Index Disruption: The Promise and Pitfalls of Self-Indexed ETFs*

Bige Kahraman† Sida Li‡ Anthony Limburg§

June 1, 2025

Abstract

Self-indexed ETFs, which track indices created and maintained by ETF issuers themselves, rather than relying on external index providers, are disrupting the index provision market. This paper asks whether this practice ultimately benefits end-investors. Contrary to proponents' expectations, self-indexing ETFs neither offer more competitive fees, nor enhance product variety in a way that improves financial returns for investors. Instead, our results highlight a conflict of interest: ETF issuers who also act as wealth managers use self-indexed ETFs to charge higher fees for otherwise comparable portfolios.

JEL Classification: G11, G14, G15, G23

Keywords: exchange-traded funds, ETFs, benchmarks, index providers, self indexing, indices, mutual funds, financial intermediation

*We thank the discussants and participants at the FMA Europe Conference, Midwest Finance Association (MFA) Annual Meeting, the SEC Conference on Financial Market Regulation (CFMR) and the University of Oxford FAME seminar for their useful comments. Anthony Limburg acknowledges support from the Clarendon Fund, the Templeton Education and Charity Trust (TECT) and the Saïd Business School Foundation. Sida Li acknowledges support from ACCESS for supercomputing allocations.

†Saïd Business School, University of Oxford. Email: bige.kahraman@sbs.ox.ac.uk.

‡Brandeis International Business School. Email: sidali@brandeis.edu.

§Saïd Business School, University of Oxford. Email: anthony.limburg@sbs.ox.ac.uk.

Conflicts of interest disclosure statements

Bige Kahraman

I have no conflicts of interest with respect to this research. For full disclosure, I hold a portfolio of ETFs, mutual funds and single stocks issued by organizations that may be affected by this research.

Sida Li

I have no conflicts of interest with respect to this research. For full disclosure, I hold a portfolio of ETFs, mutual funds and single stocks issued by organizations that may be affected by this research.

Anthony Limburg

I have no conflicts of interest with respect to this research. For full disclosure, I hold a portfolio of ETFs, mutual funds and single stocks issued by organizations that may be affected by this research.

1 Introduction

Traditionally, exchange-traded funds (ETFs) relied on indices constructed by established industry leaders, such as S&P, FTSE Russell, and MSCI. Wielding significant market power, these prominent index providers often levy substantial fees on ETF issuers which are subsequently passed onto end-investors. About one-third of all ETF fees is estimated to be paid to index providers in the form of licensing fees, and about 60% of these licensing fees are estimated to be markups.[1]

A key development disrupting the traditional index provision market has been the rise of self-indexing ETFs. Self-indexing refers to the practice where ETF issuers, or their affiliated entities, construct and manage their own proprietary indices, thus avoiding the expenses associated with licensing third-party benchmarks.[2] A crucial factor in the expansion of self-indexing ETFs was the SEC's 2013 decision to relax rules on self-indexing, making it easier for issuers to use their own benchmarks.

While relatively uncommon just a decade ago, self-indexing ETFs now account for nearly 20% of ETFs within popular investment styles like broad equity and large cap. This significant trend motivates the central question of this paper: whether self-indexing ETFs ultimately benefit their end-investors. Despite the significance of indices in passive investments, the index provision market remains understudied, and this paper aims to

[1] See An et al. (2023) for a structural estimation model.

[2] ETFs which rely on third-party index providers are referred to as public-indexing ETFs.

contribute to this literature by focusing on self-indexing in the ETF market.

Industry commentators who favor self-indexing celebrate the increased competition in the index provision market, arguing that this development can provide financial benefits to end-investors.[3] Self-indexing can empower issuers to increase product variety and offer more competitive fees due to the elimination of licensing costs. However, there are also skeptical views. One view points out that it can be difficult to compete down fees due to the inelasticity of demand for the industry leaders' indices (An et al., 2023) and another one highlights the potential conflicts of interest inherent in self-indexing which can ultimately harm investors (Xiao and Xiong, 2024).

To empirically examine these contrasting perspectives, we begin our empirical analysis by comparing the fees of self-indexed and public-indexed ETFs. Contrary to the proponents' expectation of lower fees, we find that self-indexing ETFs charge significantly *higher* net expense ratios compared to their public-indexing counterparts. Within our sample of passive U.S. equity funds, self-indexing ETFs have fees between 11% and 13% higher than those of public-indexing funds. Notably, this result is obtained after controlling for a host of fund characteristics as well as style, issuer, and time fixed effects.

Second, we test the notion that self-indexing provides issuers greater latitude in crafting investment strategies, and such differentiation potentially translates into financial advantages for investors. This may justify the higher fees charged by self-indexing funds.

[3]See, for instance, https://www.ft.com/content/e886b2d2-e852-3071-85c1-c9a57113d8a5

Analyzing daily holdings data from ETF Global within each investment style, we find no significant difference in the average pairwise portfolio similarity between self- and public-index ETFs, which contradicts the proponents' expectations. Using network analysis to visualize fund similarity, we observe that most self-index ETFs cluster closely with their public-index counterparts, reinforcing the lack of significant portfolio differentiation.[4]

Third, we evaluate the potential differences in fund performances using numerous types of portfolio and panel regression tests, and do not find any evidence that self-index ETFs perform better. Even when examining a subsample of self-index ETFs exhibiting slightly more differentiated portfolios, we find no evidence that they charge higher fees and/or deliver higher performance. A surprising finding is that while public-index ETFs resembling their peers charge lower fees (consistent with competition dynamics), more similar self-index ETFs paradoxically charge higher fees without offering any performance gains. In summary, our findings contradict the notion that the self-indexing practice offers financial advantages to end-investors through increased product variety, lower fees, or enhanced performance. Instead, our evidence reveals that investors, on average, pay higher fees for essentially comparable portfolios.

Our analysis then shifts to explore the potential conflicts of interest hypothesis. Our

[4]While our analysis show no evidence of genuine differentiation, clients may nonetheless perceive them as different, perhaps due to marketing and branding efforts by the issuing company. This would be in support of our next hypothesis on conflicts of interest.

central argument is that when self-indexing ETF issuers also operate as investment advisors, this dual role creates incentives for issuers to favor their own self-indexing ETFs, potentially at the expense of client interests. Our findings lend support to this idea.[5]

To examine this hypothesis, we start by testing whether self-index ETFs issued by issuers more susceptible to the conflicts of interest concern (due to their dual roles as investment advisors) charge higher fees. To test this idea, we classify issuers into two groups: 'specialized fund managers' and 'investment advisors'. Interestingly, investment advisors, likely due to their larger scale, generally charge significantly lower fees across their product range. However, a notable exception is their self-index ETFs, for which these issuers on average charge significantly higher fees. This aligns with their incentive to promote proprietary products.[6] An important finding is that, amongst ETFs issued by specialized fund managers, there are no significant differences in fees between self- and public-index ETFs.

Next we look at the characteristics of institutional investors holding shares of public- versus self-index ETFs. Under the conflicts of interest hypothesis, investment advisors favor their self-index ETFs in the advisory part of their businesses. These assets would appear in the 13F filings of the investment advisor to the extent that these assets are

[5]There are industry commentators who have raised concerns about other potential forms of conflicts of interest stemming from self-indexing. We do not focus on these issues as we do not find them particularly strong arguments. We instead highlight an issue which has been largely overlooked previously. We further discuss these matters in section 4.

[6]We also document that self-index ETFs offered by investment advisors tend to be more similar to their peer groups, without providing any significant performance benefits.

held in discretionary advisory accounts.[7] If this is the case, we would expect to find *Self-Ownership* by the issuing company (fraction of ETF's shares reported by the ETF issuing company in 13F filings) to be disproportionately high for self-index ETFs issued by investment advisors. On average, self-index ETFs have 18% self-ownership, while public-index ETFs have only 2%. Crucially, our regression analysis reveals that this significant difference in self-ownership is primarily attributable to self-index ETFs issued by investment advisors, further supporting the conflicts of interest hypothesis.

Finally, we close with a test in which we evaluate whether model portfolio recommendation is a channel through which issuers promote their self-index ETFs. Model portfolios are recommendation portfolios designed by asset managers, some of which play a dual role by simultaneously managing funds. Brogaard et al. (2021) show that asset manager model providers recommend their affiliated ETFs more frequently and hence enjoy significant flows especially from clients who are not particularly sensitive to fees. Building on these findings, we ask whether ETF issuers which are also model providers are more likely to include their self-index ETFs in their recommended model portfolios, compared to their public index ETFs. Our results show that this is indeed the case.

[7] Institutional investment managers in the United States that exercise investment discretion over $100 million in Section 13(f) securities are required to file Form 13F including the assets that they own or the client assets that manage. See Section 13(f)(1)(b) of the Securities Exchange Act, and 15 U.S.C. 78c(a)(35).

Related Literature

While there are numerous studies on ETFs, there is not enough research focusing on issues in the index provision market despite its relevance. Traditionally, most ETFs are public-index ETFs, that is, ETF issuers rely on external index providers and pass on the licensing fees to end-investors. Although ETF fees are seemingly low on an annual basis, even small differences in fees can have an important impact on long-term investments. Given the increasing popularity of ETFs with many households, it is crucial to examine how the index provision market affects investor welfare. Our paper contributes to this area of research by focusing on the emerging phenomenon of self-indexing.

An et al. (2023) are the first to study the competitive dynamics between index providers and ETF issuers, quantifying the impact on investor costs. Their estimates indicate that licensing fees to index providers constitute roughly one-third of ETF expense ratios, with approximately 60% of these fees representing index provider markups. Xiao and Xiong (2024) develop a general equilibrium model to explain the existence of index providers, arguing that they mitigate conflicts of interest arising from self-indexing by ETF issuers. While employing a conflict mechanism different to the one discussed in this paper, their model's central assertion, that index providers address inherent issues in self-indexing, is supported by our analysis.

Kostovetsky and Warner (2025) demonstrate that index brand significantly influences ETF investor flows, even when controlling for issuer characteristics. While acknowledging self-indexing ETFs, their study primarily focuses on externally indexed funds. This research, conversely, provides a comprehensive analysis of self-indexing ETFs, examining fees, investment strategies, performance, and potential conflicts of interest. Specifically, we delineate the conditions under which these conflicts manifest and their relationship to fee structures.[8]

This study is also related to a growing body of literature highlighting potential drawbacks within the ETF market. Focusing on high fees and significant underperformance by thematic ETFs, Ben-David et al. (2022) show that these ETFs cater to investor sentiment. In a similar vein, Huang et al. (2023) show poor performance by smart-beta funds and relate this finding to index providers' data mining practices. Akey et al. (2021), Easley et al. (2021) and Cheng et al. (2019) document closet active management of passive ETFs. Brown et al. (2024) document the surprising popularity of ETFs with returns that are highly correlated with those of cheaper, more liquid competitors. Khomyn et al. (2024) show that for a given index, ETFs with higher secondary market liquidity extract significant rents from investors who are arguably more sensitive to liquidity than to fees.

[8]Legal scholars are increasingly focused on the index market due to the significant expansion of passive investment products in the last decade. For example, Robertson (2019) offers a foundational analysis of the U.S. domestic equity index landscape, identifying the phenomenon of affiliated indices, among other observations such as "closet" passive funds. Furthermore, Mahoney and Robertson (2021) provide a valuable legal discussion concerning potential future regulations in this area. Our study diverges by addressing the economic question of whether end-investors experience net benefits (measured beyond fees) from self-indexing ETFs or are adversely affected by inherent conflicts of interest.

Finally, at a broader level, this study adds to research on fees in retail financial products, financial advice and conflicts of interest between financial advisors and their clients. This literature includes but is not limited to Bergstresser et al. (2009), Choi et al. (2010), Anagol and Kim (2012), Edelen et al. (2012), Christoffersen et al. (2013), Guercio and Reuter (2014), Gennaioli et al. (2015), Pool et al. (2016), Egan et al. (2019), Chalmers and Reuter (2020), Linnainmaa et al. (2021), Brogaard et al. (2021), Roussanov et al. (2021), Pool et al. (2022), Edelen et al. (2025). Different from these papers, we focus on the newly emerging phenomenon of self-indexing and the extent to which conflicts of interest stemming from the dual role of financial advisor and index designer can relate to costs for end investors.

2 Data and sample construction

Our main source of data is ETF Global which covers all ETFs listed in the U.S. and Canada with no survivorship biases. ETF Global provides information on a wide range of monthly ETF characteristics (e.g., fund name, issuer, inception date, benchmark index, AUM, leverage ratio, listing exchange, sector exposures, investment region, fund focus, asset class, active management dummy, currency and sector exposure, put and call options volume, short interest, management fee, and total/net expenses) as well as daily (monthly) holdings data for opaque (nonopaque) ETFs.

ETF Global has three key benefits over the CRSP Mutual Fund database. First and most importantly, it provides information on each ETF's benchmark index (variable labeled 'primary_benchmark'). This allows us to identify self-indexers as we will explain below. Second, the database provides daily ETF holdings, based on which we assess the similarity of investment strategies among ETFs. Thirdly, ETF Global updates fund characteristics monthly, enabling more accurate analyses. One downside of ETF Global is that the data are only available from 2012 onwards. However, this does not have an important impact on our study since self-indexing is a recent phenomenon. To illustrate, at the start of our sample in January 2012, we find only 13 self-indexing ETFs.

Our sample includes all passive unlevered U.S. equity ETFs from 2012 to 2020.[9] To identify the active (passive) status of the fund, we use the active management identifier provided by ETF Global. We remove all data points with missing information on fund primary benchmark and net expense ratios. We also eliminate cases where total expenses minus fee waivers are not equal to the ETFs net expenses.

We complement our data with supplementary information from Morningstar Direct, 13F institutional filings from Thomson/Refinitiv, and the CRSP stock and Mutual Fund databases. From Morningstar Direct, we extract a smart-beta identifier and model portfolio recommendations. We match approximately 90% of our sample with Morningstar using tickers and ETF names. We use 13F institutional filings and data on shares outstanding from the CRSP stock database, to calculate the share of institutional ownership

[9]It is ambiguous whether levered ETFs can be considered "passive". Nevertheless, we find that the underlying indexes (before leverage) are all public indexes in our data.

in our sample of ETFs. We use the CRSP Mutual Fund database to obtain daily, ETF share-split adjusted, NAV (net asset value) returns as well as to construct characteristics pertinent to ETF issuing investment management companies (e.g. age and total assets under management). Finally, we obtain data on factor returns from Kenneth R. French's website.

A self-indexed ETF is an exchange-traded fund that tracks an index created and maintained by the ETF issuer or one of its affiliated companies, rather than tracking an index provided by an independent third-party index provider, such as S&P Dow Jones Indices, MSCI, or FTSE Russell. A public-index ETF, on the other hand, is an ETF that licenses its index from an external third-party index provider.[10]

When the ETF market was still in its infancy, ETF sponsors had to obtain exemptive orders from the Securities and Exchange Commission to start offering self-index funds. This changed when new SEC orders, issued on July 10, 2013, permitted ETF sponsors and their underlying indexes to operate more flexibly than before. In particular, the new orders removed most of the requirements which the previous system had imposed on self-index ETFs.[11] In this way, the 2013 orders allowed self-index ETFs to operate in es-

[10]Our definition of self-indexing is based on the formal SEC definition, which states: "A "Self-Indexing Fund" is a Fund for which an Affiliated Person, or a Second-Tier Affiliate, of the Trust or a Fund, of the Adviser, of any Sub-Adviser to or promoter of a Fund, or of the Distributor (each, an "Affiliated Index Provider") will serve as the Index Provider" (SEC Release No. 30560)

[11]These conditions generally included: (i) requirements to promote transparency of the underlying index methodology (e.g., previous orders required the methodology governing the indexes to be publicly available at no charge, changes to such methodology to be made upon at least 60days' public notice, and index components to be publicly disclosed on a daily basis); (ii) use of an unaffiliated third party to calculate the index; and (iii) creation of certain firewall arrangements to separate the fund's adviser and the index provider.

sentially the same manner as ETFs based on unaffiliated, third-party indexes. Leveling the requirements for the two types of ETFs arguably made the self-indexing model more attractive, contributing to the observed trend over the last decade.[12] The 2019 “ETF Rule” by the SEC removed “exemptive order” regulations, instead modernizing the regulation of ETFs by establishing a clear and consistent framework for the majority of ETFs operating today. The 2019 rule did not change the status quo rules and requirements for self-index ETFs.[13]

We use the variable ‘primary_benchmark’ from ETF Global to identify ETFs that are self- and public-indexing. Our procedure involves multiple steps. Self-index ETFs are defined as ETFs which are using a benchmark index that is offered by the same or an affiliated company, offering the ETF. Our first step involves parsing the text reported in ‘primary_benchmark’, taking the last data point available for a given fund (that is, 2020’s cross section snapshot for alive funds and last available one for dead funds). If the text includes one of the key words corresponding to commonly used third-party indices such as “S&P”, “FTSE”, “Russell”, “CRSP”, “MSCI”, “Morningstar”, “Dow Jones”, “NYSE”, “NASDAQ”, “Nasdaq”, “StrataQuant”, we mark these observations as public-index ETFs. If the text in the benchmark does not include these keywords but instead includes the name of the issuer (or its affiliates), we then mark these as self-index ETFs.

[12] Although it is hard to pin down the exact motivation behind the SEC order, contemporary industry commentators interpreted it as conferring a significant advantage to ETF sponsors vying for market share in price-sensitive product segments. See, for instance, SEC Issues New Relief for Self-Indexing ETFs | Morgan Lewis - JDSupra.

[13] https://www.sec.gov/news/press-release/2019-190

For instance, the JP Morgan US momentum ETF (ticker: JMOM) is marked as a self-index ETF since its benchmark is "JP Morgan US Momentum Factor Index" and the issuer name is "JP Morgan". Second, for the cases that we were unable to identify through text parsing, we manually collect information on funds' benchmark indices using online sources and check whether the index is provided by the same (including its subsidiaries) or an affiliated company offering the ETF.[14] Third, to capture the changes in benchmark index over time for a given fund, we look at changes in the text reported in the 'primary_benchmark' variable and evaluate if it is a genuine change or a change that is simply capturing a name change or change in name abbreviation. We then manually update the self-index (public-index) identifiers accordingly. Through this procedure, we can identify the self-index (public-index) status of more than 99% of funds for which ETF Global provides a 'primary_benchmark'. This classification method allows us to define self-index ETFs in line with the formal definition offered by the SEC (see above). Our final sample includes 786 unique ETFs.

ETF Global also provides a 'focus' classification for ETFs which broadly classifies the investment style of a fund into one of 26 categories (e.g. 'Large Cap', 'High Dividend Yield', 'Broad Equity'), including sector-specific ETFs (e.g., 'Financials') as well as information on the issuing management company (issuer). For both variables we correct

[14]Examples of ETF issuers being affiliated with index providers include VanEck ETFs tracking Market Vectors indices, FlexShares ETFs tracking Northern Trust indices, and Lattice ETFs tracking Hartford indices.

spelling errors (e.g. 'JP Morgan' and 'JPMorgan' and 'High Dividend Yielld' and 'High Dividend Yield').

2.1 Descriptive statistics

While self-indexing is a recent phenomenon in the ETF industry, it has been gaining traction over the past decade. Figure 1 shows the number of self- versus public-index ETFs during our sample period. While there were only a few self-index funds a decade ago, by the end of 2020, this has increased to 96 ETFs in total. Table 1 shows the distribution of self- and public-index ETFs across investment styles. Self-index funds appear to be most prevalent in investment styles which are also popular among public-index funds. These include broad equity, large cap, and small cap.[15] Within these investment styles, self-index funds constitute about 20% of all funds by the end of our sample period. Due to their young age, self-index funds are still small in size compared to public-index funds, however our results on fund flows (presented later in the paper) show that their growth rates are on average comparable to those of public-index funds, conditional on similar fund characteristics.

We classify ETF issuers into three groups based on their mix of self- and public-index offerings: ones which offer only self-index funds, ones which offer only public-index funds, and ones which offer a mix of self- and public-index funds. Figure 2 shows the

[15]For expositional purposes, in this table, we group all sector-specific funds as 'sector' funds. This reduces the total number of investment styles to fifteen.

share of issuers in each of these groups over time. While most issuers only offered public-index ETFs a decade ago, this has been changing over time. The count of issuers which explicitly focus on self-index ETFs, and issuers offering both types of funds has been steadily increasing since 2012. By the end of our sample period, 51% of issuers in our sample are pure public-index issuers. 28% are pure self-index and 20% are issuers which offer a mix of public and self-index ETFs.

Internet Appendix Table IA.1 lists the 10 largest ETF issuers within each group.[16] Total issuer AUMs reported in this table reflect the total AUMs of ETFs issued by each of these issuers within our sample. Several large investment management companies such as Goldman Sachs, JP Morgan, State Street Global Advisors, and Fidelity stand out as issuers offering a mix of self- and public-index ETFs. For instance, within our sample of U.S equity passive ETFs, Fidelity offers 6 self- and 11 public-index ETFs; JP Morgan offers 5 versus 4; Goldman Sachs 3 versus 2; and State Street offers 3 versus 59. Other issuers with a mix offering include Invesco (2 versus 105), Victory Capital Management (7 versus 2) and Pacer Financial with (8 versus 4), and Northern Trust (5 versus 2), and Charles Schwab (1 versus 10).

When we look at the group of issuers which thus far offer only self-index funds, we see that these issuers tend to be smaller. John Hancock, Renaissance, and American Century are the largest ones within this group. The issuers which thus far offer only public index

[16]For brevity, tables reported in the Internet Appendix are labeled with abbreviation "IA".

funds include some of the largest players in the market such as BlackRock and Vanguard. This group however also includes smaller issuers such as Van Eck, DWS, Principal, and Nuveen.

In Table IA.2, we consider the full sample of issuers to examine the systematic relation between issuer characteristics and issuers' propensity to offer self-index ETFs. For each issuer, we extract the total AUM of all outstanding funds (across all asset-classes), as well as the age relative to the earliest offer date of any fund for each issuer from the CRSP Mutual Fund database. When we look at the role of issuer size (total issuer AUM) and issuer age, and we don't find statistically significant effects. Figure IA.1, which shows the distribution of issuer size and issuer age within each issuer group, provides insights into these findings. Self-index ETFs are primarily issued by either large and old or small and young issuers and this drives the coefficient estimates for issuer size and age towards zero in a linear regression model. The only significant effect the regression picks up is the pre-existence of a self-index ETF. Issuers with a self-index ETF are more likely to offer another one.

Panel A of Table 2 shows the descriptive statistics on various fund characteristics for self- versus public-index funds. Panel B reports summary information for our holdings similarity measure which we will define and discuss in the next section. $Net\ Expense\ Ratio\ (\%)$ is the annual net expenses after fee waivers, as provided by ETF Global. $Log(Age)$ is the natural logarithm of fund age, where age of an ETF is defined as the difference between the month-end date and the inception date, divided by 365. $Log(AUM)$ is the natural logarithm of fund's total assets under management. If information on AUM for an ETF is missing in ETF Global, we extract it from the CRSP Mutual Fund database. We then remove all remaining observations with missing or zero AUM. All variable descriptions are provided in Appendix A.

We examine ETFs' gross (before fees) NAV-based returns ($Gross\ Return\ (\%)$). Since ETFs subtract expense ratios from NAV on a daily basis, we first calculate daily gross returns by adding annual net expense ratio/252 (assuming 252 trading days in a year) to daily NAV net returns. Monthly gross returns are then calculated as the cumulative product of one plus each day's return over a full month, minus one. $Style\text{-}Adjusted\ Return\ (\%)$ is the fund's monthly return minus the average monthly fund return within the same investment style; $Gross\ Carhart\ Alpha\ (\%)$ is the fund's excess return estimated over a 36-month rolling window using the 4-factor Carhart model. Following the literature, we define fund flow as:

$$FundFlow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1} \cdot (1 + R_{i,t})}{TNA_{i,t-1}}. \qquad (1)$$

Public-index funds tend to have larger $Log(AUM)$, 19.4 versus 18.2, and they also tend to be older – $Log(Age)$ is 2 versus 1.37 on average. We examine a range of performance metrics such as raw returns, alphas and style-adjusted returns and do not observe significant differences in any of these performance metrics. In the next section, we will formally examine the systematic differences in fees and performances between the two types of ETFs in regression models where we control for various fund characteristics, investment styles, and time fixed effects.

This panel also reports summary statistics on our ownership variable (available at quarterly frequency). $Institutional\ Self\ Ownership$ is the fraction of total ETF shares reported in 13F filings of the ETF issuing company. Self-index ETFs have 14% self-ownership, while public-index ETFs had only 2% on average. Later in the paper, we formally examine the ownership differences between the two groups of ETFs.

3 Do self-indexing ETFs provide financial benefits to investors?

3.1 Differences in annual fees

We start by analyzing the extent to which fees charged by self-indexing ETFs differ from those charged by public-indexing ETFs. We regress net expense ratios on a number of explanatory variables including: $SelfIndexer?$, which is an indicator variable that equals one if the fund is identified as a self-index ETF, $Log(AUM)$, $Log(Age)$,

$Log(Holdings)$, $SmartBeta?$ (an indicator variable taken from Morningstar which equals 1 for smart beta ETFs) along with various month, style, and issuer fixed effects. Specifically, column 1 includes month-style fixed effects and column 2 adds issuer fixed effects alongside the control variables. Column 3 replaces issuer fixed effects with issuer-time fixed effects. Standard errors are clustered by month and issuer dimensions.

Table 3 reports the results. Throughout all regressions, the coefficient for self-indexing is positive and statistically significant. For instance, in column 2, the coefficient estimate for the self-index dummy variable is 0.052 and it is statistically significant at the 1% level. In column 3, the estimate is 0.047 and again it is statistically significant at the 1% level. Economically these estimates imply that self-index ETFs on average charge 11% to 13% higher than public-index ETFs. These results are interesting in that they represent contrary evidence to the popular view expressed by numerous industry commentators that self-indexing is a way for ETF issuers to offer more competitive fees.

At the end of our sample period, self-indexing ETFs held $43 billion in total assets under management. Consequently, the estimates in Table 3 suggest an approximate annual loss of $22 million for end-investors. This is currently modest since these funds are relatively new and young. However, given the growth trajectory of these funds (Figure IA.2), we believe that self-indexing ETFs warrant increased attention before larger losses are realized by end-investors.

Internet Appendix Table IA.3 reports the results of three robustness tests for our fee regressions. Our results remain similar in all tests. First, we ask whether these results might be driven by the possibility that some self-indexing ETFs are specialized ETFs which track bespoke indices (e.g. Solactive, Indxx) or specific subsets of assets within a category. To address this, we control for *Uniqueness*, which is a measure proposed in Kostovetsky and Warner (2025) and defined as the 12-month average of a fund's absolute difference in gross returns versus the value-weighted average gross return of all funds within the same style category. [17]

Second, we examine the potential role of fund activeness. As noted previously, our sample consists of passive ETF funds (as labeled by the data provider), and we control for smart-beta indicator in the regressions. However, recent papers document closet active management of passive ETFs (Akey et al., 2021; Easley et al., 2021; Cheng et al., 2019). To the extent that more self-indexed ETFs are 'closet' active, this phenomenon may contribute to higher average fees observed for self-index ETF funds. To examine this, we first turn to our holdings data and calculate annual turnover rates for all funds in our sample. The average annual turnover rates for self-indexers and public indexers are 28% and 25%, respectively. The similarly low turnover rate within the two subsamples indicates that self-indexing does not appear to be a device to 'secretly' market active ETFs. [18]

[17] As an alternative to the *Uniqueness* metric, we develop a holdings-based measure of similarity (or uniqueness) in the subsequent section. Controlling for this variable does not alter the primary finding of the fee regression. In fact, the results presented in Table 6 suggest an inverse relationship, wherein greater similarity (less differentiation) among self-indexed ETFs is associated with higher fees.

[18] We also look at the speed of fund trading around index reconstitution events and find that, similar to public-index ETFs, self-index ETFs' trade speed is also 1, meaning that they do not seem to strategically buy or sell stocks gradually, but all at once, which is typical for passive products.

Nevertheless, to remain conservative, we control for ETFs' yearly portfolio turnover ratio (*TurnRatio*) in our fee regression.

Finally, we control for the influence of ETF secondary market share liquidity, which was recently shown to be a key dimension of ETF competition (Khomyn et al., 2024). As a proxy for share liquidity, we use an ETFs bid-ask spread, *MAvgSpread*, which is defined as the difference between the ask and bid as reported in CRSP, scaled by the price. Throughout all tests, our conclusion remains intact.

To sum up, this section demonstrates that self-index ETFs charge *higher* (not lower) fees compared with their public-index counterparts. In the sections to follow, we first focus on testing whether self-index ETFs are differentiated products potentially providing superior performance, and later move to the question of conflicts of interest stemming from issuers playing the dual role of being an ETF issuer and financial advisor.

3.2 Differentiated product hypothesis

In section 3.1, we considered the possibility of self-index ETFs being specialized. In this section, we dive deeper into testing the hypothesis that self-indexing provides issuers greater flexibility in setting the investment strategy and whether such differentiation is associated with financial benefits to customers. Our analysis reveals that there is no significant difference in holdings between self-index and public-index ETFs, and that self-index ETFs do not provide performance benefits to customers. Interestingly, self-index ETFs that have less differentiated portfolios compared to other ETFs charge higher fees. Taken altogether, the results do not support the differentiated products hypothesis.

3.2.1 Differences in holdings

We start by introducing a cosine similarity measure between each pair of ETFs within our sample, to examine the differences in holdings. We pull the end-of-month portfolios of all ETFs in the sample from ETF Global and use all equity holdings with a valid ticker. For each pair of ETFs within our sample, we identify months where both ETFs have holding data (index for overlapped months is m) and calculate the cosine similarity between the two ETF's portfolios. Specifically, i is the index of stocks held by either ETF, ETF A's holding for stock i is $A_{i,m}$ dollars and ETF B's holding for stock i is $B_{i,m}$, where $A_{i,m}$ and $B_{i,m}$ equal zero if the ETF does not hold stock i. The cosine similarity is calculated as:

$$Cos_{A,B,m} = \frac{\sum_{i=1}^{N} A_{i,m} B_{i,m}}{\sqrt{\sum_{i=1}^{N} A_{i,m}^2} \cdot \sqrt{\sum_{i=1}^{N} B_{i,m}^2}} \tag{2}$$

N is the total number of stocks that either ETF A or B holds. By definition $Cos_{A,B,m} = Cos_{B,A,m}$. We take the numerical average of $Cos_{A,B,m}$ across all months m, and obtain our pairwise cosine similarity measure, $Cos_{A,B}$.[19]

[19]To confirm our calculations we checked the cosine similarity between well-established S&P 500 ETFs such as SPY, IVV, and VOO and find that their measures of cosine similarity are greater than 0.998.

Panel B in Table 2 reports the summary statistics of fund-level cosine similarity for self- vs public-index ETFs and tests the statistical significance of the difference between the two groups. To obtain a fund-level measure, for each fund, we average out all pairwise cosine similarities with all other funds within the same style. We do not find a significant difference between public and self-index ETFs in holdings similarity (0.263 vs. 0.267) in the overall sample. When we focus on popular investment styles (e.g., broad equity, large cap, and small cap), we observe the same finding. Except for the thematic style where the similarities are overall lower for both groups, we do not find significant differences in holdings similarities between the two groups of ETFs.

We next conduct network analyses to obtain further insights into the variation across funds, specifically, by visualizing the variation in holdings similarity across funds. Based on $Cos_{A,B}$, we calculate a minimum spanning tree representing the network of all ETFs in our sample. Intuitively, this method calculates a parsimonious network which connects all funds, minimizing the necessary connections, where more similar funds become bunched together, forming clusters (i.e. 'communities'). In Figure 3, we plot this network using the Fruchterman Reingold layout, which minimizes overlapping connections and spreads the network evenly on a canvas. Panel A shows the network structure in the full sample while Panels B to F focus on the most popular investment styles (listed above).

Figure 3 reveals that, self-index ETFs are generally centrally located within many clusters, in that their holdings are not significantly different from those of public-index counterparts. Had their holdings been substantially different compared with public-index ETFs, they would be systematically located outside the clusters, perhaps even forming their own. We observe this pattern when analyzing the full sample of ETFs (Panel A) as well as the subsamples based on investment styles (Panels B to F). For instance, in broad equity (Panel B), we observe two main clusters and self-index funds are generally bunched with public-index funds within these clusters. In a similar vein, for large cap and small cap ETFs, the majority of self-index funds seem to be located within the main cluster along with most public-index funds.[20]

To sum up, we do not find meaningful differentiation in portfolio holdings between self- and public-index ETFs within the same investment style. Although our analysis show no evidence of real differentiation, clients may nonetheless perceive them as different, perhaps due to marketing and branding efforts by the issuing company. This perception would align with our conflict of interest hypothesis, which we will examine in the next section.

3.2.2 Differences in performance

We next move to evaluating the potential differences in fund performances to see whether the higher expense ratios of self-index ETFs can be justified with better performance outcomes. Despite the similarity in portfolio holdings (as we show in section 3.2.1),

[20]In Table IA.4 we reproduce Table 3 (fee regression) with style fixed effects constructed based on the clusters from the network analysis. The results remain similar.

factors such as trading strategies could potentially lead to performance differences.

We first conduct portfolio tests where we estimate and contrast the monthly before-fee alphas of self- and public-indexing ETF portfolios over our sample period. We report results for both value- and equal-weighted portfolio returns, and we use CAPM, Fama-French and Carhart factor models to estimate alphas. In total, we estimate alphas under 6 different scenarios. All returns are before fees and we require at least 20 observations in each portfolio. Standard errors are Newey-West standard errors allowing correlation up to 6 months.

Table 4 reports the results. Panels A, B, and C show CAPM, 3-factor and 4-factor alphas, respectively. Columns 1 and 2 report monthly alphas. When we look at the monthly alphas of self-index ETF portfolios, we see that alphas are statistically indistinguishable from zero. A similar pattern is observed also for public-index ETF portfolios, except for one scenario where CAPM monthly alpha (with equal-weighting) is weakly negative. Finally, alphas of the spread portfolio – the portfolio that goes long on self-indexers and short on public-indexers – are statistically insignificant throughout all tests. Columns 3 and 4 reproduces the analysis using daily portfolio returns. Given the relatively short sample period of our study, this allows us to increase the statistical power of the tests. The results remain unchanged.

In Internet Appendix Table IA.5, we examine the potential differences between performances in a regression setting where we can control for the potential influence of fund characteristics on fund performance. The dependent variables in Panel A and B, respectively, are gross fund returns and gross risk-adjusted returns where returns are adjusted using Carhart 4-factor model with loadings to each factor estimated from 36-months rolling regressions. Standard errors are clustered by month and issuer. The main variable of interest is $SelfIndexer?$ and other explanatory variables include $Log(AUM)$, $Log(Age)$, $Log(Holdings)$, $SmartBeta?$, $Volatility$ as well as a host of month, style, and issuer fixed effects. The results are consistent with the portfolio tests. There are no significant performance differences between self- and public-index ETFs.

In Table 5, we consider the degree to which self- and public-index funds vary on volatility and Sharpe ratio. In panel A, the dependent variable is $Volatility$, which is defined as the standard deviation of daily net fund returns within a month, and in panel B, it is $Sharpe\ Ratio$, which is the monthly net fund return divided by the fund volatility. Self-index ETFs appear to have somewhat lower volatility, however estimates are economically modest (around a 2-4% difference with respect to the mean value of public-index ETFs). In a similar vein, self-index ETFs seem to have either similar or weakly lower Sharpe Ratios. Taken together with Table 4, our findings highlight that self-indexing ETFs do not deliver superior performance outcomes compared to their public-index counterparts.

Finally, in this section, we examine whether those self-index ETFs which are somewhat differentiated perform better and (or) charge higher fees. Table 6 shows that this is not the case. In columns 1-2 of Table 6, respectively, we introduce a $SelfIndexer?$ x $CosineSimilarity$ interaction term to the baseline regression on fees (Table 3) and regression on gross ETF returns. $CosineSimilarity$ is the average cosine similarity of holdings of an ETF compared to all other ETFs in the same style category in that month. There is no evidence that differentiated self-index ETFs perform any differently, but interestingly they charge differently. In column 1 where the dependent variable is fees, $CosineSimilarity$ is significantly negative, showing that public-index ETFs that are more similar to their peers charge lower fees, as one would expect. However, the coefficient estimate reverses for self-index ETFs ($SelfIndexer?$ x $CosineSimilarity$ is significantly positive). Self-index ETFs that are more similar (not differentiated) charge higher fees.[21] Overall, the evidence in this section does not support the differentiated products hypothesis.

4 Conflicts of interest hypothesis

Our analyses reveal that on average, self-index funds charge significantly higher fees without providing any obvious performance benefits in return. In this section, we examine the hypothesis that conflicts of interest can arise from the fact that some of the issuers of self-index ETFs also offer wealth management advisory services, and this may give them the incentive to favor their self-index ETFs at the expense of their clients. The

[21] In Section 4.0.1 we shed light on contributing factors behind this result.

results of this section support this hypothesis.

4.0.1 Do investment advisors' self-index ETFs charge more?

We start by analyzing whether self-index ETFs issued by issuers which are likely to suffer from conflicts of interest charge higher fees. To test this idea, we classify issuers into two groups: 'specialized fund managers' and 'investment advisors'. While the former group explicitly focuses on fund management (e.g., Global X, Defiance ETFs, Van Eck), the latter group offer a range of financial services including fund management as well as wealth management advisory services (e.g., Fidelity, BlackRock, State Street, Goldman Sachs) which arguably makes them more susceptible to the conflicts of interest concern. The advisory services part of the business can create the opportunity of marketing expensive self-index ETFs which might be otherwise difficult. There are 105 unique issuers in our sample, with 34 classified as investment advisor and 71 as specialized fund managers. A full list of these issuers is provided in Internet Appendix Table IA.6.

Column 1 of Table 7 tests the differences in fees between the two groups of issuers. We take the specification of column 1 of Table 3 and introduce the interaction term $SelfIndexer?$ x $InvAdvisor?$, where $InvAdvisor?$ is an indicator variable which equals 1 for ETFs issued by issuers classified as 'investment advisors'. The regression includes month-style fixed effects alongside all control variables. The results are striking. Investment advisors, on average, charge lower fees, which is consistent with these companies being larger and hence being able to offer more competitive prices for their products. However, their self-index ETFs systematically charge higher fees. The $SelfIndexer?$ x $InvAdvisor?$ interaction term is significantly positive. $SelfIndexer?$ alone, however, is indistinguishable from zero: among ETFs issued by specialized fund managers, there is no significant differences in fees between self- vs public-index ETFs.

In columns 2 and 3, respectively, we check whether self-indexed ETFs from investment advisors perform differently and have more differentiated portfolios. There is no such evidence. There is no significant difference in performance, and if anything, self-indexed ETFs by investment advisors are more similar to other ETFs in the same style. Altogether, the findings in this section support the conflicts of interest hypothesis.

4.0.2 Investor demand for self-index ETFs

In this section, we analyze the differences in investor flows and the characteristics of investors holding shares of public- versus self-index ETFs. If it is the case that expensive self-index ETFs are primarily used to extract rents from conflicted issuers' advisory clients, this should be reflected in fund flows and ETF holdings data. Fund flows can be expected to be lower for self-index ETFs issued by investment advisors since these ETFs will primarily attract flows from advisory clients of the issuing company. Furthermore, these self-index ETFs should appear to have disproportionately high institutional ownership by the issuing company ('institutional self-ownership') as investment advisors report client assets held in discretionary advisory accounts.

We start by examining fund flows. In Table 8, we regress next period ETF fund flows on explanatory variables including the indicator variable $SelfIndexer?$, $Log(AUM)$, $Log(Age)$, and $AvgNetRet(Yr)$ $(\%)$ which is the average fund return over the past 12 months. Regressions include month-style and issuer fixed effects. Standard errors are clustered by month and issuer dimensions.

As expected, $AvgNetRet(Yr)$ $(\%)$ is significantly positive as investors have a tendency to chase past performance. $Log(AUM)$ and $Log(Age)$ are significantly negative, capturing the fact that smaller and younger funds tend to receive more investor capital proportional to their total net assets. These results are consistent with other studies in the literature (e.g., Evans (2010), Pastor et al. (2015)). Our main variable of interest, $SelfIndexer?$, is statistically indistinguishable from zero. Next, we remove issuer fixed effects and include the interaction term, $SelfIndexer?$ x $InvAdvisor?$. ETFs issued by investment advisors on average have higher flows ($InvAdvisor?$ is positive) but their self-indexed ETFs which tend to be more expensive (in Table 7) attract significantly less investor capital. The $SelfIndexer?$ x $InvAdvisor?$ interaction is negative and statistically significant at the 1% level. This result shows that, in the aggregate, investors penalize the higher fee self-index ETFs offered by investment advisor issuers.

We next look at the characteristics of investors holding shares of public- versus self-index ETFs. Due to data limitations, we do not observe all investor holdings, but we can observe ETF holdings reported in 13F filings. Under the conflicts of interest hypothesis, investment advisors favor their self-index ETFs in the advisory part of their businesses. These assets would appear in the 13F filings of the investment advisor to the extent that these assets are purchased via (and held in) discretionary advisory accounts.[22] If this is the case, we would expect to find $Institutional\ Self\text{-}Ownership$ (fraction of total shares outstanding which are reported by the ETF issuer, or an affiliated company, in 13F filings) to be disproportionately high for self-index ETFs issued by investment advisors. Table 9 confirms that this is indeed the case. In column 1, the self-index estimate is 11%, revealing a significant difference in $Institutional\ Self\text{-}Ownership$ on average. In column 2, we include the $SelfIndexer?$ x $InvAdvisor?$ interaction and find that this difference largely stems from investment advisors. In column 2, the self-index dummy is positively significant but economically small (around 3.5%), however the estimate for $SelfIndexer?$ x $InvAdvisor?$ stands at 9.5% which is both statistically significant and economically important. The investment advisor estimate is statistically not different from zero. Overall, these results confirm our conjecture.

In summary, in this section, we document supporting evidence for the conflict of interest hypothesis. Broad investment advisors that issue self-index ETFs are prone to conflicts of interest since they additionally offer wealth management advisory services to clients.

[22]Institutional investment managers in the United States that exercise investment discretion over $100 million in Section 13(f) securities are required to file Form 13F including the assets that they own or the client assets that manage. See Section 13(f)(1)(b) of the Securities Exchange Act, and 15 U.S.C. 78c(a)(35).

The higher fees levied by self-indexing ETFs stem mainly from this type of ETF issuer. There is no fee differential between public- vs. self-index ETFs issued by the remaining issuers. Moreover, self-index ETFs attract less investor flows and have disproportionately high institutional ownership by the issuing company, arguably in part driven by client assets held in discretionary advisory accounts.

Our central analysis highlights conflicts of interest that stem from the dual roles of some ETF issuers as wealth management advisors. Some industry commentators and regulators have also raised concerns about other potential conflicts of interest which can stem from self-indexing. One of these concerns is the potential manipulation of constituent pricing by the ETF issuer to inflate their net asset values.[23] This concern is not particularly relevant to our sample, which primarily comprises funds investing in liquid US equities whose valuations are not easy to manipulate. Another posited concern is that internal production of an index can potentially incentivize funds offered by the same management company front run the company's own ETF tracking the index, due to information leaks within the company about the index reconstitution events and rebalancing dates (these trades lead to predictable price movements as evidenced by Chinco and Sammon (2024) and Li (2021)). We find this argument unfounded. First of all, it is important to note that such events of public indices are pre-announced to the market, which makes them significantly more susceptible to front running by arbitrageurs than self-indexing

[23]See, for instance, https://www.etfstream.com/articles/self-indexing-etf-issuers-must-answer-conflict-of-interest-question

ETFs. Second, for completeness, we examine if self-indexing provides unfair advantage to other funds in the same management company and find no such evidence (results available upon request).[24]

4.0.3 Model portfolio recommendations and self-index ETFs

In this section, our objective is to evaluate the idea that issuers promote their self-index ETFs using model portfolio recommendations. Model portfolios are recommendation portfolios designed by asset managers and strategists for financial advisors, and some asset management companies play a dual role by simultaneously managing funds and providing investment model recommendations. Brogaard et al. (2021) show that asset manager model providers recommend their affiliated ETFs more frequently and therefore enjoy significant flows, especially from clients who are not particularly sensitive to fees. Building on these findings, we are interested in examining whether ETF issuers which are model providers are more likely to include their self-index ETFs in their recommended model portfolios compared with their public-index ETFs. Brogaard et al. (2021) reports that more than half (54%) of advised assets are allocated in model portfolios.[25] Therefore, the data in model portfolios serve as a useful proxy for client portfolios that are otherwise

[24] For each self-indexing ETF, we first look at trading activity by all ETFs offered by the same issuer around its rebalancing date and find no trace of strategic trading. In addition, we examine the relative performance of mutual funds offered by firms issuing self-indexed ETFs. The absence of superior performance in these funds suggests a lack of significant front-running activity that would otherwise be expected to enhance their returns.

[25] https://www.broadridge.com/_assets/pdf/broadridge-fa-model-portfolio-may-2019.pdf

confidential.

For each global and US-focused model portfolio in the Morningstar Direct database, we extract quarterly portfolio holdings between 2012 and 2020. 133 unique model providers report 1125 model portfolios with at least one ETF. Within our sample, 322 out of 785 unique ETFs are at least part of one model portfolio, and 29 out of 105 unique issuers are identified as model portfolio providers. Of these 29 ETF issuers, 20 are investment advisors. This means that a much higher percentage of investment advisors provide model portfolios. Although only 12% of the specialized fund managers offer model portfolios (9 out of 71), this figure is 58% for investment advisors (20 out of 34).

We classify an ETF as affiliated to a model provider if the ETF issuer is affiliated with the model provider (as in Brogaard et al. (2021)). For example, an iShares S&P 500 ETF in the "BlackRock 50/50 Long-Horizon Allc" model portfolio counts as an affiliated ETF. Out of 29 model providing ETF issuers within our sample, 17 include their ETFs in their model portfolios.[26] Out of these 17 issuers, 11 are classified as investment advisors. In sum, 32% of investment advisors within our sample (11 out of 34) offer model portfolios which include at least one of their own or affiliated ETFs.

For each ETF quarter, we count the number of affiliated model portfolios an ETF is part of. If we cannot match an ETF to any model portfolio, then the ETF is taken to have

[26] The 17 issuers which include their affiliated ETFs in their model portfolios are Vanguard, First Trust, SSgA, WisdomTree, Goldman Sachs, Northern Trust, John Hancock, BlackRock, Charles Schwab, Fidelity, Franklin Templeton Investments, Inspire Investing, Nuveen, Global X, Invesco, Columbia, and JPMorgan.

zero affiliated model portfolios. This is the main dependent variable for the regressions in Table 10. Explanatory variables include $Log(AUM)$, $Log(Age)$, $Net\ Expense\ Ratio\ (\%)$, $AvgNetRet(Yr)\ (\%)$ and a smart-beta flag. Standard errors are double clustered (by quarter and issuer).

Issuers which are model providers tend to recommend their self-indexed ETFs more often than their public-index ETFs. (Note that for issuers that are not model providers the dependent variable equals zero for both self- and public-index ETFs.) This effect is economically large, given that the average count of affiliated model portfolio membership is around 0.6 (the estimates for $SelfIndexer?$ are 0.788 and 1.171 in columns 2-3). In the last column, we introduce $SelfIndexer?$ x $InvAdvisor?$ to examine the behavior of investment advisors. $InvAdvisor?$ is significantly positive – ETFs issued by investment advisors have more affiliation to model portfolios, consistent with most model providers being classified as investment advisors. Moreover, $SelfIndexer?$ x $InvAdvisor?$ is significantly positive, that is, self-index ETFs issued by investment advisors have more affiliation to model portfolios, compared to other self-index ETFs. To sum up, this section provides supporting evidence that model portfolio recommendations appear to be one way that these advisors promote self-index ETFs to investors.

5 Conclusion

The market for index providers is a concentrated market where the three largest providers serve about 75 percent of the market. Index providers use their market power to charge relatively high fees to ETF issuers and these costs are then passed on to end investors (An et al., 2023). With the 2013 SEC orders allowing ETFs that create their own indices in-house (self-index ETFs) to operate essentially in the same way as ETFs relying on external index providers (public-index ETFs), there has been a gradual shift in the marketplace.

We investigate whether self-indexing ultimately benefits end-investors. Our findings contradict the notion that self-indexing offers financial advantages through increased product variety, lower fees, or improved performance. Instead, our evidence support the conflicts of interest concern, that the ETF issuers which also operate as wealth management advisors can utilize self-indexed ETFs to favor their proprietary products, possibly at the expense of client interests.

References

Akey, Pat, Adriana Robertson, and Mikhail Simutin, "Closet active management of passive funds," *Rotman School of Management Working Paper*, 2021.

An, Yu, Matteo Benetton, and Yang Song, "Index providers: Whales behind the scenes of ETFs," *Journal of Financial Economics*, 2023, *149* (3), 407–433.

Anagol, Santosh and Hugh Hoikwang Kim, "The impact of shrouded fees: Evidence from a natural experiment in the Indian mutual funds market," *American Economic Review*, 2012, *102* (1), 576–593.

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim, and Rabih Moussawi, "Competition for Attention in the ETF Space," *The Review of Financial Studies*, 08 2022, *36* (3), 987–1042.

Bergstresser, Daniel, John M. R. Chalmers, and Peter Tufano, "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry," *The Review of Financial Studies*, 05 2009, *22* (10), 4129–4156.

Brogaard, Jonathan, Nataliya Gerasimova, and Ying Liu, "Advising the Advisors: Evidence from ETFs," *Available at SSRN 3972562*, 2021.

Brown, David C, Scott Cederburg, and Mitch Towner, "Dominated ETFs," *forthcoming at Critical Finance Review*, 2024.

Chalmers, John and Jonathan Reuter, "Is conflicted investment advice better than no advice?," *Journal of Financial Economics*, 2020, *138* (2), 366–387.

Cheng, Si, Massimo Massa, and Hong Zhang, "The unexpected activeness of passive investors: a worldwide analysis of ETFs," *The Review of Asset Pricing Studies*, 2019, *9* (2), 296–355.

Choi, James J, David Laibson, and Brigitte C Madrian, "Why does the law of one price fail? An experiment on index mutual funds," *The Review of Financial Studies*, 2010, *23* (4), 1405–1432.

Christoffersen, Susan EK, Richard Evans, and David K Musto, "What do consumers' fund flows maximize? Evidence from their brokers' incentives," *The Journal of Finance*, 2013, *68* (1), 201–235.

Easley, David, David Michayluk, Maureen O'Hara, and Talis J Putnins, "The active world of passive investing," *Review of Finance*, 2021, *25* (5), 1433–1471.

Edelen, Roger M., Kingsley Y.L. Fong, and Jingyi Han, "Regulating inattention in fee-based financial advice," *Journal of Financial Economics*, 2025, *164*, 103985.

_ , Richard B. Evans, and Gregory B. Kadlec, "Disclosure and agency conflict: Evidence from mutual fund commission bundling," *Journal of Financial Economics*, 2012, *103* (2), 308–326.

Egan, Mark, Gregor Matvos, and Amit Seru, "The market for financial adviser misconduct," *Journal of Political Economy*, 2019, *127* (1), 233–295.

Evans, Richard B, "Mutual fund incubation," *The Journal of Finance*, 2010, *65* (4), 1581–1611.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, "Money doctors," *The Journal of Finance*, 2015, *70* (1), 91–114.

Guercio, Diane Del and Jonathan Reuter, "Mutual fund performance and the incentive to generate alpha," *The Journal of Finance*, 2014, *69* (4), 1673–1704.

Huang, Shiyang, Yang Song, and Hong Xiang, "The Smart Beta Mirage," *Journal of Financial and Quantitative Analysis*, 2023, p. 1–32.

Khomyn, Marta, Talis Putnins, and Marius Zoican, "The value of ETF liquidity," *The Review of Financial Studies*, 2024, *37* (10), 3092–3148.

Kostovetsky, Leonard and Jerold Warner, "Investor heterogeneity and the market for fund benchmarks: Evidence from passive ETFs," *Journal of Banking & Finance*, 2025, *173*, 107412.

Linnainmaa, Juhani T, Brian T Melzer, and Alessandro Previtero, "The misguided beliefs of financial advisors," *The Journal of Finance*, 2021, *76* (2), 587–621.

Mahoney, Paul G and Adriana Z Robertson, "Advisers by another name," *Harvard Business Law Review*, 2021, *11*, 311.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, "Scale and skill in active management," *Journal of Financial Economics*, 2015, *116* (1), 23–45.

Pool, Veronika K, Clemens Sialm, and Irina Stefanescu, "It pays to set the menu: Mutual fund investment options in 401 (k) plans," *The Journal of Finance*, 2016, *71* (4), 1779–1812.

_ , _ , and _ , "Mutual fund revenue sharing in 401 (k) plans," *National Bureau of Economic Research*, 2022.

Robertson, Adriana Z, "Passive in name only: Delegated management and index investing," *Yale Journal on Regulation*, 2019, *36*, 795.

Roussanov, Nikolai, Hongxun Ruan, and Yanhao Wei, "Marketing mutual funds," *The Review of Financial Studies*, 2021, *34* (6), 3045–3094.

Xiao, Yizhou and Yan Xiong, "Designing Index Provision," *Available at SSRN 4815898*, 2024.

Figures

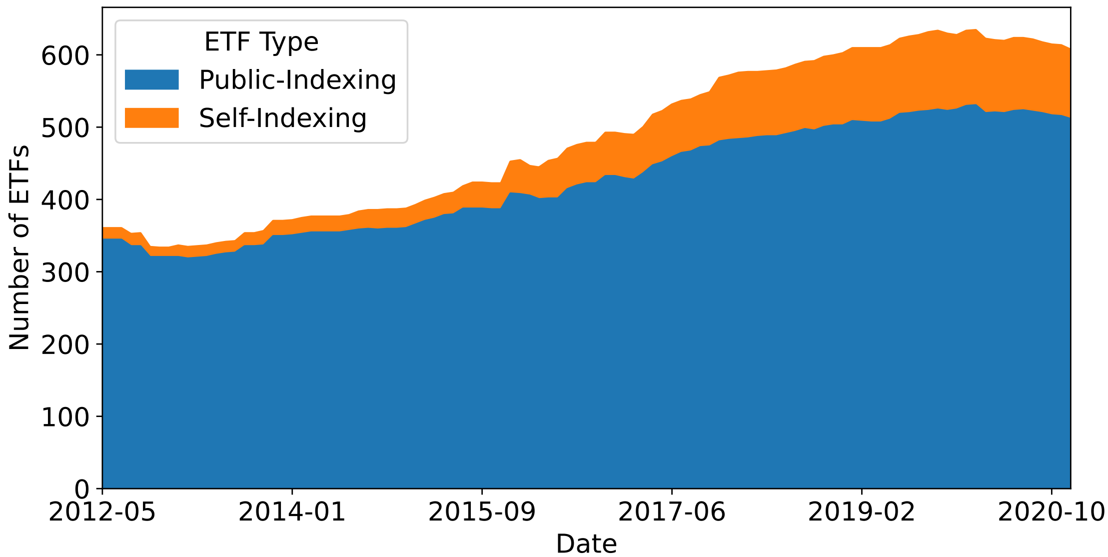

Figure 1. Number of self- and public-index ETFs over time.
This figure presents the monthly number of self-index and public-index ETFs in our sample from May 2012 until December 2020. Self-index ETFs are ETFs which track an index created and maintained by the ETF issuer themselves or one of its affiliated companies. Public-index ETFs are ETFs which track indices from unaffiliated index providers such as S&P Dow Jones or FTSE Russell.

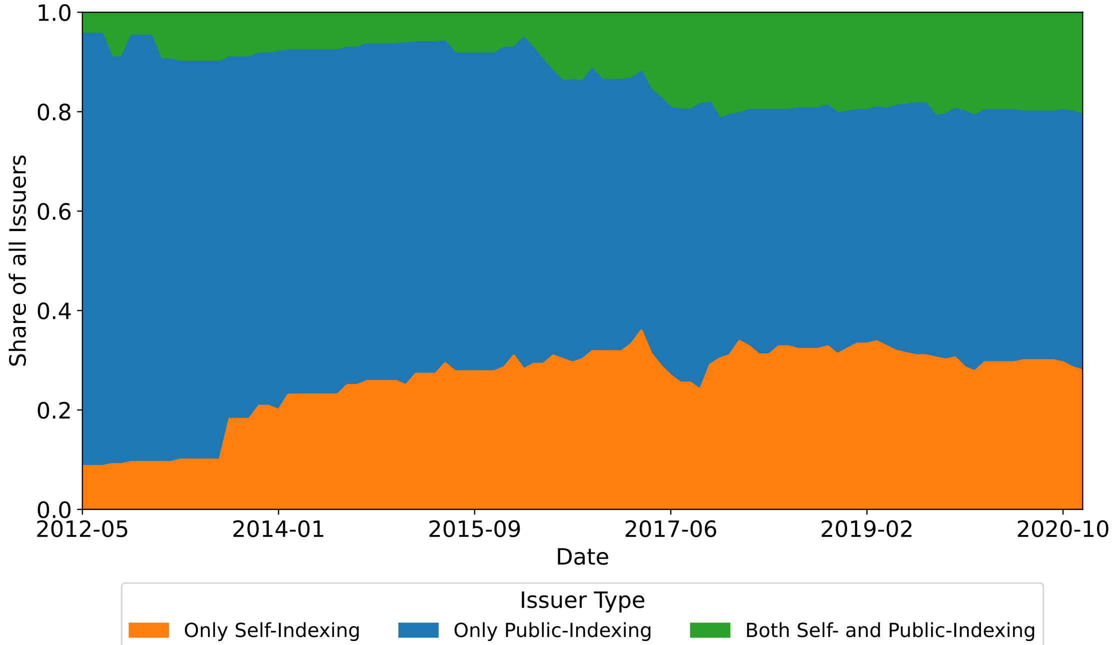

Figure 2. Share of issuer types over time.
This figure presents the monthly share of issuers in our sample from May 2012 until December 2020 which provide only self-index, only public-index, or both types of ETFs.

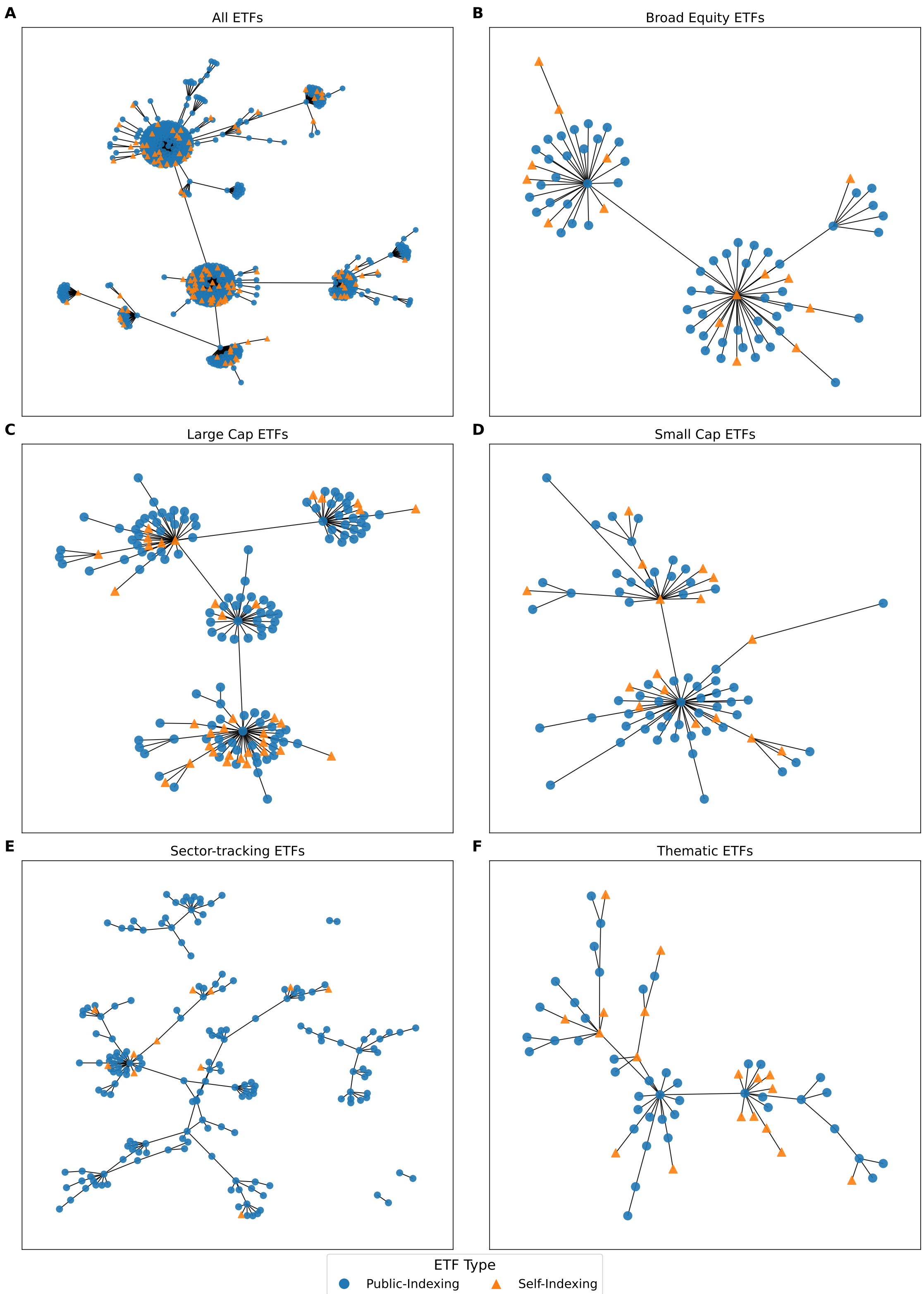

Figure 3. ETF networks based on cosine similarity of stock holdings.
This figure presents minimum spanning tree network plots of ETFs based on their monthly average pair-wise cosine similarity of stock holdings. The networks are plotted using the Fruchterman Reingold Layout. Panel A presents the network of all ETFs whereas panels B to F plot networks within ETF subgroups.

Tables

Table 1. Number of ETFs by style and indexing group.
Self-Index ETFs are ETFs which track indices provided by the ETF issuer or an affiliated company. Public-Index ETFs track indices provided by independent third parties. Styles are obtained from the focus grouping within ETF Global. For simplicity, in this table we bunch all sector-related styles into one group.

Style	Public-Indexer (N)	Public-Indexer (%)	Self-Indexer (N)	Self-Indexer (%)	Total (N)
Alpha-Seeking	21	84.0	4	16.0	25
Broad Equity	77	79.4	20	20.6	97
Buywrite	2	100.0	0	0.0	2
High Dividend Yield	27	69.2	12	30.8	39
Large Cap	165	79.3	43	20.7	208
Long/Short	12	85.7	2	14.3	14
Micro Cap	4	100.0	0	0.0	4
Mid Cap	43	87.8	6	12.2	49
Preferred Stock	9	100.0	0	0.0	9
Real Estate	1	50.0	1	50.0	2
Sector	198	92.1	17	7.9	215
Size and Style	1	100.0	0	0.0	1
Small Cap	74	79.6	19	20.4	93
Target Outcome	1	100.0	0	0.0	1
Theme	41	70.7	17	29.3	58

Table 2. Summary statistics

In Panel A we report summary statistics per ETF type. $NetExpenseRatio(\%)$ is the annual net expense ratio. $GrossReturn(\%)$ is the net (NAV) return adjusted for the net expense ratio. $Style\text{-}AdjustedReturn(\%)$ is the $GrossReturn$ minus the style-average gross return in the same time period. $Volatility$ is the standard deviation of net (NAV) returns in a month. $GrossCarhartAlpha(\%)$ is the excess return over a Carhart four-factor model prediction with a 3-year rolling window. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $FundFlow$ measures the net investment inflow in proportion to the ETFs previous AUM, adjusted for the ETFs returns. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $Avg.ReturnCorrtoStyle\text{-}Peers$ is an ETFs average return correlation compared to other ETFs in the same style category. $Inst.SelfOwnership$ is 13f ownership by the same or related investment management companies. $\#AffiliatedModels$ is the number of model portfolios by related investment management companies which report holding the ETF. Ownership variables and model portfolios are available at quarterly frequency only. Panel B reports the average cosine similarity of ETFs holdings compared to other ETFs in the same style-group, grouped by ETF type. We report the statistics for the full sample and the five largest style groups. For both panels we report the F-statistic for a comparison of means between Self-Index and Public-Index ETFs.

	Public-Indexing			Self-Indexing			
Variable	N	Mean	Std. Dev.	N	Mean	Std. Dev.	Test
Panel A: Summary Statistics of Key Variables							
Net Expense Ratio (%)	44261	0.419	0.541	6173	0.428	0.265	F=1.701
Gross Return (%)	44111	1.05	5.44	6091	1.02	5.41	F=0.155
Style-Adjusted Return (%)	44111	0.00311	2	6091	0.00187	2.3	F=0.002
Volatility	44104	0.169	0.13	6086	0.164	0.136	F=9.623***
Gross Carhart Alpha (%)	24942	-0.166	2.69	2314	-0.184	2.07	F=0.104
Log(AUM)	43876	19.4	2.42	6081	18.2	1.9	F=1509.446***
Log(Age)	44256	2	0.748	6173	1.37	0.693	F=3800.009***
Log(Holdings)	43414	4.88	1.4	6009	5.07	1.46	F=103.954***
Fund Flow	43454	0.0198	0.119	5963	0.0319	0.124	F=53.316***
Smart Beta?	38888	0.533	0.499	5381	0.775	0.418	F=1152.968***
Avg. Return Corr to Style-Peers	42797	0.702	0.218	5833	0.737	0.218	F=133.456***
Inst. Self Ownership	15036	0.0183	0.0937	2083	0.14	0.283	F=1557.11***
# Affiliated Models	15050	0.695	4.01	2104	0.874	4	F=3.682*
Panel B: Average Cosine Similarity of Holdings Compared to Style-Peers							
Full Sample	551	0.263	0.147	106	0.267	0.146	F=0.068
Large Cap	119	0.318	0.137	34	0.328	0.133	F=0.132
Mid Cap	40	0.267	0.109	5	0.288	0.037	F=0.183
Small Cap	58	0.257	0.112	15	0.231	0.0815	F=0.728
Broad Equity	61	0.249	0.14	11	0.246	0.107	F=0.004
Theme	40	0.201	0.13	16	0.139	0.0968	F=2.888*

Table 3. Explaining net expense ratios.
The dependent variable is an ETFs annual Net Expense Ratio in percentage terms. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	Net Expense Ratio (%)		
	(1)	(2)	(3)
$SelfIndexer?$	0.055***	0.052***	0.047***
	(0.006)	(0.008)	(0.009)
$Log(AUM)$		−0.023***	−0.025***
		(0.001)	(0.001)
$Log(Holdings)$		−0.012***	−0.017***
		(0.001)	(0.001)
$Log(Age)$		0.070***	0.075***
		(0.004)	(0.004)
$SmartBeta?$		0.001	0.001
		(0.010)	(0.010)
Num. Obs.	50 434	43 214	43 214
R^2	0.23	0.47	0.50
FE: Issuer		X	
FE: Month-Style	X	X	X
FE: Month-Issuer			X

* p < 0.1, ** p < 0.05, *** p < 0.01

Table 4. Alphas from portfolio tests with monthly and daily returns.
Self-Indexers represents the portfolio of all self-indexed ETFs in a period. Public-Indexers represents the portfolio of all public-indexed ETFs in a period. Gross returns are calculated by backing out net expense ratios from net (NAV) returns. We report alphas from predictions of gross returns in excess of the risk-free rate. To predict excess gross returns we use the CAPM (Panel A), the Fama-French three-factor model (Panel B) and the Carhart four-factor model (Panel C). Only observations with at least 20 ETFs per portfolio are included. We use Newey-West Standard Errors with six lags and report them in parentheses.

	Monthly Returns (%)		Daily Returns (%)	
	Gross, VW	Gross, EW	Gross, VW	Gross, EW
Panel A: CAPM				
Self-Indexers	-0.119	-0.164	0.008	-0.008
	(0.116)	(0.117)	(0.007)	(0.006)
Public-Indexers	0.015	-0.226*	0.001	-0.008
	(0.039)	(0.134)	(0.003)	(0.007)
Self-Indexers (-) Public-Indexers	-0.135	0.061	0.006	0.001
	(0.088)	(0.049)	(0.007)	(0.002)
Panel B: FF3				
Self-Indexers	0.037	-0.005	0.008	-0.001
	(0.088)	(0.064)	(0.006)	(0.003)
Public-Indexers	0.048	-0.038	0.001	-0.001
	(0.034)	(0.055)	(0.002)	(0.003)
Self-Indexers (-) Public-Indexers	-0.011	0.033	0.004	0.000
	(0.069)	(0.045)	(0.006)	(0.002)
Panel C: FF3 + Mom				
Self-Indexers	0.037	-0.005	0.007	-0.001
	(0.089)	(0.064)	(0.006)	(0.003)
Public-Indexers	0.049	-0.038	0.001	-0.002
	(0.034)	(0.055)	(0.002)	(0.003)
Self-Indexers (-) Public-Indexers	-0.012	0.033	0.004	0.001
	(0.069)	(0.045)	(0.006)	(0.002)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Explaining other ETF characteristics.
In Panel A we predict ETF volatility, defined as the standard deviation of daily net (NAV) returns within a month. In Panel B we predict Sharpe Ratios, defined as the monthly net (NAV) returns divided by the volatility. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	(1)	(2)	(3)
Panel A: Volatility			
$SelfIndexer?$	-0.005***	-0.007***	-0.008***
	(0.001)	(0.001)	(0.001)
$Log(AUM)$		-0.003***	-0.003***
		(0.000)	(0.000)
$Log(Holdings)$		-0.002***	-0.003***
		(0.000)	(0.000)
$Log(Age)$		0.007***	0.006***
		(0.001)	(0.001)
$SmartBeta?$		-0.006***	-0.006***
		(0.001)	(0.001)
Num. Obs.	50190	43074	43074
R^2	0.90	0.92	0.93
Panel B: Sharpe Ratio			
$SelfIndexer?$	-3.266	-7.239*	-9.197*
	(2.386)	(3.989)	(5.490)
$Log(AUM)$		-0.455	-0.589
		(0.480)	(0.511)
$Log(Holdings)$		2.180	3.031
		(2.040)	(2.724)
$Log(Age)$		0.068	0.420
		(1.136)	(1.126)
$SmartBeta?$		1.166	1.416
		(1.526)	(1.707)
Num. Obs.	50189	43073	43073
R^2	0.14	0.13	0.18
FE: Issuer		X	
FE: Month-Style	X	X	X
FE: Month-Issuer			X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6. Holdings similarity, fees and performance.
In column (1) we predict annual Net Expense Ratios in percentage terms and in column (2) we predict gross returns, defined as net (NAV) returns adjusted for net expense ratios. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $CosineSimilarity$ is the average cosine similarity of holdings of an ETF compared to all other ETFs in the same style category. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $Volatility$ is the standard deviation of daily net (NAV) returns within a month. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	Net Expense Ratio (%)	Gross Return (%)
	(1)	(2)
$SelfIndexer?$	−0.001	0.120
	(0.010)	(0.195)
$CosineSimilarity$	−0.432***	0.590**
	(0.037)	(0.292)
$Log(AUM)$	−0.025***	−0.012
	(0.001)	(0.012)
$Log(Holdings)$	−0.016***	−0.028
	(0.002)	(0.023)
$Log(Age)$	0.069***	0.018
	(0.005)	(0.068)
$SmartBeta?$	−0.058***	−0.060
	(0.012)	(0.037)
$SelfIndexer?{:}CosineSimilarity$	0.305***	−0.770
	(0.030)	(0.699)
$Volatility$		1.450*
		(0.759)
Num. Obs.	29 902	18 663
R^2	0.51	0.55
FE: Issuer	X	X
FE: Month-Style	X	X

* p < 0.1, ** p < 0.05, *** p < 0.01

Table 7. Exploring the role of investment advisors.

In column (1) we predict annual Net Expense Ratios in percentage terms. In column (2) we predict Gross Returns, defined as net (NAV) returns adjusted for net expense ratios. In column (3) we predict the average cosine similarity of holdings of an ETF compared to all other ETFs in the same style category. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $InvAdvisor?$ is a dummy variable which is one if the ETF issuer is classified as investment advisor as opposed to a fund specialist. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $Volatility$ is the standard deviation of daily net (NAV) returns within a month. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	Net Expense Ratio (%)	Gross Return (%)	Cosine Similarity
	(1)	(2)	(3)
$SelfIndexer?$	0.002	0.136	−0.020***
	(0.013)	(0.086)	(0.002)
$InvAdvisor?$	−0.320***	−0.006	0.015***
	(0.011)	(0.042)	(0.002)
$Log(AUM)$	−0.024***	0.018*	0.004***
	(0.001)	(0.009)	(0.000)
$Log(Holdings)$	−0.045***	0.001	0.032***
	(0.003)	(0.014)	(0.001)
$Log(Age)$	0.061***	−0.022	−0.010***
	(0.004)	(0.025)	(0.001)
$SmartBeta?$	0.049***	−0.050**	−0.011***
	(0.008)	(0.025)	(0.001)
$SelfIndexer?{:}InvAdvisor?$	0.090***	−0.120	0.013***
	(0.016)	(0.093)	(0.003)
$Volatility$		1.141	
		(0.735)	
Num. Obs.	43 214	43 074	29 902
R^2	0.21	0.88	0.70
FE: Month-Style	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8. Explaining fund flows.
The dependent variable is next-period Fund Flow, defined as the net investment inflow in proportion to the ETFs previous AUM, adjusted for returns. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $AvgNetRet(Yr)(\%)$ is the 12-month rolling average net (NAV) ETF return. $InvAdvisor?$ is a dummy variable which is one if the ETF issuer is classified as investment advisor as opposed to a fund specialist. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	Fund Flows (t+1)		
	(1)	(2)	(3)
$SelfIndexer?$	−0.003	−0.002	0.011**
	(0.005)	(0.004)	(0.005)
$Log(AUM)$	−0.001	−0.003***	−0.002***
	(0.001)	(0.001)	(0.000)
$Log(Age)$	−0.030***	−0.018***	−0.017***
	(0.002)	(0.002)	(0.002)
$SmartBeta?$	0.001	0.002	0.000
	(0.002)	(0.002)	(0.002)
$AvgNetRet(Yr)(\%)$		0.017***	0.017***
		(0.001)	(0.001)
$InvAdvisor?$			0.005**
			(0.002)
$SelfIndexer?$:$InvAdvisor?$			−0.018***
			(0.006)
Num. Obs.	43 096	36 145	36 145
R^2	0.16	0.18	0.17
FE: Issuer	X	X	
FE: Month-Style	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9. Explaining institutional self ownership. The dependent variable is Institutional Self Ownership, defined as 13f ownership by the same or related investment management companies. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $AvgNetRet(Yr)(\%)$ is the 12-month rolling average net (NAV) ETF return. $InvAdvisor?$ is a dummy variable which is one if the ETF issuer is classified as investment advisor as opposed to a fund specialist. Standard errors are two-way clustered by quarter-issuer interaction and reported in parentheses.

	Institutional Self Ownership	
	(1)	(2)
$SelfIndexer?$	0.108***	0.035**
	(0.017)	(0.015)
$Log(AUM)$	0.008***	0.008***
	(0.001)	(0.001)
$Log(Holdings)$	0.011***	0.009***
	(0.003)	(0.002)
$Log(Age)$	−0.052***	−0.053***
	(0.006)	(0.006)
$SmartBeta?$	0.033***	0.029***
	(0.005)	(0.005)
$AvgNetRet(Yr)(\%)$	−0.001	−0.001
	(0.002)	(0.002)
$InvAdvisor?$		0.000
		(0.004)
$SelfIndexer?{:}InvAdvisor?$		0.095***
		(0.025)
Num. Obs.	12 307	12 307
R^2	0.18	0.19
FE: Quarter-Style	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10. Explaining affiliated model portfolio membership. The dependent variable is the number of affiliated model portfolios an ETF is a part of. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $NetExp.Ratio(\%)$ is the annual net expense ratio. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $AvgNetRet(Yr)(\%)$ is the 12-month rolling average net (NAV) ETF return. $InvAdvisor?$ is a dummy variable which is one if the ETF issuer is classified as investment advisor as opposed to a fund specialist. Standard errors are two-way clustered by quarter-issuer interaction and reported in parentheses.

	# Affiliated Model Portfolios			
	(1)	(2)	(3)	(4)
$SelfIndexer?$	−0.083	0.788*	1.171**	−1.190***
	(0.184)	(0.454)	(0.513)	(0.164)
$NetExp.Ratio(\%)$	−0.018	−0.021	−0.029	0.088***
	(0.032)	(0.042)	(0.041)	(0.029)
$Log(AUM)$	0.540***	0.737***	0.769***	0.530***
	(0.063)	(0.091)	(0.094)	(0.062)
$Log(Age)$	−0.788***	−2.036***	−2.156***	−0.853***
	(0.175)	(0.315)	(0.330)	(0.186)
$SmartBeta?$	0.367***	−0.442***	−0.425***	0.336***
	(0.080)	(0.089)	(0.090)	(0.082)
$AvgNetRet(Yr)(\%)$	−0.091	−0.084	−0.043	−0.088
	(0.082)	(0.080)	(0.085)	(0.082)
$InvAdvisor?$				0.339***
				(0.130)
$SelfIndexer?{:}InvAdvisor?$				1.351***
				(0.258)
Num. Obs.	12 385	12 385	12 385	12 385
R^2	0.10	0.16	0.19	0.11
FE: Issuer		X		
FE: Quarter-Style	X	X	X	X
FE: Quarter-Issuer			X	

* p < 0.1, ** p < 0.05, *** p < 0.01

Appendix A: Variable Definitions

Table A.1. Variable Definitions

Variable	Definition
$SelfIndexer?$	Dummy variable that equals one if an ETF is self-indexed. Self-indexed ETFs are ETFs which track an index created and maintained by the ETF issuer themselves. ETF indices obtained from 'primary_benchmark' column in ETF Global.
$NetExp.Ratio(\%)$	Yearly net expense ratio after fee waivers from ETF Global.
$Log(AUM)$	Natural logarithm of an ETF's assets under management from ETF Global. If this information is missing in ETF Global, we use data from the CRSP MF database.
$Log(Age)$	Natural logarithm of fund age, where age of an ETF is defined as the difference between the month-end date and the inception date, divided by 365. The inception date is from ETF Global.
$Log(Holdings)$	Natural logarithm of the unique number of held assets by an ETF, obtained from ETF Global.
$FundFlow$	Increase in total net assets of an ETF not attributable to asset performance, as share of previous period's total net assets (see equation 1).
$Volatility$	Standard deviation of daily net (NAV) ETF returns within a calendar month. Daily NAV returns are obtained from the CRSP daily stock file.
$GrossReturn(\%)$	We add $NetExp.Ratio(\%)/252$ to daily net (NAV) returns. Monthly returns are then calculated as the cumulative product of one plus each day's return over a full month, minus one. Daily NAVs are from the CRSP daily stock file.
$Style\text{-}AdjustedReturn(\%)$	$GrossReturn(\%)$ minus the style-average $GrossReturn(\%)$ in the same month. Style categories are obtained from the 'focus' columns in ETF Global.
$AvgNetRet(Yr)(\%)$	12-month rolling average net (NAV) ETF return, calculated using NAVs from the CRSP MF database.
$GrossCarhartAlpha(\%)$	ETF's excess $GrossReturn(\%)$ estimated over a 36-month rolling window using the 4-factor Carhart model.
$SharpeRatio$	Monthly net (NAV) ETF return divided by the $Volatility$, using NAVs from the CRSP MF database.
$Avg.RetCorrtoStyle\text{-}Peers$	Average of all pairwise daily net (NAV) return correlations of a given ETF with all other style-peers in each month. Daily NAV returns are obtained from the CRSP daily stock file and style refers to the column 'focus' in ETF Global.
$CosineSimilarity$	Monthly average cosine similarity of ETF holdings compared to all other ETFs in the same style. Data is obtained from ETF Global where style refers to column 'focus'.
$Inst.SelfOwnership$	Fraction of ETF shares held only by the ETF issuing company. Quarterly holdings are from Thomson/Refinitiv's S34 file.
$SmartBeta?$	Dummy variable that is one if an ETF is classified as smart beta. The smart beta classification is taken from Morningstar Direct.
$Uniqueness$	12-month average of an ETFs absolute difference in $GrossReturn(\%)$ versus the value-weighted average $GrossReturn(\%)$ of all ETFs within the same style category. Style categories are obtained from the 'focus' columns in ETF Global.
$TurnoverRatio$	Minimum of aggregate sales or purchases of securities by the ETF, divided by the 12-month average TNA. Obtained from the CRSP MF database.
$MavgSpread$	Monthly average difference between the daily ask and bid, scaled by the price. All variables obtained from the CRSP daily stock file.
$InvestmentAdvisor?$	Dummy variable which is one if the ETF issuer is manually classified as investment advisor as opposed to a fund specialist.
$IssuerLog(AUM)$	Natural logarithm of all assets under management of an issuer across all asset classes in the CRSP MF database.
$IssuerLog(Age)$	Natural logarithm of the number of days divided by 365 since the first observation of an issuer in the CRSP MF database.
$IssuerNrETFs$	Number of passive equity ETFs operated by an issuer in the CRSP MF database.
$IssuerNrSelfIndexETFs$	Number of passive self-index ETFs operated by an issuer. Self-index ETFs are defined using data from ETF Global (see above).

Internet Appendix to

Index Disruption: The Promise and Pitfalls of Self-Indexed ETFs

(Not for Publication)

Figures

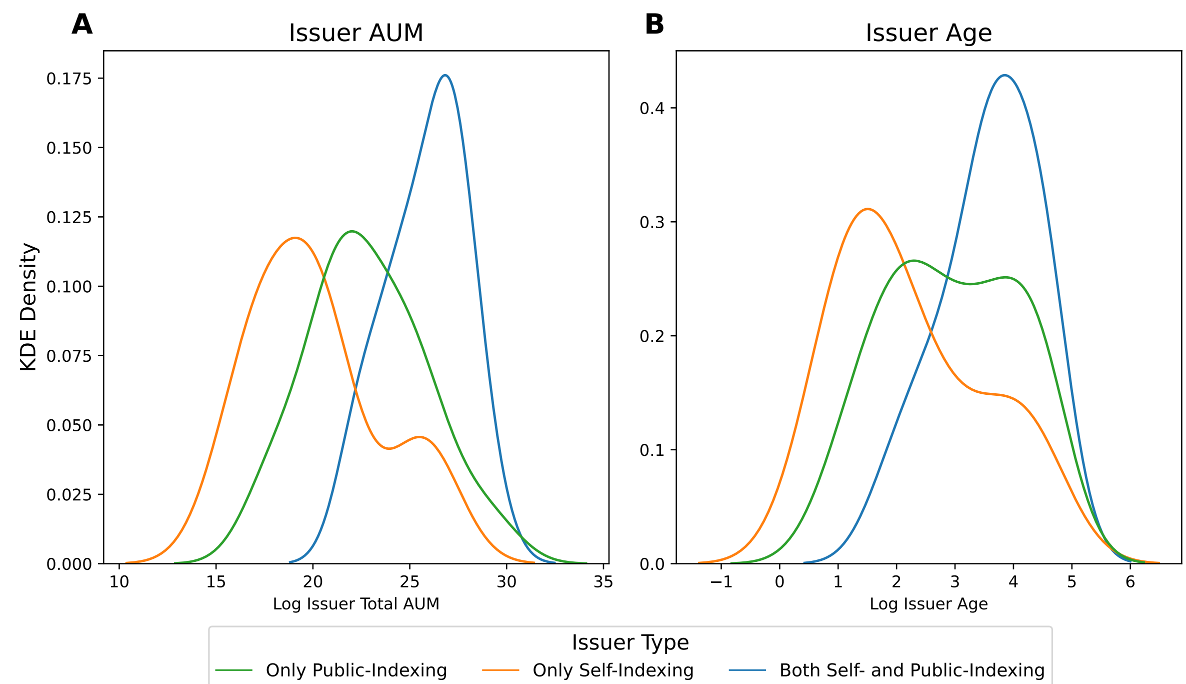

Figure IA.1. Density plots of issuer characteristics by issuer type.
This figure presents KDE density plots for key issuer characteristics, grouped by issuer type, as of December 2020. Panel A plots the densities for the logarithm of total issuer AUM, calculated as the total AUM of all outstanding funds by an issuer (across all asset-classes) from the CRSP MF database. Panel B plots the densities for the logarithm of issuer age, defined as the difference in days between December 31st 2020 and the first observation of the issuer in the CRSP MF database, divided by 365.

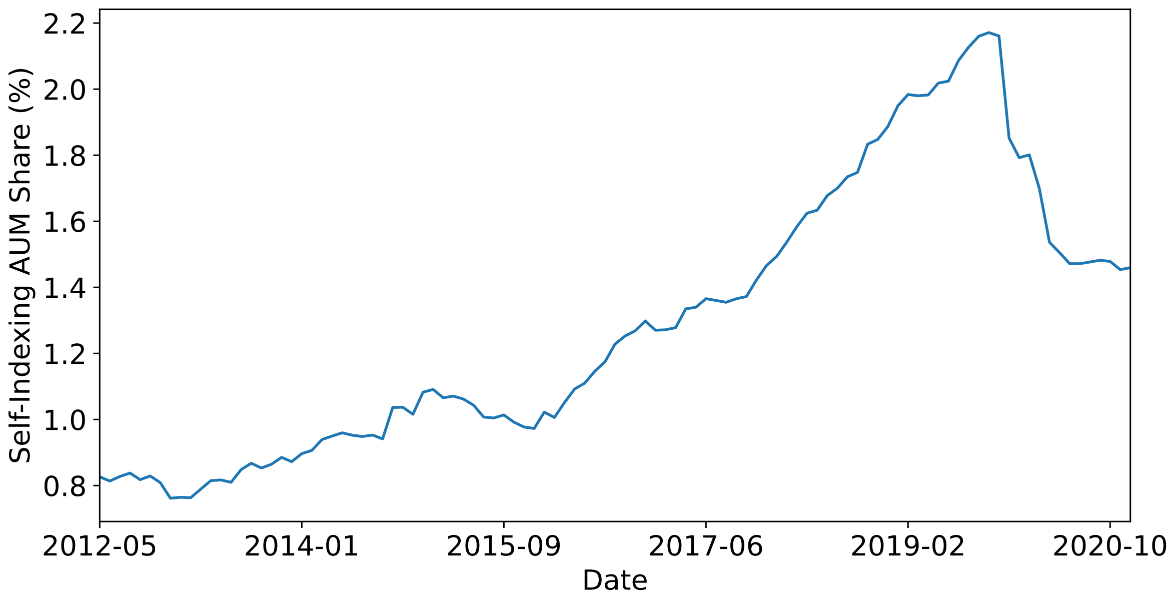

Figure IA.2. Share of self-index ETF AUM.
This figure presents the monthly share of assets managed (AUM) in self-index ETFs as a fraction of total ETF AUM from May 2012 until December 2020. The full sample includes all passive unlevered equity ETFs in the ETF Global Database.

Tables

Table IA.1. Largest issuers of ETFs by AUM as of December 2020.
AUM and the number of unique ETFs are from our sample of non-active equity ETFs in ETF Global. Mixing issuers are issuers which offer both self- and public-index ETFs.

	Self-Indexing		Public-Indexing	
Issuer	AUM ($Mn.)	Unique ETFs	AUM ($Mn.)	Unique ETFs
Panel A: Only Self-Indexing Issuers				
John Hancock	854.76	11	-	-
Renaissance	734.30	1	-	-
American Century	384.50	2	-	-
Motley Fool Asset Management	367.76	1	-	-
Tortoise	360.65	1	-	-
Hartford	269.96	2	-	-
Inspire Investing	237.59	2	-	-
Janus Henderson	236.10	2	-	-
Alpha Architect	218.68	2	-	-
Distillate Capital Partners	201.09	1	-	-
Panel B: Mixing Issuers				
WisdomTree	15588.73	13	1006.86	1
Goldman Sachs	11400.19	3	621.38	2
Pacer Financial	2722.28	8	871.15	4
Victory Capital Management	2365.95	7	448.53	2
Northern Trust	2080.11	5	1452.45	2
Charles Schwab	1320.58	1	107147.66	10
Fidelity	980.20	6	17625.00	11
SSgA	976.39	3	633741.90	59
JPMorgan	737.89	5	1280.29	4
ProShares	278.79	2	9032.34	13
Panel C: Only Public-Indexing Issuers				
Blackrock	-	-	961544.19	93
Vanguard	-	-	866575.55	43
First Trust	-	-	55674.09	54
Global X	-	-	6132.46	15
Alps	-	-	5936.66	6
Van Eck	-	-	4300.81	4
DWS	-	-	3444.60	4
Principal	-	-	2858.50	5
Nuveen	-	-	1904.51	6
Franklin Templeton Investments	-	-	1520.72	3

Table IA.2. Explaining the issuance of self-index ETFs at the issuer-month level.
The dependent variable is a dummy which equals one if the issuer increases the amount of outstanding self-index ETFs. In columns (1) - (3) we use an OLS estimation and in columns (4) - (6) we use probit estimation. $IssuerLog(Age)$ is the natural logarithm of the number of days divided by 365 since the first observation of an issuer in the CRSP MF database. $IssuerLog(AUM)$ is the natural logarithm of all assets under management of an issuer across all asset classes in the CRSP MF database. $IssuerNrETFs$ is the number of non-active equity ETFs operated by an issuer and $IssuerNrSelfIndexETFs$ is the number of self-index ETFs operated by an issuer in our sample from ETF global. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	OLS			Logit		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.042			−1.918		
	(0.051)			(2.457)		
$IssuerLog(Age)$	0.004	0.004	0.004	0.312	0.237	−1.749
	(0.007)	(0.006)	(0.011)	(0.357)	(0.343)	(2.538)
$IssuerLog(AUM)$	−0.002	−0.002	−0.012*	−0.197	−0.157	−1.020
	(0.003)	(0.003)	(0.007)	(0.155)	(0.116)	(1.068)
$IssuerNrETFs$	0.000	0.000	0.001	0.020	0.019	0.047
	(0.000)	(0.000)	(0.001)	(0.014)	(0.014)	(0.042)
$IssuerNrSelfIndexETFs$	0.007***	0.007***	0.010	0.264***	0.266***	0.760
	(0.002)	(0.002)	(0.006)	(0.052)	(0.048)	(0.757)
Num. Obs.	1521	1521	1521	1521	705	286
R^2	0.03	0.05	0.12	0.13	0.17	0.16
FE: Month		X			X	
FE: Issuer			X			X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table IA.3. Explaining net expense ratios with extended controls.
The dependent variable is an ETFs annual Net Expense Ratio. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $Uniqueness$ is the 12-month rolling average absolute difference of ETF gross returns versus the value-weighted average ETF gross return in the same style-category, as in Kostovetsky and Warner (2025). $TurnRatio$ is an ETFs yearly portfolio turnover ratio. $MAvgSpread$ is the monthly average ETF secondary market price spread scaled by price. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	Net Expense Ratio (%)			
	(1)	(2)	(3)	(4)
$SelfIndexer?$	0.057***	0.050***	0.043***	0.031***
	(0.011)	(0.011)	(0.011)	(0.011)
$Log(AUM)$	−0.024***	−0.025***	−0.022***	−0.023***
	(0.002)	(0.002)	(0.002)	(0.002)
$Log(Holdings)$	0.000	0.001	0.001	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
$Log(Age)$	0.075***	0.079***	0.068***	0.074***
	(0.004)	(0.004)	(0.004)	(0.004)
$SmartBeta?$	0.001	0.004	−0.004	−0.004
	(0.012)	(0.012)	(0.013)	(0.013)
$Uniqueness$	0.056***	0.060***	0.055***	0.059***
	(0.004)	(0.005)	(0.005)	(0.005)
$TurnRatio$			0.032***	0.041***
			(0.007)	(0.008)
$MAvgSpread$			−0.008	−0.008
			(0.006)	(0.006)
Num. Obs.	35 644	35 644	35 000	35 000
R^2	0.50	0.53	0.51	0.53
FE: Issuer	X		X	
FE: Month-Style	X	X	X	X
FE: Month-Issuer		X		X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table IA.4. Explaining net expense ratios with community fixed effects.
The dependent variable is an ETFs annual Net Expense Ratio in percentage terms. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $Uniqueness$ is the 12-month rolling average absolute difference of ETF gross returns versus the value-weighted average ETF gross return in the same style-category, as in Kostovetsky and Warner (2025). $TurnRatio$ is an ETFs yearly portfolio turnover ratio. $MAvgSpread$ is the monthly average ETF secondary market price spread scaled by price. Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	Net Expense Ratio (%)			
	(1)	(2)	(3)	(4)
$SelfIndexer?$	0.069***	0.036***	0.033***	0.035***
	(0.007)	(0.006)	(0.007)	(0.009)
$Log(AUM)$		−0.018***	−0.019***	−0.017***
		(0.001)	(0.001)	(0.001)
$Log(Holdings)$		−0.026***	−0.030***	−0.020***
		(0.001)	(0.001)	(0.001)
$Log(Age)$		0.063***	0.065***	0.062***
		(0.003)	(0.003)	(0.003)
$SmartBeta?$		0.024***	0.027***	0.037***
		(0.005)	(0.005)	(0.006)
$Uniqueness$				0.023***
				(0.001)
Num. Obs.	46 933	41 486	41 486	34 480
R^2	0.77	0.91	0.91	0.92
FE: Issuer		X		X
FE: Month-Issuer			X	
FE: Month-Community	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table IA.5. Monthly performance regressions.
In Panel A we predict the Gross Return of ETFs, calculated by backing out net expense ratios from net (NAV) returns. In Panel B we predict Gross Carhart four-factor excess returns, defined as the gross returns minus predicted returns using factor loadings from a 36 month rolling windows. $SelfIndexer?$ is a dummy variable that equals one if an ETF is self-indexed. $Log(AUM)$ is the natural logarithm of an ETFs assets under management. $Log(Holdings)$ is the natural logarithm of an ETFs number of unique portfolio holdings. $Log(Age)$ is the natural logarithm of an ETFs age as measured in days since inception divided by 365. $SmartBeta?$ is a dummy variable that is one if an ETF is classified as smart beta. $Volatility$ is the standard deviation of daily net (NAV) returns within a month. $Uniqueness$ is the 12-month rolling average absolute difference of ETF gross returns versus the value-weighted average ETF gross return in the same style-category, as in Kostovetsky and Warner (2025). Standard errors are two-way clustered by month-issuer interaction and reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Gross Returns (%)						
$SelfIndexer?$	0.047	-0.085	-0.089	0.082*	-0.144	-0.120
	(0.038)	(0.061)	(0.066)	(0.048)	(0.114)	(0.113)
$Log(AUM)$	0.018**	0.012	0.007	0.020**	0.008	0.004
	(0.009)	(0.009)	(0.008)	(0.010)	(0.009)	(0.009)
$Log(Holdings)$	-0.001	0.009	0.010	0.002	0.010	0.014
	(0.014)	(0.015)	(0.019)	(0.016)	(0.017)	(0.021)
$Log(Age)$	-0.023	0.005	0.005	-0.032	0.022	0.023
	(0.025)	(0.027)	(0.027)	(0.029)	(0.031)	(0.032)
$SmartBeta?$	-0.054**	-0.049*	-0.054*	-0.034	-0.014	-0.015
	(0.025)	(0.029)	(0.029)	(0.028)	(0.034)	(0.034)
$Volatility$	1.152	1.174	0.698	0.600	0.669	0.365
	(0.734)	(0.731)	(0.788)	(0.691)	(0.702)	(0.758)
$Uniqueness$				0.041	0.044	0.046
				(0.048)	(0.050)	(0.052)
Num. Obs.	43074	43074	43074	35641	35641	35641
R^2	0.88	0.88	0.90	0.87	0.87	0.90
Panel B: Gross Carhart Four-Factor Excess Returns (%)						
$SelfIndexer?$	0.032	-0.045	-0.042	0.048	-0.142	-0.153
	(0.055)	(0.088)	(0.099)	(0.064)	(0.112)	(0.137)
$Log(AUM)$	-0.005	-0.012	-0.013	-0.008	-0.017	-0.018
	(0.009)	(0.010)	(0.011)	(0.013)	(0.013)	(0.014)
$Log(Holdings)$	-0.013	-0.020	-0.022	-0.034	-0.037	-0.031
	(0.019)	(0.020)	(0.021)	(0.023)	(0.023)	(0.024)
$Log(Age)$	0.009	0.002	0.000	-0.006	-0.035	-0.027
	(0.042)	(0.054)	(0.057)	(0.048)	(0.062)	(0.068)
$SmartBeta?$	-0.086**	-0.088**	-0.096***	-0.090**	-0.092**	-0.105**
	(0.034)	(0.034)	(0.035)	(0.039)	(0.041)	(0.043)
$Volatility$	-0.351	-0.063	-0.544	-0.387	-0.071	-0.718
	(0.643)	(0.653)	(0.675)	(0.742)	(0.760)	(0.755)
$Uniqueness$				-0.019	-0.029	-0.017
				(0.046)	(0.051)	(0.050)
Num. Obs.	23929	23929	23929	20162	20162	20162
R^2	0.54	0.54	0.60	0.55	0.55	0.62
FE: Issuer		X			X	
FE: Month-Style	X	X	X	X	X	X
FE: Month-Issuer			X			X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table IA.6. Sample of issuers classified as investment advisors and fund specialists.
We classify issuers as 'Investment Advisors' if our manual search shows they offer a range of financial services including fund management and (wealth) advisory services. 'Fund Specialists' are issuers which focus explicitly on fund management.

Investment Advisors	Fund Specialists	
Advisors Asset Management	AGFiQ	LocalShares
Blackrock	Alpha Architect	M-CAM International
Calamos	AlphaClone	Metaurus Advisors
Cboe Vest Financial	Alps	Motley Fool Asset Management
Charles Schwab	American Century	Nuveen
Deutsche Bank	Amplify	O'Shares
Diamond Hill	Aptus Capital Advisors	Oppenheimer
DWS	Beyond Investing	Pacer Financial
Fidelity	BioShares	Pax World
Goldman Sachs	BNY Mellon Investment Management	ProShares
Guggenheim	Cambria	QuantX Funds
Hartford	Change Finance	Reality Shares
Invesco	Columbia	Recon Capital
Janus Henderson	Compass	Renaissance
John Hancock	Cushing	RevenueShares
JPMorgan	Defiance ETFs	Salt Financial
Legg Mason	Direxion	SerenityShares
Nationwide	Distillate Capital Partners	SL Advisors
Northern Trust	Elkhorn	SP Funds
PIMCO	EntrepreneurShares	Sprott
Point Bridge Capital	ETF Managers Group	Syntax
Principal	ETF Securities	Timothy Plan
Redwood	Exchange Traded Concepts	TriLine Index Solutions
Royal Bank of Scotland	Exponential ETFs	TWM Funds
Russell	Falah Capital	US Commodity Funds
Scottrade	FFCM	USAA
SoFi	First Trust	ValueShares
SSgA	Franklin Templeton Investments	Van Eck
Tortoise	Global Beta Advisors	VelocityShares
Vanguard	Global X	Vident Financial
Victory Capital Management	GraniteShares	Virtus
Wahed Invest	Horizons	WBI
WisdomTree	Hoya Capital	-
-	Impact Shares	-
-	IndexIQ	-
-	Innovation Shares	-
-	Innovator Capital Management	-
-	Inspire Investing	-
-	KraneShares	-
-	Lattice	-