper*.

Bond, P. and D. García (2022). The equilibrium consequences of indexing. *Review of Financial Studies*.

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

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

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

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

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

Chen, A. and T. Zimmermann (2021). Open source cross-sectional asset pricing. *Critical Finance Review*.

Chinco, A. (2022). Model identification vs. market efficiency. *Working Paper*.

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

Cochrane, J. (2017). Macro-finance. *Review of Finance*.

Coles, J., D. Heath, and M. Ringgenberg (2022). On index investing. *Journal of Financial Economics*.

CRSP Indexes (2022, March). *Quarterly Performance: US Total Market Index*. CRSP Indexes.

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*.

FTSE Russell (2022, February). *Russell US Equity Indices*. FTSE Russell.

Gabaix, X. and R. Koijen (2022). In search of the origins of financial fluctuations: The inelastic markets hypothesis. *Working Paper*.

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

Gould, S. J. and E. Vrba (1982). Exaptation—a missing term in the science of form. *Paleobiology*.

Greenwich Associates (2016). Institutional investment in ETFs: Versatility fuels growth. *White Paper*.

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

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

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., J. Kubik, and J. Stein (2004). Social interaction and stock-market participation. *Journal of Finance*.

Hong, H. and J. Stein (2007). Disagreement and the stock market. *Journal of Economic Perspectives*.

Huang, S., M. O'Hara, and Z. Zhong (2021). Innovation and informed trading: Evidence from industry ETFs. *Review of Financial Studies*.

ICI Fact Book (2022).

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, C. and M. Ready (1991). Inferring trade direction from intraday data. *Journal of Finance*.

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*.

Ljungqvist, A., C. Malloy, and F. Marston (2009). Rewriting history. *Journal of Finance*.

Lo, A. (2004). The adaptive markets hypothesis. *Journal of Portfolio Management*.

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.

Madhavan, A., J. Ribando, and N. Udevbulu (2022). Demystifying index rebalancing: An analysis of the costs of liquidity provision. *Journal of Portfolio Management*.

Malmendier, U. and S. Nagel (2011). Depression babies: do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*.

Mehra, R. and E. Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics 15*(2), 145–161.

Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica*.

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*.

Nasdaq (2021, October). *Nasdaq-100 Index Methodology*. Nasdaq.

Odean, T. (1999). Do investors trade too much? *American Economic Review*.

Pavlova, A. and T. Sikorskaya (2022). Benchmarking intensity. *Review of Financial Studies*.

Pedersen, L. (2018). Sharpening the arithmetic of active management. *Financial Analysts Journal*.

Piskorski, T., A. Seru, and J. Witkin (2015). Asset-quality misrepresentation by financial intermediaries: Evidence from the RMBS market. *Journal of Finance*.

Pörtner, H., D. Roberts, H. Adams, C. Adler, P. Aldunce, E. Ali, R. Ara Begum, R. Betts, R. Bezner Kerr, R. Biesbroek, J. Birkmann, K. Bowen, E. Castellanos, G. Cissé, A. Constable, W. Cramer, D. Dodman, S. Eriksen, A. Fischlin, M. Garschagen, B. Glavovic, E. Gilmore, M. Haasnoot, S. Harper, T. Hasegawa, B. Hayward, Y. Hirabayashi, M. Howden, K. Kalaba, W. Kiessling, R. Lasco, J. Lawrence, M. Lemos, R. Lempert, D. Ley, T. Lissner, S. Lluch-Cota, S. Loeschke, S. Lucatello, Y. Luo, B. Mackey, S. Maharaj, C. Mendez, K. Mintenbeck, M. Moncassim Vale, M. Morecroft, A. Mukherji, M. Mycoo, T. Mustonen, J. Nalau, A. Okem, J. Ometto, C. Parmesan, M. Pelling, P. Pinho, E. Poloczanska, M.-F. Racault, D. Reckien, J. Pereira, A. Revi, S. Rose, R. Sanchez-Rodriguez, E. Schipper, D. Schmidt, D. Schoeman, R. Shaw, C. Singh, W. Solecki, L. Stringer, A. Thomas, E. Totin, C. Trisos, D. Viner, M. van Aalst, M. Wairiu, R. Warren, P. Yanda, and Z. Zaiton (2022). *Climate change 2022: Impacts, adaptation and vulnerability*. IPCC.

Poterba, J. and J. Shoven (2002). Exchange-traded funds: A new investment option for taxable investors. *American Economic Review*.

Robertson, A. (2019). Passive in name only: Delegated management and index investing. *Yale J. on Reg.*.

Sammon, M. (2022). Passive ownership and price informativeness. *Working Paper*.

Scheinkman, J. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*.

S&P Dow Jones Indices (2022, June). *S&P 500 Fact Sheet*. S&P Dow Jones Indices.

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*.

Zitzewitz, E. (2012). Forensic economics. *Journal of Economic Literature*.