# Index Providers: Whales Behind the Scenes of ETFs*

Yu An † Matteo Benetton ‡ Yang Song §

July 30, 2021

**Abstract**

Most ETFs passively replicate the performance of an index that is constructed and maintained by an index provider. We show that index providers wield strong market power and charge large markups to ETFs, which are passed on to investors through management fees. We document three stylized facts about index providers: (*i*) the ETF indexing market is highly concentrated; (*ii*) when choosing ETFs, investors care about the identities of index providers, although index providers explain little variation in ETF returns; and (*iii*) about one-third of all ETF management fees are paid as index licensing fees to index providers. Using a structural model that incorporates two-tiered competition between index providers for ETFs and between ETFs for investors, we estimate that 60% of licensing fees are markups charged by index providers. Eliminating index providers' market power can reduce ETF management fees by 30%.

*We thank Zahi Ben-David, Philip Bond, Jarrad Harford, Avi Kamara, Nagpurnanand Prabhala, and Zhaogang Song, as well as seminar participants at Johns Hopkins University and University of Washington for helpful comments. We thank Haoliang Jiang and Zheng (Jacky) Zhang for excellent research assistance.

†Carey Business School, Johns Hopkins University; yua@jhu.edu.

‡Haas School of Business, University of California at Berkeley; benetton@berkeley.edu.

§Foster School of Business, University of Washington; songy18@uw.edu.

## 1 Introduction

Exchange-traded funds (ETFs) have experienced remarkable growth in recent years. According to the 2020 Investment Company Institute (ICI) Fact Book, total assets under management (AUM) in ETFs have increased from $992 billion in 2010 to $4.4 trillion by the end of 2019. By design, the vast majority of ETFs passively replicate the performance of an underlying index, which in most cases is constructed and maintained by a designated index provider.[1] As S&P Dow Jones, the world's largest index provider, writes on its website, "an index provider is a specialized firm that is dedicated to creating and calculating market indices and licensing its intellectual capital as the basis of passive products."[2] Thus, most ETFs exhibit a two-tier organizational structure: (*i*) an index provider builds and maintains the index that underlies an ETF and charges index licensing fees to an ETF sponsor, and (*ii*) the ETF sponsor services ETF investors and charges management fees to ETF investors.

Take the largest ETF in the world, the SPDR S&P 500 ETF (SPY), as an example. In this case, the ETF sponsor is State Street (SPDR), and the index provider is S&P Dow Jones, which owns the underlying ETF index —the S&P 500 index. State Street charges SPY investors 9 basis points (bps) per year, and in turn, pays 3 bps of the ETF assets plus a flat fee of $600,000 per year to S&P Dow Jones. In other words, more than one-third of SPY's total revenue is paid to the index provider as index licensing fees.[3] For another well-known ETF, the Invesco QQQ Trust, 9 bps out of the 20 bps management fees that the ETF sponsor (Invesco) charges to ETF investors are paid in the form of licensing fees to the index provider (NASDAQ), who owns the underlying NASDAQ-100 Index.

Even though index providers play an indispensable role in the ETF marketplace and capture a substantial fraction of the total ETF business revenue, the competitive landscape

[1]More recently, ETF sponsors started offering so-called "actively-managed" ETFs, which do not passively track indexes (Akey, Robertson, and Simutin, 2021).

[2]See https://www.spglobal.com/spdji/en/index-literacy/who-s-behind-the-index/.

[3]For example, in 2019 the SPY AUM totaled about $300 billion, implying that the total management fee collected by State Street from SPY is roughly $270 million, with more than $90 million paid to S&P Dow Jones in index licensing fees.

between ETF sponsors and index providers or how their interaction influences ETF investors have not been studied so far. Our paper takes on this task through both reduced-form analysis and structural modeling.

We document that the index provider market is highly concentrated and dominated by a few large players. Moreover, when choosing ETFs, investors care about the identities of index providers, although index providers' identities explain little of the variation in ETF returns. We estimate that about one-third of all ETF management fees are paid to index providers in the form of licensing fees. Our structural estimation reveals that 60% of the index licensing fees charged by index providers to ETF sponsors are markups, and the remaining 40% of the index licensing fees reflect the marginal costs of index provision. Overall, our findings show that index providers wield strong market power, and their high indexing licensing fees are passed onto ETF investors through management fees.[4] Through a counterfactual analysis, we estimate that eliminating index providers' market power can reduce ETF management fees by about 30%.

Our paper is structured in two parts. In the first part of the paper, we establish three stylized facts about index providers in the U.S. equity ETF market over a 10-year period from 2010 through 2019. First, the ETF indexing business is highly concentrated among a few large index providers. For example, about 53% of all ETF assets in our sample track the indexes built by S&P Dow Jones. The five largest index providers in the equity ETF market, S&P Dow Jones, CRSP, FTSE Russell, MSCI, and NASDAQ, capture in aggregate about 95% of the entire ETF market. Specifically, over our sample period, the time-series average of the Herfindahl-Hirschman index (HHI) of the index provider industry is 3,294, which is deemed highly concentrated according to the U.S. Department of Justice and the Federal Trade Commission.[5]

[4]There is ample evidence of an increased role of market power in the U.S. economy; see Philippon (2019) for a full treatment of this concern.

[5]Markets are classified as unconcentrated if the HHI is below 1500, as moderately concentrated if the HHI is between 1500 and 2500, and as highly concentrated if the HHI is above 2500. See Section 5.3 of Horizontal Merger Guidelines (2010).

Second, we find that, when choosing among ETFs, investors care about the identities of index providers, although there is no material difference in return profiles between indexes that various index providers deliver. Indeed, as the global head of iShares and index investments at BlackRock noted, "One of our close partners is MSCI. Often it'll be MSCI that brings us to a client."[6] Consistent with the "brand-value" view expressed by this senior market participant, we find that index-provider fixed effects alone can explain about 21% of variations in ETF assets. Even after controlling for ETF-sponsor, time, and ETF-category fixed effects, management fees, and past returns, index providers can still explain 8% of additional variations in ETF assets. In contrast, we find that the index-provider fixed effects have literally zero explanatory power for ETF returns. This finding is also consistent with the conclusion drawn in an industry report by BNY Mellon: "There is minimal difference between several index providers that serve the U.S. and global equity markets in terms of performance; while methodology varies among indexes, those variances are largely tempered by capitalization weighting."[7]

Third, we show that a large fraction of ETF sponsors' revenues are paid to index providers in the form of index licensing fees. Specifically, we collect the first data on the licensing fees between index providers and ETF sponsors by reading all ETF filings on the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) of the SEC. Since licensing fees are disclosed by ETF sponsors on a voluntary basis,[8] only about 10% of ETFs in our sample disclose their licensing fees. Despite this limitation and possible selection bias, our novel data enable us to conduct the first analysis of ETF index licensing fees.[9] Based on the best available information that we can obtain, we find that more than 95% of the licensing fees are imposed in the form of "percentage-of-AUM" fees, with the remaining licensing fees applied as flat fees. In other words, index licensing fees are mostly tied to the assets of ETFs. We estimate that the index

[6]See https://www.bloomberg.com/news/articles/2017-11-27/index-providers-rule-the-world-for-now-at-least.

[7]See https://www.morningstar.com/lp/asset-management-in-an-era-of-cost-pressure.

[8]Licensing fees are operating expenses of the ETF, which are reflected in its management fees. However, because the SEC does not consider index providers to be advisers, licensing fees are not disclosed separately.

[9]Our sample does include some of the large and heavily traded ETFs, such as the SPDR S&P 500 ETF (SPY), the Invesco QQQ ETF (QQQ), and the SPDR Dow Jones Industrial Average ETF (DIA).

licensing fees comprise about one-third of all ETF management fees that ETF sponsors collect from ETF investors. This fraction has also increased steadily, from 31.4% in 2010 to 35.7% in 2019. Not surprisingly, this trend leads ETF sponsors to complain about the index licensing fees.[10]

In the second part of the paper, we build a structural model that incorporates the two-tiered competition between index providers for ETFs and between ETFs for investors. This structural approach allows us to (*i*) quantitatively assess the (un)competitiveness among index providers behind ETFs; (*ii*) decompose index providers' costs and markups, which are unobserved from the data; and (*iii*) conduct counterfactual analysis of index providers' market power and study the influence on ETF management fees paid by investors.

Our structural model consists of a discrete number of index providers, a discrete number of ETF sponsors, and a continuum of investors. In the first stage, each index provider lists a licensing fee for using its index, and each ETF sponsor chooses among all available index providers to form an ETF. The competition structure is modeled using the technology of Eaton and Kortum (2002). That is, if an ETF sponsor expects a higher profit from using an index provider's index, the probability that ETF sponsor chooses the index provider is higher. Because of market frictions, however, such as persistent relationships and switching costs, the ETF sponsor may not always choose the index provider that generates the highest expected profit.

In the second stage, each ETF, which is formed by a pair consisting of an index provider and an ETF sponsor, competes for investors. Specifically, each ETF sponsor optimally chooses the ETF management fee that maximizes its own profit. Because the index licensing agreement is signed in the first stage, each ETF sponsor treats the licensing fee as part of its marginal costs when determining ETF management fees. We model investors' choices of ETFs using a

[10]For example, the CEO of Amundi was quoted by Financial Times "Index fees are a real problem. These providers are an oligopoly and the prices they charge are out of line with the value they add." See https://www.ft.com/content/e886b2d2-e852-3071-85c1-c9a57113d8a5. A Global Head of Vanguard was quoted by Morningstar.com, "What we have seen over the last several years is that a larger and larger percentage of the total expense ratio has been eaten by index licensing fees." See https://www.morningstar.com/articles/569429/vanguard-index-swap-all-about-cost.

standard discrete choice framework. Investors care about ETF management fees, past returns, ETF categories, as well as the identities of index providers and ETF sponsors.

We structurally estimate the model using the top twenty U.S. equity ETFs, while taking the remaining ETFs as an outside option. We choose the top twenty ETFs as of December 2019, and they hold about 60% of all U.S. equity ETF assets. We explicitly model the top twenty ETFs because they are mostly broad-market ETFs and, importantly, there exists significant market segmentation between board-market ETFs and smaller and more specialized ETFs, such as thematic ETFs (Ben-David, Franzoni, Kim, and Moussawi, 2021a). It is worth noting that our results are not sensitive to this particular choice. In Appendix C, we estimate the model using the top fifty ETFs, which hold about 80% of all equity ETF assets, and we obtain similar conclusions.

Our structural estimation reveals several results. First, the key structural parameter shows that the index provider market is highly uncompetitive. Specifically, if index provider A can offer 1% higher profit for ETFs than index provider B, the probability that an ETF chooses index provider A is only 0.53% higher than the probability that the ETF chooses index provider B. In contrast, if index providers were perfectly competitive, index provider A should be always be chosen over B. Such a low elasticity implies very limited substitutability across index providers, which is consistent with persistent indexing relationships and significant market power wielded by index providers.

Second, about 60% of index licensing fees are markups. Over our sample period, the estimated licensing fees are about 4.7 bps of ETF's AUM. We estimate that the marginal costs of index provision are about 1.9 bps, while markups are about 2.8 bps. Thus, the Lerner index (=markup/licensing fees) of index providers is about 60%, indicating that index providers charge very high markups for index provision. Aligned with our estimate, Financial Times (2019) estimates the profit margin of the top three index providers to be about 65% as of 2019. Moreover, we find that markups charged by index providers are very stable, although the marginal cost of index provision has decreased over time. As a result, the Lerner index for index providers has increased from 55.6% in 2012 to 63.4% in 2019. Consistent with this increasing trend as we estimate, a major index provider, MSCI, reports that the operating margin for its index segment increased by about 12% from 2014 to 2019 (MSCI, 2020). In comparison, we estimate that about 30% to 40% of management fees that investors pay to ETF sponsors reflect markups of ETF sponsors. This is also aligned with the profit margin of BlackRock, which is reported to be about 37% over our sample period (BlackRock, 2020).

Third, we conduct a counterfactual analysis in which index providers are competitive and have no market power. In the counterfactual scenario, index providers set licensing fees equal to their marginal costs. The twenty ETFs, while keeping their equilibrium index providers, jointly change management fees optimally under the counterfactual licensing fees. In the counterfactual scenario, we find that ETF marginal costs decrease by about 2.8 basis points and the markup charged by ETF sponsors is similar to that in the baseline scenario. As a result, the ETF management fees decline by 2.8 basis points, a 30% reduction relative to the baseline scenario. Our analysis shows that increasing the competitiveness of index providers reduces ETF marginal costs, which are passed on to ETF investors through lower management fees at an almost one-to-one pass-through rate.

Overall, our results have several potential policy implications. First, licensing fees are currently disclosed on a voluntary basis currently, and the SEC could require mandatory disclosure of such fees. Although investors do not pay the licensing fees directly, our results show that licensing fees are effectively passed on to investors through higher management fees. Improved disclosure could help investors, regulators, and academics better understand the composition of management fees. Second, although ETF management fees have been trending downwards in recent years (see, for example, the 2020 ICI Fact Book), our results indicate that high licensing fees hinder the further reduction of management fees. Better policies that would improve competition could be proposed to reduce frictions between index providers and ETF sponsors.

**Related literature.** Our paper contributes to the growing literature on ETFs by unpacking the black box of index providers. To the best of our knowledge, we are the first to study the structure of competition between index providers and ETF sponsors and to show that matching and contracting between index providers and ETF sponsors matter to the first order of ETF management fees charged to investors. Relatedly, Robertson (2019) finds that the index providers of 81 of 571 U.S. equity ETFs are affiliated with ETF sponsors (so-called "self-indexing") and that these ETFs charge relatively higher management fees. While affiliated index providers are indeed relevant to small ETFs, the large ETF sponsors and index providers, which capture over 95% of total AUM, are not affiliated with each other. Mahoney and Robertson (2021) discuss the legal aspects of index providers as investment advisers. Akey, Robertson, and Simutin (2021) show that about 20% of ETFs track proprietary indexes and these ETFs charge higher management fees but generate worse returns. Kostovetsky and Warner (2021) show that ETF benchmarks with larger index providers are able to attract more capital from investors, consistent with our stylized fact regarding the brand value of index providers. The competition between index providers and its effect on index licensing fees and ETF management fees, which are the key to our analysis, are not studied in these papers.

Our paper is also related to the recent research on the bright and dark sides of ETFs. Azar, Schmalz, and Tecu (2018) study the implications of passive investment for corporate governance and corporate power. Huang, Song, and Xiang (2020) find that index providers and ETF sponsors conduct extensive data mining when constructing smart beta indexes so as to attract investment flows, while Ben-David, Franzoni, Kim, and Moussawi (2021a) find evidence that thematic ETFs are constructed and offered to cater to investors' sentiment. Brown, Cederburg, and Towner (2021) find that ETFs that have similar returns but higher management fees and less liquid than their competitors attract excess capital, and Khomyn, Putniņš, and Zoican (2020) show that more liquid ETFs attract shorter horizon investors and charge higher management fees. Moreover, some argue that ETFs increase asset volatility and harm liquidity (e.g., Israeli, Lee, and Sridharan, 2017; Ben-David, Franzoni, and Moussawi, 2018; Da and Shive, 2018; Agarwal, Hanouna, Moussawi, and Stahel, 2019; Pan and Zeng, 2019), while others find evidence that ETFs improve market efficiency (e.g. Box, Davis, Evans, and Lynch, 2020; Glosten, Nallareddy, and Zou, 2020; Huang, O'Hara, and Zhong, 2021).

Finally, our paper contributes to the growing literature that explores the industrial organization of financial markets with structural techniques (Bao and Ni, 2017; Egan, Hortaçsu, and Matvos, 2017; Benetton, 2018; Buchak, Matvos, Piskorski, and Seru, 2018; Koijen and Yogo, 2019; Buchak, Matvos, Piskorski, and Seru, 2020). Our paper is related to Hortaçsu and Syverson (2004), who develop a search model to understand fund proliferation and fee dispersion in S&P 500 index funds; Egan, MacKay, and Yang (2020), who study the ETF market with a structural demand model to infer investors' expectations from ETF demand; and Jiang (2020), who builds a quantitative model to understand how relationship lending between shadow and traditional banks affects competition in the downstream mortgage market.

The rest of the paper is structured as follows. Section 2 describes the data. Section 3 documents the three stylized facts about index providers in U.S. equity ETF markets. Section 4 presents a structural model of index providers and ETF sponsors. Section 5 estimates the model and presents results. Section 6 concludes. The appendices provide additional results and robustness checks.

## 2 Data

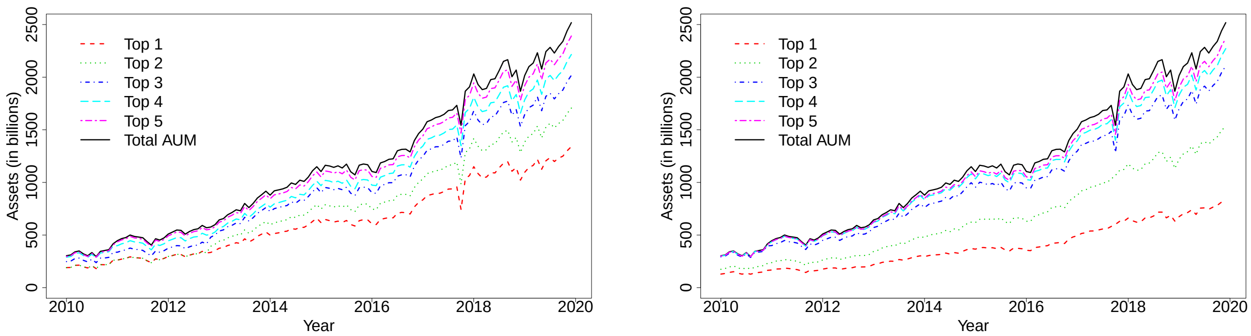
We take several steps to construct the sample. First, we obtain a list of U.S. equity ETFs from Morningstar spanning a 10-year period from January 2010 through December 2019. Specifically, we exclude leverage ETFs and synthetic ETFs. Second, for each ETF we identify its underlying index manually and collect the information on the index from its official website or from professional third-party websites (e.g., ETF.com). We then merge the list of ETFs with the CRSP mutual fund database to obtain monthly returns, expense ratios, and AUM. After this step, we obtain 598 U.S. equity ETFs and provide summary statistics in Table 1.

**Table 1**
## Summary statistics

This table reports the summary statistics of our sample at the ETF, ETF sponsor, and index provider levels. Our sample includes the U.S. equity ETFs (excluding leverage and synthetic ETFs) and spans from January 2010 to December 2019.

| | Mean | SD | 25th | 50th | 75th |
|---|---|---|---|---|---|
| **Panel A: ETF level (obs = 598)** | | | | | |
| AUM ($ million) | 2037.29 | 9121.63 | 47.70 | 209.16 | 814.13 |
| Monthly return (%) | 1.07 | 0.47 | 0.91 | 1.06 | 1.20 |
| Management fee (%) | 0.37 | 0.20 | 0.20 | 0.35 | 0.50 |
| Turnover ratio (%) | 62.85 | 209.15 | 16.19 | 32.32 | 61.14 |
| **Panel B: ETF sponsor level (obs = 68)** | | | | | |
| Total AUM ($ million) | 16939.70 | 68657.61 | 35.43 | 142.25 | 1154.38 |
| # of ETF | 6.79 | 15.56 | 1.00 | 1.77 | 4.14 |
| # of matched index providers | 1.68 | 1.58 | 1.00 | 1.00 | 1.89 |
| **Panel C: Index provider level (obs = 77)** | | | | | |
| Total AUM ($ million) | 15635.70 | 77678.42 | 45.52 | 126.43 | 1136.98 |
| # of ETF | 5.91 | 18.16 | 1.00 | 1.00 | 3.00 |
| # of matched ETF sponsors | 1.43 | 1.45 | 1.00 | 1.00 | 1.00 |

The results reported in Panel A of Table 1 indicate the average ETF AUM of the ETFs of $2.04 billion with a standard deviation of $9.12 billion. The distribution of ETF AUM is highly skewed, with a median AUM of $209 million. The average expense ratio is 37 bps per year with a standard deviation of 20 bps. The results reported in Panel B of Table 1 show that the ETFs in our sample are offered by 68 ETF sponsors. Each ETF sponsor offers, on average, 6.79 ETF products, which track the indexes constructed by 1.68 index providers. The results reported in Panel C indciate that the ETFs in our sample track indexes constructed by 77 index providers. Each index provider has, on average, about 5.91 ETFs tracking their constructed indexes and works with 1.43 ETF sponsors. In the next section, we provide a more detailed analysis of the matching between ETF sponsors and index providers.




**(A) Assets tracking the top five index providers**

**(B) Assets invested in the top five ETF sponsors**

**Figure 1. Assets related to the top five index providers and ETF sponsors.** Panel (A) shows the total assets of ETFs that use indexes constructed by the top five index providers (by tracking assets as of December 2019): S&P Dow Jones, CRSP, FTSE Russell, MSCI, and NASDAQ. Panel (B) shows the total assets of ETFs offered by the top five ETF sponsors (by AUM as of December 2019): iShares, Vanguard, State Street, Invesco, and Schwab.

## 3 Stylized Facts about Index Providers

In this section, we document three stylized facts about index providers in ETF markets: (*i*) the ETF indexing market is highly concentrated among only a few large index providers; (*ii*) investors care about the identities of index providers when choosing ETFs, even though there are no significant differences in returns among indexes constructed by various index providers; and (*iii*) about one-third of all ETF management fees are paid to index providers in the form of index licensing fees.

### 3.1 Concentration of the Index Provider Industry

We begin by showing that ETF markets and index markets are highly concentrated. Specifically, Figure 1 plots the total assets tracking indexes provided by the top five index providers (S&P Dow Jones, CRSP, FTSE Russell, MSCI, and NASDAQ) and the total assets managed by the top five ETF sponsors (iShares, Vanguard, State Street, Invesco, and Schwab). As we can see, the top index providers and ETF sponsors capture a very large market share. The extremely high market shares of top index providers and ETF sponsors is especially striking given that the total AUM of all ETFs has grew more than fivefold from 2010 to 2019. In Table

**Table 2**
**Market share of top index providers and ETF sponsors in December 2019**

In this table, we provide the individual and cumulative market shares of the top five index providers and ETF sponsors in the U.S. equity ETF market as of December 2019.

| Index provider | | | ETF sponsor | | |
|---|---|---|---|---|---|
| Name | Market share | Cum. market share | Name | Market share | Cum. market share |
| S&P Dow Jones | 53.24% | 53.24% | iShares | 33.17% | 33.17% |
| CRSP | 14.51% | 67.75% | Vanguard | 27.82% | 60.99% |
| FTSE Russell | 12.37% | 80.12% | State Street | 22.69% | 83.68% |
| MSCI | 7.86% | 87.98% | Invesco | 6.52% | 90.20% |
| NASDAQ | 6.97% | 94.95% | Schwab | 3.87% | 94.07% |

2, we also report the market share captured by the top index providers and ETF sponsors, measured by total AUM, as of December 2019. The top five index providers and the top five ETF sponsors both capture about 95% of the market. The top ETF sponsor, iShares, has captured about 33% market share, and the top index provider for U.S. equity ETFs, S&P Dow Jones, itself has captured more than 50% of the market.

To quantify market concentration, we calculate the Herfindahl-Hirschman index (HHI) of ETF sponsors and ETF index providers for each month. Over our sample period, the monthly average of the HHI of ETF sponsors is 2,527.31, and the HHI of index providers is even higher, averaging 3,293.59, much higher than the 2,500 level, which the U.S. Department of Justice and the Federal Trade Commission regard as a highly concentrated industry (Horizontal Merger Guidelines, 2010).

Next, we show that most ETF sponsors match with one major index provider and that most index providers match with one major ETF sponsor. In Table 3, we report the matching between index providers and ETF sponsors.[11] Panel A lists the distributions of AUM across various index providers from a given ETF sponsor's perspective, and the panel should be read left to right. For example, the top left cell indicates that 57.1% of iShares' AUM uses S&P Dow Jones as index providers. We highlight cells that are over 50%. As can be seen, every top ETF sponsor has a major index providing partner. Specifically, iShares, State State, and

[11] The results are similar when using total revenue=AUM×management fees, and shown in Table B.1.

**Table 3**

## Matching between index providers and ETF sponsors

In this table, we report matching between index providers and ETF sponsors. We use "others" to represent all index providers or ETF sponsors other than the top five. Panel A reports the distribution of AUM across various index providers from a given ETF sponsor's perspective. Panel B reports the distribution of AUM across various ETF sponsors from a given index provider's perspective. We highlight cells where the figure is above 50%. The sample period is December 2019.

**Panel A: From ETF sponsors' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | **57.1%** | 0.0% | 29.3% | 9.3% | 1.2% | 3.1% |
| Vanguard | 21.1% | **52.2%** | 5.8% | 14.9% | 6.0% | 0.0% |
| State Street | **97.7%** | 0.0% | 0.2% | 0.1% | 0.0% | 2.0% |
| Invesco | 33.2% | 0.0% | 5.1% | 0.0% | **58.1%** | 3.6% |
| Schwab | **88.4%** | 0.0% | 10.7% | 0.0% | 0.0% | 1.0% |
| Others | 11.3% | 0.0% | 3.9% | 10.1% | 18.9% | **55.8%** |

**Panel B: From index providers' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | 35.6% | 0.0% | **78.6%** | 39.3% | 5.5% | 20.4% |
| Vanguard | 11.0% | **100.0%** | 13.1% | **52.8%** | 24.0% | 0.0% |
| State Street | 41.6% | 0.0% | 0.5% | 0.3% | 0.0% | 8.8% |
| Invesco | 4.1% | 0.0% | 2.7% | 0.0% | **54.3%** | 4.6% |
| Schwab | 6.4% | 0.0% | 3.3% | 0.0% | 0.0% | 0.7% |
| Others | 1.3% | 0.0% | 1.9% | 7.6% | 16.1% | **65.5%** |

Schwab rely mainly on S&P Dow Jones, Vanguard uses CRSP, and Invesco uses NASDAQ. In Panel B we report the distribution of AUM across various ETF sponsors from a given index provider's perspective, and the panel should be read top to bottom. For example, the top left cell indicates that 35.6% of S&P Dow Jones' AUM use iShares as ETF sponsors. With the exception of S&P Dow Jones, all other index providers rely mainly on one ETF sponsor. This matching between index providers and ETF sponsors could be caused by persistent relationships over time, which can have benefits such as improving collaboration between index providers and ETF sponsors. Such persistent matching also, however, likely leads to large switching costs and high market power for index providers.

Regarding the results we report in Table B.2 and Table B.3 of Appendix B, we find that the results reported in Table 3 do not change much when we use a time snapshot other than December 2019, such as December 2013 or December 2016. The matching between ETF sponsors and index providers is rather stable over time, because during our sample period most ETFs never switch index providers.

### 3.2 The Identity of Index Providers Matters for Investor Choice

We proceed now to show that the identity of index providers matters for investor choices. Specifically, we explore the role of index providers in the ETF market using a regression framework, and we estimate various variations of the following regression specification:

$$y_{kt} = \beta X_{kt} + \gamma_i + \gamma_j + \gamma_c + \gamma_t + \epsilon_{kt}, \tag{1}$$

where $X_{kt}$ are characteristics of ETF $k$ offered by index provider $i$ and ETF sponsor $j$ in category $c$ and month $t$, and $\gamma_i$, $\gamma_j$, $\gamma_c$, and $\gamma_t$ are index-provider, ETF-sponsor, category, and month fixed effects, respectively. By shutting down various fixed effects and comparing the corresponding adjusted $R^2$s, we study the contribution of multiple variables in explaining variations in the outcome variable $y_{kt}$.

Table 4 reports the results derived from regression (1) on our main variable of interest, (log) AUM $y_{kt}$ of ETF $k$ in month $t$. The first column shows that index-provider fixed effects alone can explain more than 20% of the variation in AUM. ETF sponsors are also important in capturing variation in AUM with an $R^2$ of around 30%. Category and time fixed effects are less important than index provider and ETF sponsor in explaining variation in AUM. The $R^2$ with category fixed effects is 0.05, while time fixed effects account for only about 1% of the variation in AUM, suggesting that aggregate time-series trends mask a lot of cross-sectional heterogeneity.

A key empirical concern is that the return profiles of indexes can vary across index providers. Investors do not care about the identities of index providers per se but do care about index returns, which correlates with index providers. To address this concern, we include additional controls in regression (1).

In column (5) of Table 4 we show the estimate of equation (4) with ETF sponsor, category, and time fixed effects, while controlling for ETF management fees and past returns.

**Table 4**

**Index providers matter for investor choices**

This table reports the estimates of equation (1) with various sets of fixed effects and controls. The dependent variable is (log) AUM. We report the interquartile range (IQR) of the $y$ variable and residuals. The sample consists of each ETF×month observation of U.S. equity ETFs from January 2010 through December 2019.

| | Separate fixed effects | | | | Role of index providers | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Index provider | ETF sponsor | Category | Time | ETF sponsor<br>Category<br>Time | Index provider<br>ETF sponsor<br>Category<br>Time |
| Management fees (bps) | | | | | $-0.039^{***}$ | $-0.052^{***}$ |
| | | | | | (0.001) | (0.001) |
| Past 1-year return (%) | | | | | $0.193^{***}$ | $0.211^{***}$ |
| | | | | | (0.011) | (0.010) |
| $R^2$ | 0.21 | 0.30 | 0.05 | 0.01 | 0.43 | 0.50 |
| Adjusted $R^2$ | 0.21 | 0.30 | 0.05 | 0.01 | 0.42 | 0.50 |
| Y IQR | 2.9 | 2.9 | 2.9 | 2.9 | 2.9 | 2.9 |
| Residuals IQR | 2.3 | 2.2 | 2.7 | 2.7 | 1.9 | 1.6 |
| Observations | 38,757 | 38,757 | 38,757 | 38,757 | 38,757 | 38,757 |

As expected, if investor demand is downward sloping in price, higher management fees are associated with lower AUM. Also, consistent with Dannhauser and Pontiff (2019), higher past ETF returns are associated with higher AUM.[12] The overall $R^2$ is 0.43. We also report the interquartile range (IQR) as a measure of dispersion in (log) AUM. In the data the interquartile range of (log) AUM is 2.9. The interquartile range of the residuals from the estimates reported in column (5) is 1.9, which represents approximately a 35% decline in dispersion.

Finally, in column (6) of Table 4 we show the estimates after adding index-provider fixed effects to the specification of column (5). After controlling for ETF sponsor, category, time, management fees, and past returns, index-provider fixed effects increases the $R^2$ by about 0.07, from 0.43 to 0.50. Additionally, the dispersion in the IQR of the residuals declines to 1.6, which represents an additional 10-percentage-point decline relative to the specification without index-provider fixed effects.

Overall, the results reported in Table 4 show that index providers contribute significantly

[12]It is well documented that investors chase past performance (e.g. Chevalier and Ellison, 1997; Ben-David, Li, Rossi, and Song, 2021b).

to explaining dispersion in AUM. The identity of index providers matters even after controlling for ETF management fees and past returns, suggesting that investors value non-price characteristics of index providers such as brand reputation. Consistent with our findings, Mahoney and Robertson (2021) find that large index providers can help to attract ETF flows.

To further understand the role of index providers we estimate equation (1) using management fees and monthly returns as the dependent variable. Table 5 shows the results. Column (1) of Panel A shows that index-provider fixed effects alone explain about 64% of the variation in management fees. The IQR of management fees is about 30 basis points. Controlling for index providers alone reduces the IQR by about two-thirds to about 10 basis points. The large explanatory power of index providers for management fees can come from two channels: (a) index providers' licensing fees affect ETFs' costs, which are then passed on to investors via management fees and (b) index providers affects the attractiveness of ETFs to investors, which allows ETFs to charge differential management fees. We incorporate both effects in our structural model.

Panel A also shows that ETF sponsor fixed effects have considerable explanatory power with an $R^2$ equal to 0.67. As is the case with the results obtained using AUM as the dependent variable, here category and time fixed effects have weaker explanatory power. The $R^2$s for category or time fixed effects are 0.19 and 0.01, respectively. Aggregate time-series variation in fees hides much of the cross-sectional dispersion, as also documented in Ben-David, Franzoni, Kim, and Moussawi (2021a). Comparing the results reported in columns (5) and (6) shows that adding index-provider fixed effects to ETF sponsor, category, and time fixed effects raises the $R^2$ by about 0.08 and reduces the IQR of the residuals from 12.8 to 8.9.

Finally, Panel B studies ETF returns. In contrast to what the results for AUM and management fees imply, index-provider and ETF-sponsor fixed effects have little explanatory power for returns. In both cases, the $R^2$ is about 0.01. Category fixed effects have an $R^2$ of about 0.06. The single most important variable in explaining dispersion in returns are time fixed effects, which alone capture almost 50% of the variation in returns.

**Table 5**

## Index providers: fees and returns

This table reports the estimates of equation (1) with various sets of fixed effects and controls. The dependent variable for Panel A is ETF management fees. The dependent variable for Panel B is ETF monthly returns. We report the interquartile range (IQR) of the $y$ variable and residuals. The sample consists of each ETF×month observation of U.S. equity ETFs from January 2010 through December 2019.

| | Separate fixed effects | | | | Role of index providers | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Index provider | ETF sponsor | Category | Time | ETF sponsor<br>Category<br>Time | Index provider<br>ETF sponsor<br>Category<br>Time |
| **Panel A: Management fees** | | | | | | |
| $R^2$ | 0.64 | 0.67 | 0.19 | 0.01 | 0.76 | 0.84 |
| Adjusted $R^2$ | 0.64 | 0.66 | 0.19 | 0.00 | 0.76 | 0.84 |
| Y IQR (bps) | 30 | 30 | 30 | 30 | 30 | 30 |
| Residuals IQR (bps) | 10.4 | 17.6 | 22.7 | 32.2 | 12.8 | 8.9 |
| Observations | 38,757 | 38,757 | 38,757 | 38,757 | 38,757 | 38,757 |
| **Panel B: Returns** | | | | | | |
| $R^2$ | 0.01 | 0.01 | 0.06 | 0.49 | 0.56 | 0.56 |
| Adjusted $R^2$ | 0.01 | 0.01 | 0.06 | 0.49 | 0.56 | 0.56 |
| Y IQR (%) | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 |
| Residuals IQR (%) | 1.2 | 1.2 | 1.2 | 0.6 | 0.6 | 0.6 |
| Observations | 38,757 | 38,757 | 38,757 | 38,757 | 38,757 | 38,757 |

To summarize, we find that index providers' identities: (*i*) matter for AUM even after controlling for other determinants of investors demand (e.g., management fees, past returns, ETF sponsor) and (*ii*) explain a large (tiny) fraction of dispersion in fees (returns). These findings are consistent with the views expressed by market participants as quoted in Petry, Fichtner, and Heemskerk (2019):"At the end of the day, those products (i.e., indexes) are homogeneous and exchangeable. It's like water, there are small differences why Evian is more expensive. Those are minimal differences, but the price tags are very different! MSCI is famous for being expensive — not because they have better data or indices, but because they are the brand that is most used in the world. Brand is everything!"

**Table 6**
**Comparing ETFs with and without licensing fee disclosure**

This table compares ETFs that report licensing fees and ETFs that do not report licensing fees. Specifically, we search all ETF filings on the EDGAR of the SEC. Out of the 598 ETFs in our sample, 52 ETFs report the index licensing fees.

| | Mean | SD | 25th | 50th | 75th |
|---|---|---|---|---|---|
| **Panel A: 52 ETFs with licensing fees reported** | | | | | |
| AUM ($ million) | 6915.85 | 24682.11 | 213.40 | 714.65 | 3410.53 |
| Monthly return (%) | 1.00 | 0.40 | 0.89 | 1.05 | 1.16 |
| Management fees (%) | 0.50 | 0.23 | 0.19 | 0.60 | 0.66 |
| **Panel B: 546 ETFs without licensing fees reported** | | | | | |
| AUM ($ million) | 1568.37 | 5582.83 | 44.32 | 165.75 | 740.98 |
| Monthly return (%) | 1.08 | 0.47 | 0.91 | 1.07 | 1.20 |
| Management fees (%) | 0.35 | 0.19 | 0.20 | 0.35 | 0.47 |

## 3.3 Analysis of Index Licensing Agreements

In this section, we provide an analysis of index licensing fees. To this end, we collect index licensing agreements and fees between index providers and ETF sponsors by manually searching ETF filings on the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) of the U.S. SEC.[13] Specifically, we look for the keywords "licensing fee" and "license fee" within the ETF prospectus. Because ETFs disclose licensing fees on a voluntary basis, we obtain licensing fees for 52, or about 9%, of the U.S. equity ETFs in our sample. Admittedly, whether an ETF discloses licensing fees is an endogenous choice and our data, despite our best effort, suffers selection bias. Table 6 provides a comparison of various ETF characteristics that disclose licensing fees and ETFs that do not. As we can see, ETFs that disclose licensing fees have, on average, larger AUM and charge higher management fees to investors than ETFs that do not disclose licensing fees. The return profiles of these two types of ETFs are similar.

Across the 52 ETFs for which we are able to obtain licensing fees, the typical licensing fee contract is "$x$ bps of AUM + \$y" per year, where $x$ can have various breakpoints depending on AUM, and $y$ can be 0. The other less common contractual form, which is used by only

[13] See deHaan, Song, Xie, and Zhu (2021) for more details on EDGAR.

**Table 7**

## Analysis of licensing fees

This table presents the results of an analysis of index licensing fees. Columns (1) to (3) calculate the AUM-weighted average, the simple average, and the median licensing fees as a fraction of ETF management fees. Columns (4) and (5) report the fractions of licensing fees related to ETF AUM and the fractions of fixed licensing fees, respectively.

| | Licensing fees as fractions of management fees | | | Decomposition of licensing fees | |
|---|---|---|---|---|---|
| Year | AUM-weighted mean (%) | Simple mean (%) | Median (%) | AUM-based fee (%) | Fixed fee (%) |
| | (1) | (2) | (3) | (4) | (5) |
| 2010 | 31.4 | 23.2 | 19.3 | 97.3 | 2.7 |
| 2011 | 32.5 | 20.3 | 16.7 | 98.1 | 2.0 |
| 2012 | 32.6 | 19.7 | 16.7 | 97.9 | 2.1 |
| 2013 | 32.7 | 17.8 | 16.7 | 95.8 | 4.2 |
| 2014 | 33.9 | 21.6 | 19.3 | 91.7 | 8.3 |
| 2015 | 33.7 | 21.7 | 19.8 | 93.4 | 6.6 |
| 2016 | 34.4 | 20.8 | 17.8 | 94.9 | 5.1 |
| 2017 | 35.0 | 21.1 | 19.0 | 97.3 | 2.7 |
| 2018 | 35.7 | 21.3 | 18.5 | 98.3 | 1.7 |
| 2019 | 35.7 | 21.3 | 19.3 | 98.6 | 1.4 |

three out of 52 ETFs, is "max of $x$ bps of AUM and \$y" per year. For example, consider three well-known ETFs:

- SPDR S&P 500 ETF has a licensing fee of $x = 3$ bps of AUM plus a flat fee $y = \$600,000$
- SPDR Dow Jones Industrial Average ETF has a licensing fee of $x = 4$ bps of AUM and the flat fee $y = \$0$
- Invesco QQQ ETF has no flat fee ($y = \$0$) and charges $x = 9$ bps for AUM under \$25 billion and $x = 8$ bps for AUM above \$25 billion. So the formula for the licensing fee for Invesco QQQ ETF is $9\text{bps} \times \min(\text{AUM}, \$25\text{b}) + 8\text{bps} \times \max(\text{AUM} - \$25\text{b}, 0)$

In Table 7, we provide summary statistics for licensing fees for each year from 2010 through 2019. As can be seen in the last two columns, the AUM-based component comprises more than 95% of the total licensing fee, and the flat-fee component is just a tiny fraction of the licensing fee.

Columns (1) to (3) of Table 7 further report index licensing fees as a fraction of the total ETF management fees that ETF investors pay. The ETF licensing fee is on average 21% of the ETF management fee, and the AUM-weighted average ranges from about 32% to about 36%, suggesting that larger ETFs pay out a higher fraction of total management fees to index providers. We further regress licensing fees as a fraction of ETF management fees on (log) ETF AUM and find a highly significant positive relationship ($t = 3.5$). In other words, larger ETFs pay a higher fraction of total management fees to index providers in the form of index licensing fees. Another striking pattern revealed in Table 7 is that, as a fraction of the ETF management fee, the AUM-weighted licensing fee increases steadily over time, from about 31% in 2010 to 36% in 2019.

In summary, this section shows that index providers capture a large fraction of the total revenue of the ETF business. In the next section, we build a structural model to further analyze the market power of index providers.

## 4 A Model of Index Providers and ETF Sponsors

Based on the three stylized facts documented above, in this section, we present a structural equilibrium model of index providers and ETF sponsors. This structural approach allows us to (*i*) quantitatively assess the (un)competitiveness among index providers behind ETFs, (*ii*) decompose costs and markups of index providers, which are unobserved from the data, and (*iii*) conduct counterfactual analysis of index providers' market power and study the influence on ETF management fees paid by investors.

The model works as follows. In each period $t$ a continuum mass $L_t$ of investors, indexed by $l$, choose among a discrete number of differentiated ETFs, indexed by $k = 1, 2, \ldots, K_t$. Each ETF $k$ consists of an ETF sponsor $j = 1, \ldots, J_t$ and an index provider $i = 1, \ldots, I_t$. Within the each period $t$, the timing is as follows:

- Each index provider $i$ sets licensing fees $\rho_i$ for using its index
- Each ETF $k$, which is set up by a given ETF sponsor, chooses an index provider
- Each ETF $k$ sets management fees $f_k$

- Investors choose the ETFs in which they invest their money

In what follows, we specify each agent's optimization problem in turn. Our model is static for each period $t$. In the notation, for simplicity we omit the subscript $t$, which indexes time.

### 4.1 Investors

Each investor $l$ seeks to buy one indivisible unit of an ETF. The indirect utility enjoyed by investor $l$ for choosing ETF $k$ that is sponsored by $j$ with the index provided by $i$ is given by:

$$u_{lk} = -\alpha f_k + \beta X_k + \gamma_{ij} + \xi_k + \epsilon_{lk}, \tag{2}$$

where $f_k$ is the management fee charged by ETF $k$; $X_k$ corresponds to vectors of observable features of ETF $k$, such as past returns; the interacted fixed effects $\gamma_{ij}$ for index provider $i$ and ETF sponsor $j$ capture observable and unobservable characteristics such as index provider brand value and ETF sponsor quality, as well as potential synergies between index providers and ETF sponsors; $\xi_k$ is an error term capturing additional unobservable characteristics of ETF $k$; and $\epsilon_{lk}$ is an idiosyncratic shock that varies across investors and ETFs.

Investor $l$ chooses the ETF $k$ that delivers the highest utility among the $K$ ETFs that are available on the market. As an alternative to choosing one of the $K$ ETFs, each investor also has the option of not choosing any ETF and investing its money in other asset classes. We normalize the utility of such a choice to zero ($u_{l0} = 0$). Hence, the probability that investor $l$ chooses to invest in ETF $k$ is given by:

$$s_{lk} = \text{Prob}(u_{lk} > u_{lk'}, \forall k' \in \{0, 1, \ldots, K\}). \tag{3}$$

When taking the model to the data, we assume that unobservables $\epsilon_{lk}$ in equation (2) follow an i.i.d. type-1 extreme value distribution. Hence the probability that investor $l$ chooses to invest in ETF $k$ is given by:

$$s_{lk} = s_k = \frac{e^{-\alpha f_k + \beta X_k + \gamma_{ij} + \xi_k}}{1 + \sum_{k'=1}^{K} e^{-\alpha f_{k'} + \beta X_{k'} + \gamma_{i'j'} + \xi_{k'}}}, \tag{4}$$

where the first equality comes from the common parameters across investors. Summing across the continuum mass $L$ of investors in the market, we obtain the AUM of ETF $k$: $\text{AUM}_k = \sum_l s_{lk} = s_k L$.

### 4.2 ETF Sponsors

ETF sponsors maximize profits by setting optimal management fees $f_k$ for the ETFs they offer given their costs, which depend on the index provider they choose.[14]

Given the choice of index provider $i$, the profit of ETF $k$, which is sponsored by $j$, is given by:

$$\pi_{ki}(\rho_i) = \max_{f_k} (f_k - c_k(\rho_i)) s_k L, \tag{5}$$

where $c_k(\rho_i)$ is the marginal cost of offering ETF $k$, conditional on the choice of index provider $i$; $\rho_i$ is the licensing fee that index provider $i$ charges as a fraction of ETF's AUM; and $s_k$ is the market share of ETF $k$ from equation (4). Consistent with the results reported in last two columns of Table 7, we assume that licensing fee $\rho_i$ is paid as a percentage-of-AUM fee. In equation (5), both licensing fee $\rho_i$ and the equilibrium market share $s_k$ depend on index provider $i$, because the market share $s_k$ implicitly depends on the interacted fixed effect $\gamma_{ij}$ in equation (2).

The first-order condition of profits in equation (5) relative to management fees gives the standard markup pricing formula:

$$f_k(\rho_i) = c_k(\rho_i) + \overbrace{\frac{1}{\alpha(1 - s_k)}}^{\text{markup}}. \tag{6}$$

[14] In practice, large ETF sponsors usually offer multiple ETFs (see Table 1). For tractability, we assume that ETFs make the profit-maximization decision independently regardless of whether they belong to the same sponsor. We leave the investigation of multi-product strategies adopted by ETF sponsors to future research.

We assume that the marginal cost of ETFs consists of two components:

$$c_k(\rho_i) = \tilde{c}_k + \rho_i. \tag{7}$$

First, ETFs incur a marginal operating cost $\tilde{c}_k$, which could vary across ETFs. The component $\tilde{c}_k$ is exogenous in the model and does not depend on index providers. Second, ETF sponsors pay a licensing fee $\rho_i$ to index provider $i$ as a fraction of ETF's AUM. The licensing fees are optimally chosen by index providers, and ETF sponsors choose between various providers. We model this choice parsimoniously using the technology of Eaton and Kortum (2002), which has also recently being applied by Jiang (2020) in structural work on the U.S. mortgage market.

Formally, ETF $k$ chooses among index providers $i = 1, \ldots, I$ to maximize its total profits:

$$\Pi_{ki} = \pi_{ki}(\rho_i) \times \xi_{ki}, \tag{8}$$

where $\pi_{ki}(\rho_i)$ are the profits conditional on choosing index provider $i$ given in equation (5) and $\xi_{ki}$ is a structural error term capturing, for example, relationship persistence or synergies between index providers and ETF sponsors (see Table 3).

ETF $k$ chooses the index provider $i$ that delivers the highest total profits $\Pi_{ki}$ among the $I$ index providers that exist in the market. Hence the probability that ETF $k$ chooses index provider $i$ is given by:

$$q_{ki}(\rho_i) = \text{Prob}(\Pi_{ki} > \Pi_{ki'}, \forall i' \in \{1, 2, \ldots, I\}). \tag{9}$$

When taking the model to the data, we assume that unobservables $\xi_{ki}$ in equation (8) follow an i.i.d. type-2 extreme value distribution $G(\xi, \sigma) = e^{(-\gamma\xi)^{(-\sigma)}}$, where $\gamma = \Gamma(1 - 1/\sigma)$. Hence, the probability that ETF $k$ chooses index provider $i$ is given by:

$$q_{ki}(\rho_i) = \frac{\pi_{ki}(\rho_i)^\sigma}{\sum_{i'=1}^{I} \pi_{ki'}(\rho_{i'})^\sigma}. \tag{10}$$

The parameter $\sigma \in [0, \infty)$ is a key structural parameter that captures the competitive landscape of index providers. At one extreme of $\sigma = \infty$, ETF $k$ chooses the index provider $i$ that offers the highest profit $\pi_{ki}(\rho_i)$. At the other extreme of $\sigma = 0$, ETF $k$ chooses any index provider $i$ with equal probability regardless of the profit $\pi_{ki}(\rho_i)$. In general, a higher $\sigma$ implies stronger competition between index providers.

### 4.3 Index Providers

We now characterize the problem of index providers. Each index provider $i$ optimally chooses the licensing fee $\rho_i$ that maximizes its profits. The total profit of index provider $i$ is given by:

$$\pi_i = \max_{\rho_i} (\rho_i - \kappa_i) L \sum_k q_{ki}(\rho_i) s_k^*(\rho_i), \tag{11}$$

where $\kappa_i$ is the marginal cost of index provider $i$, $q_{ki}(\rho_i)$ is the probability that ETF $k$ that chooses index provider $i$ given by equation (10), and $s_k^*(\rho_i)$ is the market share of ETF $k$ when choosing (potentially counterfactual) index provider $i$ with licensing fee $\rho_i$. This market share is evaluated under the corresponding optimal choice of management fee $f_k^*(\rho_i)$ given by (6).

In (11), we model the costs of providing an index as the per-AUM marginal costs $\kappa_i$. These costs could arise from, for example, higher operational costs for educating a larger investor base about the index and the higher litigation risk.[15] In practice, there could also be a fixed cost for providing an index that does not vary with AUM. Such a fixed cost is likely, however, to be a nonsignificant component of index licensing fees, which are mostly AUM-based fees and have barely changed (in percentage terms) over the past ten years despite impressive growth in AUM in ETFs (see Section 3.3). Therefore, we model the per-AUM marginal cost $\kappa_i$ in (11).

The implicit assumption underlying (11) is that each ETF $k$ observes only the licensing fee contracts offered by various index providers to itself, but not to other ETFs. This is

[15] For example, SEC recently fined S&P $9 million for failing to update the VIX index in a timely fashion. See https://www.sec.gov/news/press-release/2021-84.

reasonable because, in practice, index agreements are rarely disclosed (see Section 3.3). Under this assumption, if index provider $i$ offers a licensing fee $\tilde{\rho}_i$ that deviates from equilibrium $\rho_i$ to an ETF $k$, the ETF interprets this deviation as specific to itself. ETF $k$ calculates the optimal (counterfactual) management fee $f_k^*(\tilde{\rho}_i)$ and market share $s_k^*(\tilde{\rho}_i)$ using the deviated licensing fee $\tilde{\rho}_i$, but assumes that other ETFs' index provider matching and management fees remain as equilibrium outcomes.

The first-order condition of index provider $i$'s profiting from equation (11) relative to licensing fees yields[16]

$$\rho_i = \kappa_i + \overbrace{\frac{\sum_k q_{ki}(\rho_i) s_k^*(\rho_i)}{\sum_k \alpha q_{ki}(\rho_i) s_k^*(\rho_i)(1 - s_k^*(\rho_i)) \left(\sigma(1 - q_{ki}(\rho_i)) + 1 - s_k^*(\rho_i)\right)}}^{\text{markup}}. \tag{12}$$

Two aspects of the index provider's first-order condition are worth emphasizing. First, index providers internalize the fact that setting a higher licensing fee reduces both the probability $q_{ki}(\rho_i)$ of being selected by an ETF and the market share $s_k^*(\rho_i)$ of the ETF itself, which passes on some of the higher licensing fees to investors in terms of higher management fees. Second, if ETFs are perfect substitutes (i.e., investors are perfectly elastic, $\alpha = \infty$) or if index providers are perfect substitutes ($\sigma = \infty$), licensing fees equal the marginal costs that index providers pay. In our model, although index providers do not face investors directly, the competitive landscape $\alpha$ for ETF investors affects the optimal licensing fee of index providers indirectly.

Index providers' optimal licensing fees, ETFs' optimal management fees and choice of index provider, and investors' optimal ETF choices, characterize the equilibrium in the ETF market.

[16] For the detailed derivation, see Appendix A.

## 5 Estimation, Results, and Counterfactual Analysis

In this section, we estimate our structural model, report results, and present counterfactual analyses.

### 5.1 Estimation

We structurally estimate the model using the top twenty U.S. equity ETFs (based on AUM in December 2019) while taking the remaining ETFs as an outside option.[17] We use a monthly panel for the period running from January 2010 through December 2019. Our focus on the largest ETFs is motivated by three main factors. First, the ETF market is quite concentrated. Despite the increase in the number of ETFs in the last ten years, the twenty largest ETFs as of December 2019 hold almost 60% of total U.S. equity ETF AUM.[18] Second, the top twenty ETFs are mostly broad-market ETFs that have long been established, and significant market segmentation and product differentiation exist between broad-market ETFs and smaller or more specialized ETFs (Ben-David et al. (2021a)). Focusing on the top twenty ETFs allows us to study the impact of index providers across relatively homogeneous products. Finally, investors in ETF markets may experience search frictions, which can limit investors' knowledge of product availability (Hortaçsu and Syverson (2004)). Hence the standard assumption that investors know the products in their choice set may be less likely to be satisfied if we include less popular ETFs. Restricting our sample to the top twenty ETFs alleviates this concern, as investors are likely aware of and able to compare the top ETFs.[19]

We estimate our model in several steps. In the first step, we estimate investors' preferences.

[17] Our results are not sensitive to this particular choice. In Appendix C, for example, we obtain similar estimation results using the top fifty ETFs, which in aggregate hold more than 80% of total U.S. equity ETF AUM.

[18] In Figure B.1 in Appendix B we plot the distribution of market shares of the largest twenty ETFs used in our structural estimation.

[19] An alternative modeling approach for investor demand for ETFs could be a search model along the lines of Hortaçsu and Syverson (2004). Given our main question of interest—understanding the role of large index providers' brand value and licensing fees for the equilibrium in the ETF market—a discrete choice approach with differentiated ETFs whose heterogeneity is a function of index providers, is a reasonable and simple approach.

The logit demand system in equation (4) results in the following linear regression specification:

$$\ln(s_{kt}) = -\alpha f_{kt} + \beta X_{kt} + \gamma_{ij} + \gamma_t + \xi_{kt}, \tag{13}$$

where we also include fixed effects for time (month-year) $t$ to absorb the outside option. In the estimation of (13) we are controlling for ETF sponsor and index provider time-invariant quality with fixed effects $\gamma_{ij}$. Changes in unobserved ETF quality ($\xi_{kt}$) that are correlated with contemporaneous changes in management fees ($f_{kt}$) could however be a source of bias for our estimates. For example, if an ETF expects a negative shock to its own quality $\xi_{kt}$, it may reduce the management fee $f_{kt}$ as a response. This endogeneity causes the OLS estimate of $\alpha$ to be biased downwards.

To address this endogeneity concern, we adopt an instrumental variable approach. Specifically, we instrument management fees with 1) average management fees of other ETFs in other categories of non-top ETFs offered by the same index providers; 2) number of ETFs in other categories of non-top ETFs offered by the same ETF sponsor; 3) interactions of the two. These instruments are likely to be exogenous to an ETF's own quality $\xi_{kt}$, because we explore variations in other ETFs of the same index provider or ETF sponsor. To mitigate the endogeneity concern of ETFs competing for customer demands, we especially use variations of *non-top ETFs in other categories*, so that these ETFs are less likely to directly compete with our given ETF $k$. The idea for the first instrument is that variations in management fees of other ETFs using the same index provider likely reflect common shocks to the index provider's licensing fees, which then affect the ETF's management fee $f_{kt}$.[20] The idea for the second instrument is that sponsors that offer more ETFs could potentially spread fixed operational costs across multiple ETFs, resulting in a lower average marginal cost per ETF, which passes to investors through a lower management fee $f_{kt}$.

In the second step, we estimate ETFs' cost parameters. Using the estimated investors'

[20] This idea resembles the common approach in industrial organization to use the price of a specific brand in other cities as an instrument for the price in a given city, under the assumption that correlation in prices is due to common marginal costs (Nevo (2001); Hausman (2008)).

demand parameters together with observed management fees and market shares, we can back out the marginal cost of ETF $k$ at time $t$ from (6), as follows:

$$c_{kt} = f_{kt} - \frac{1}{\hat{\alpha}(1 - s_{kt})}, \tag{14}$$

where $\hat{\alpha}$ represents the estimated coefficients on management fees; and $f_{kt}$ and $s_{kt}$ are the observed equilibrium management fees and market share of ETF $k$. We then project estimated marginal costs on index-provider, ETF-sponsor, and time fixed effects as follows:

$$c_{kt} = \gamma_i^c + \gamma_j^c + \gamma_t^c + \omega_{kt}, \tag{15}$$

where $\gamma_i^c$, $\gamma_j^c$, and $\gamma_t^c$ are index-provider, ETF-sponsor and time fixed effects; and $\omega_{kt}$ are structural error terms capturing unobservable determinants of costs.

In our last step we estimate index providers' costs $\kappa_{it}$ by inverting the first-order condition (12). This last step is the most challenging because licensing fees $\rho_{it}$ are also unobservable in most cases.[21] To address this problem, we combine our hand-collected data on licensing fees with our structural estimates for ETFs.

Most notably, we assume that the fixed effects on index providers $\gamma_i^c$ in equation (15) capture the effect of licensing fees on ETF's marginal costs. This assumption is consistent with our evidence that index licensing fees represent the single most important cost for ETFs when interacted with index providers.[22] The fixed effect estimates of index providers give only the *relative* magnitude of licensing fees. We then use our estimates in Section 3.3 to pin down the average level of $\rho_{it}$ in each period $t$. Specifically, we assume

$$\rho_{it} = \tau_t + \widehat{\gamma}_i^c, \tag{16}$$

[21] As we noted in the introduction, licensing fees are disclosed on a voluntary basis. In the standard inversion of the first-order conditions, prices are observables and, together with estimated markups, allow us to back out marginal costs. This is the approach we adopt in the second step of the estimation to infer ETFs' marginal costs using observable management fees.

[22] There may be other costs when ETFs interact with index providers, such as infrastructure costs of ETFs tracking specific indices. However, these costs are likely to be small relative to the licensing fees.

where $\widehat{\gamma}_i^c$ are the estimated fixed effects from the ETF-marginal-cost regression (15). For each month $t$, we choose parameter $\tau_t$ such that the AUM-weighted average fraction of licensing fees over management fees equals the empirical estimates reported in column (2) of Table 7.

With $\rho_{it}$, we then identify the structural parameter $\sigma$ via maximum likelihood. Most notably, for each period $t$ we construct the log-likelihood of observing ETFs choosing their index providers:

$$\mathcal{L}_t = \sum_k \sum_i \mathbb{I}_{kit} \left( \log(q_{kit}(\rho_{it})) \right), \tag{17}$$

where $\mathbb{I}_{kit}$ is an indicator variable that equals one if ETF $k$ chooses index provider $i$ in month $t$, and zero otherwise; and the probability $q_{kit}$ is given by (10).

Notice that, to calculate $q_{kit}$, we need to compute the (counterfactual) optimal profit $\pi^*_{kt}(\rho_{i't})$ of ETF $k$ for all possible index providers $i' = 1, \ldots, I_t$. Specifically, we use (4) and (6) to calculate the (counterfactual) market share $s_{kt}(\rho_{i't})$ and optimal management fee $f^*_{kt}(\rho_{i't})$, when ETF $k$ chooses index provider $i'$.[23] We then compute the optimal profits $\pi^*_{kt}(\rho_{i't})$ for each possible match with different index providers and construct the index provider choice probabilities given by (10).

Finally, using the estimated structural parameters $\alpha$ and $\sigma$, index providers' choice probabilities $q_{kit}(\rho_{it})$, ETFs' market shares $s_{kt}(\rho_{it})$, and calibrated licensing fees $\rho_{it}$, we can compute index providers' markup and back out unobservable marginal costs $\kappa_{it}$ using equation (12).

### 5.2 Results

In this section, we report the results obtained by estimating the structural model. Table 8 shows the results for investor demand parameters (columns (1) and (2)) and ETFs' cost parameters (columns (3) and (4)).

[23] Some index provider×ETF sponsor interacted fixed effects $\gamma_{ij}$ cannot be estimated from regression (13), because the corresponding index providers and ETF sponsors do not match with each other (see Table 3). To address this issue, we regress the observed interacted fixed effects on separate fixed effects ($\gamma_{ij} = \gamma_i + \gamma_j + \psi_{ij}$), and use this regression to estimate unobserved interacted fixed effects, which are used to calculate counterfactual market shares and management fees.

For column (1), we shows the OLS estimates of equation (13) with time fixed effects and interacted fixed effects for ETF sponsors and index providers. As expected, we find that higher fees are associated with a lower market share. The coefficient is now highly significant and implies an elasticity of about 2.2. Past returns enter negatively, but the coefficient is always estimated very imprecisely. While the interacted ETF sponsors and index-provider fixed effects capture all time-invariant characteristics that can affect demand, time-varying demand shocks to specific ETFs that are correlated with management fees could still bias our estimates.

Hence, in column (2) of Table 8, we report the IV estimates of equation (13). Our first stage is strong, with an F statistics on the excluded instruments about 29. Once we instrument for management fees, the coefficient increases in (absolute) magnitude, consistent with a downward bias in the OLS specification. The average elasticity to management fees is about 3.0.

Columns (3) and (4) of Table 8 show the estimates from equation (15). The dependent variable is the marginal costs at the ETF-month level. We find an average marginal cost of about 7.4 basis points. For column (3), we control only for time fixed effects to absorb changes in costs during our sample period. We find that, on average, CRSP (the excluded dummy) is associated with the lowest-cost ETFs (which are offered exclusively by Vanguard), followed by the S&P Dow Jones. This is consistent with the evidence that CRSP is the latest entry into the index providing market, with an aim to provide low-cost ETFs to investors.[24]

The estimates for index-provider fixed effects in column (3) may be capturing differential marginal costs of ETF sponsors. To isolate the effect of index providers with the aim of capturing the impact of licensing fees, for column (4), we add ETF-sponsor fixed effects. Now S&P Dow Jones is associated with the lowest ETF marginal costs, followed by CRSP, while FTSE Russell and MSCI involve the highest and second-highest marginal costs, respectively. Given that marginal cost $\widehat{\gamma}_i^c$ reflects the relative level of licensing fees $\rho_i$ as in (16), this ranking

[24] See https://www.cnbc.com/2012/10/04/vanguard-move-highlights-littleknown-index-costs.html.

**Table 8**

**Structural parameters**

This table reports the structural parameters for investor demand from equation (13) and ETF marginal costs from equation (15). In columns (1) and (2) the dependent variable is the (log) market share. For columns (3) and (4) the dependent variable is ETF marginal costs in basis points. Past return is the average of monthly returns in the past 12 months. The excluded dummy for index providers is for CRSP. All standard errors are clustered at the ETF sponsor level.

| | Investors demand parameters<br>Dep Var: Market share (log) | | ETF cost parameters<br>Dep Var: Marginal costs (bps) | |
|---|---|---|---|---|
| | (1)<br>OLS | (2)<br>IV | (3)<br>OLS | (4)<br>OLS |
| Management fees (bps) | -0.197*** | -0.261*** | | |
| | (0.062) | (0.030) | | |
| Past return (12 months) | -6.672 | -4.840 | | |
| | (14.238) | (3.575) | | |
| FTSE Russell | | | 10.351*** | 5.071*** |
| | | | (2.872) | (1.393) |
| MSCI | | | 5.949*** | 3.977*** |
| | | | (1.180) | (0.665) |
| NASDAQ | | | 7.534* | 2.197*** |
| | | | (3.985) | (0.004) |
| S&P Dow Jones | | | 3.010*** | -3.512*** |
| | | | (1.067) | (0.033) |
| FE year-month | Yes | Yes | Yes | Yes |
| FE ETF sponsor × IP | Yes | Yes | No | No |
| FE ETF sponsor | | | No | Yes |
| Elasticity to fees | 2.24 | 2.97 | | |
| First-stage F stat | | 28.72 | | |
| Mean dep. var. | -3.92 | -3.92 | 7.37 | 7.37 |
| SD dep. var. | 0.82 | 0.82 | 5.54 | 5.54 |
| $R^2$ | 0.66 | 0.63 | 0.52 | 0.81 |
| Observations | 2,100 | 2,100 | 2,100 | 2,100 |

also applies to licensing fees. That is, S&P Dow Jones has the lowest licensing fees, CRSP the second-lowest, and FTSE Russell and MSCI charge the highest licensing fees.[25]

Table 9 presents the main estimation results. We report three cross sections of the model, for December 2012, 2016, and 2019, in turn. The AUM of the top twenty ETFs increases significantly over time, consistent with the aggregate trend documented in Figure 1. By

[25] This finding is consistent with the interview of a former asset manager in New York, "MSCI is famous for being expensive—not because they have better data or indices, but because they are the brand that is most used in the world." (Petry, Fichtner, and Heemskerk (2019))

December 2019, the top twenty ETFs have almost \$1.5 trillion in AUM, up more than 300% from December 2012. We also report the structural parameter $\sigma$, which affects the competition among index providers. The parameter is stable across years at around 0.53. By (10), we have

$$\sigma = \frac{\ln(q_{ki}(\rho_i)) - \ln(q_{ki'}(\rho_{i'}))}{\ln(\pi_{ki}(\rho_i)) - \ln(\pi_{ki'}(\rho_{i'}))}. \tag{18}$$

Hence, if an index provider $i$ can offer a 1% higher profit for ETFs than another index provider $i'$, the probability that the ETF chooses index provider $i$ is only 0.53% higher than the probability that the ETF chooses index provider $i'$. If index providers were perfectly competitive, index provider $i$ should be always be chosen over $i'$ ($\sigma = \infty$). Such a low elasticity implies very limited substitutability across index providers, which is consistent with the persistence of indexing relationships and significant market power of index providers.

In Panel B of Table 9, we report results for the ETF variables. Over time, management fees declined from on average 12.7 basis points in 2012 to 9.3 basis points by the end of 2019. The decreases in management fees are driven mostly by marginal costs, which drop from 8.8 basis points to 5.4 basis points. The markup charged by ETFs, on the other hand, is rather stable, at around 3.9 basis points over time. As a result, the Lerner index (=markup/management fees) increased over time from about 31% in 2012 to 42% in 2019.

In Panel C, we report results for index provider variables. The licensing fees charged by index providers decreased over time, from 5.2 basis points in 2012 to 4.4 basis points in 2019. While licensing fees are indeed dropping over time, the decrease is lesser in magnitude than that of management fees, in both percentages and absolute terms. The drops in licensing fees are also driven mostly by marginal costs, which decrease from 2.3 basis points to 1.6 basis points. The reduction in marginal costs over time suggests that index providers have increasing expertise in providing cheap indexes. The markup charged by index providers, on the other hand, is rather stable, at around 2.8 basis points.

According to our calibration, index providers experience very high margins, with an average Lerner index of around 60%. More importantly, although the marginal costs of index

**Table 9**
**Estimation results**

This table reports several variables of interest for December 2012, 2016, and 2019. The Lerner index is defined as the difference between price and marginal cost divided by price. We calculate management fees, licensing fees, marginal costs, markups, and the Lerner index for each ETF, and then report average across ETFs. The last two columns report differences in levels and in percentages between December 2019 and 2012.

| | 2012 | 2016 | 2019 | Change 2019-2012 | Change (%) 2019-2012 |
|---|---|---|---|---|---|
| **Panel A: Data & parameters** | | | | | |
| AUM of all ETFs ($ billions) | 598 | 1,464 | 2,514 | 1,916 | 320.4 |
| AUM of top twenty ETFs ($ billions) | 356 | 842 | 1,464 | 1,108 | 311.6 |
| $\sigma$ | 0.54 | 0.53 | 0.53 | -0.01 | -1.1 |
| **Panel B: ETFs** | | | | | |
| Management fees (bps) | 12.7 | 10.4 | 9.3 | -3.4 | -26.9 |
| Marginal costs (bps) | 8.8 | 6.5 | 5.4 | -3.4 | -38.8 |
| Markups (bps) | 3.9 | 3.9 | 3.9 | -0.0 | -0.1 |
| Lerner index (%) | 30.7 | 37.6 | 42.0 | 11.3 | 36.7 |
| **Panel C: Index providers** | | | | | |
| Licensing fees (bps) | 5.2 | 4.7 | 4.4 | -0.7 | -14.2 |
| Marginal costs (bps) | 2.3 | 1.9 | 1.6 | -0.7 | -29.2 |
| Markups (bps) | 2.9 | 2.8 | 2.8 | -0.1 | -2.1 |
| Lerner index (%) | 55.6 | 60.2 | 63.4 | 7.8 | 14.0 |

providers have decreased significantly over time, their markup is rather stable. This leads to an increasing Lerner index over time, reaching 63% by the end of 2019. Our estimate is aligned with Financial Times (2019) that estimates the profit margin of the top three index providers to be about 65% as of 2019, and it is also consistent with the increasing trend in profit margin for index provision as reported by MSCI (2020). As a result, index providers have benefited enormously from the significantly increased AUM in ETFs.

### 5.3 Counterfactual Analysis

Our estimation results show that index providers wield high market power and a large fraction of index licensing fees are markups. We now consider a counterfactual scenario with perfectly competitive index providers and study the equilibrium effect on the ETF market.

**Table 10**

## Increasing competition among index providers

This table reports several variables of interest for December 2019 in the baseline case and the perfectly competitive index provider case. The Lerner index is defined as the difference between the price and marginal costs divided by the price. We calculate management fees, licensing fees, marginal costs, markups, and the Lerner index for each ETF, and then report average across ETFs. In the last two columns we report the differences in levels and percentages between the perfectly competitive scenario and the baseline scenario.

| Year 2019 | Baseline | Competitive IP | Change | Change (%) |
|---|---|---|---|---|
| **Panel A: ETFs** | | | | |
| Management fees (bps) | 9.3 | 6.5 | -2.8 | -29.8 |
| Marginal costs (bps) | 5.4 | 2.6 | -2.8 | -52.0 |
| Markups (bps) | 3.9 | 3.9 | 0.0 | 0.9 |
| Lerner index (%) | 42.0 | 60.3 | 18.3 | 43.7 |
| **Panel B: Index providers** | | | | |
| Licensing fees (bps) | 4.4 | 1.6 | -2.8 | -63.4 |
| Marginal costs (bps) | 1.6 | 1.6 | 0.0 | 0.0 |
| Markups (bps) | 2.8 | 0.0 | -2.8 | -100.0 |
| Lerner index (%) | 63.4 | 0.0 | -63.4 | -100.0 |

We implement the counterfactual scenario as follows. We assume that each index provider $i$ charges a licensing fee that is equal to the marginal cost (i.e., $\rho_{it} = \kappa_{it}$). This scenario corresponds to setting $\sigma = \infty$ and provides quantitative estimates of the upper bound of potential gains from significant improvement in competitiveness in the index providing markets. With new licensing fee $\kappa_{it}$, we solve for the equilibrium management fee $f_k^*(\kappa_{it})$ and market share $s_k^*(\kappa_{it})$ for all ETFs $k = 1, \cdots, K$ jointly, using (3) and (6). In the counterfactual scenario, each ETF's choice of index providers is kept the same as the observed equilibrium outcomes. We preserve equilibrium matching to focus on the effects of reducing licensing fees on management fees.

Table 10 shows the results for December 2019.[26] Comparing the estimates in the baseline with those in the perfectly competitive index provider case, we see that increasing the competitiveness of index providers reduces ETFs' marginal costs by about 2.8 basis points, which corresponds to a 52% decline. The lower marginal costs are passed on to investors, as

[26]In unreported tables, we find similar counterfactual results when using the snapshot of December 2012 and December 2016.

management fees also decline by 2.8 basis points, from 9.3 to 6.5 basis points. This represents approximately a 30% decline. Because of decreasing marginal costs and stable markups, the Lerner index increases by about twenty percentage points, from 42% to 60%. Panel B of Table 10 shows the outcomes for index providers. In a perfectly competitive market for index providers, licensing fees are equal to marginal costs, and profits are zero. The licensing fees decrease from 4.4 basis points to 1.6 basis points, which corresponds to a 63% decline.

## 6 Conclusion

Most ETFs passively replicate the performance of an index that is constructed and maintained by an index provider. In this paper, we provide the first analysis of the competition structure between index providers and ETF sponsors and the consequences for ETF management fees charged to investors. We find that the index provider market is highly concentrated and dominated by a few large players and that about one-third of ETF management fees are paid as index licensing fees to index providers. Moreover, we find that ETF investors care about the identities of index providers, although the identities of index providers explain little of the variations in ETF returns. Using a structural model that incorporates the two-tiered competition between index providers for ETFs and between ETFs for investors, we show that index providers wield very strong market power and about 60% of index licensing fees are markups charged by index providers. Eliminating the market power of index providers could reduce ETF management fees by 30%. The potential policy implications of our results include the need to require disclosure of licensing fees for all ETFs and the recommendation to promote competition between index providers.

## References


Agarwal, Vikas, Paul Hanouna, Rabih Moussawi, and Christof Stahel, 2019, Do ETFs increase the commonality in liquidity of underlying stocks?, Working Paper, Georgia State University.

Akey, Pat, Adriana Robertson, and Mikhail Simutin, 2021, Closet active management of passive funds, Working paper, University of Toronto.

Azar, José, Martin C Schmalz, and Isabel Tecu, 2018, Anticompetitive effects of common ownership, *The Journal of Finance* 73, 1513–1565.

Bao, Weining, and Jian Ni, 2017, Could good intentions backfire? an empirical analysis of the bank deposit insurance, *Marketing Science* 36, 301–319.

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim, and Rabih Moussawi, 2021a, Competition for attention in the ETF space, Technical report, National Bureau of Economic Research.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do ETFs increase volatility? *Journal of Finance* 73, 2471–2535.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2021b, What do mutual fund investors really care about?, *Review of Financial Studies, Forthcoming* .

Benetton, Matteo, 2018, Leverage regulation and market structure: An empirical model of the UK mortgage market, *Available at SSRN 3247956* .

BlackRock, 2020, the 2019 annual report.

Box, Travis, Ryan Davis, Richard B. Evans, and Andrew A. Lynch, 2020, Intraday arbitrage between ETFs and their underlying portfolios, *Journal of Financial Economics* Forthcoming.

Brown, David, Scott Cederburg, and Mitch Towner, 2021, Dominated ETFs, Working paper, University of Arizona.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2020, Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy, Technical report, National Bureau of Economic Research.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Da, Zhi, and Sophie Shive, 2018, Exchange-traded funds and asset return correlations, *European Financial Management* 24, 136–168.

Dannhauser, Caitlin D, and Jeffrey Pontiff, 2019, Flow, Working paper, Boston College.

deHaan, Ed, Yang Song, Chloe Xie, and Christina Zhu, 2021, Obfuscation in mutual funds, *Journal of Accounting and Economics, Forthcoming* .

Eaton, Jonathan, and Samuel Kortum, 2002, Technology, geography, and trade, *Econometrica* 70, 1741–1779.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, 2017, Deposit competition and financial fragility: Evidence from the us banking sector, *American Economic Review* 107, 169–216.

Egan, Mark L, Alexander MacKay, and Hanbin Yang, 2020, Recovering investor expectations from demand for index funds, Technical report, National Bureau of Economic Research.

Financial Times, 2019, Index companies to feel the chill of fund managers' price war.

Glosten, Lawrence, Suresh Nallareddy, and Yuan Zou, 2020, ETF activity and informational efficiency of underlying securities, *Management Science,* Forthcoming.

Hausman, Jerry A, 2008, 5 valuation of new goods under perfect and imperfect competition, *The Economics of New Goods* 58, 209.

Horizontal Merger Guidelines, 2010, The U.S. Department of Justice and the Federal Trade Commission.

Hortaçsu, Ali, and Chad Syverson, 2004, Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds, *Quarterly Journal of Economics* 119, 403–456.

Huang, Shiyang, Maureen O'Hara, and Zhuo Zhong, 2021, Innovation and informed trading: Evidence from industry ETFs, *Review of Financial Studies* Forthcoming.

Huang, Shiyang, Yang Song, and Hong Xiang, 2020, The smart beta mirage, Working paper, The University of Hong Kong and University of Washington.

Israeli, Doron, Charles M.C. Lee, and Suhas A. Sridharan, 2017, Is there a dark side to exchange-traded funds? An information perspective, *Review of Accounting Studies* 22, 1048–1083.

Jiang, Erica Xuewei, 2020, Financing competitors: Shadow banks' funding and mortgage market competition, Working paper, University of Southern California.

Khomyn, Marta, Tālis J Putniņš, and Marius Zoican, 2020, The value of ETF liquidity, Working paper, University of Technology Sydney and University of Toronto.

Koijen, Ralph SJ, and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Kostovetsky, Leonard, and Jerold B Warner, 2021, The market for fund benchmarks: Evidence from ETFs, *Available at SSRN 3804002* .

Mahoney, Paul G, and Adriana Robertson, 2021, Advisers by another name, *Virginia Law and Economics Research Paper* .

MSCI, 2020, the 2019 annual report. New York: MSCI.

Nevo, Aviv, 2001, Measuring market power in the ready-to-eat cereal industry, *Econometrica* 69, 307–342.

Pan, Kevin, and Yao Zeng, 2019, ETF arbitrage under liquidity mismatch, Working paper, Goldman Sachs and University of Washington.

Petry, Johannes, Jan Fichtner, and Eelke Heemskerk, 2019, Steering capital: the growing private authority of index providers in the age of passive asset management, *Review of international political economy* 1–25.

Philippon, Thomas, 2019, *The great reversal* (Harvard University Press).

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

## Appendices

## A Derivations of Structural Model

In this appendix, we provide detailed derivations of formulas that we omit in the main text for the structural model.

We derive equation (12). From (4), we have

$$\begin{aligned}\frac{\partial s_k(f_k)}{\partial f_k} =& \frac{1+\sum_{k'} e^{-\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}} - e^{-\alpha f_k+\beta X_k+\gamma_{ij}+\xi_k}}{(1+\sum_{k'} e^{\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}})^2} \frac{\partial e^{-\alpha f_k+\beta X_k+\gamma_{ij}+\xi_k}}{\partial f_k} \\ =& \frac{1+\sum_{k'} e^{-\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}} - e^{-\alpha f_k+\beta X_k+\gamma_{ij}+\xi_k}}{(1+\sum_{k'} e^{\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}})^2} e^{-\alpha f_k+\beta X_k+\gamma_{ij}+\xi_k}(-\alpha) \\ =& -\alpha s_k(f_k)(1-s_k(f_k)). \end{aligned} \tag{A.1}$$

In addition, we have

$$\begin{aligned}\frac{\partial s_k^*(\rho_i)}{\partial \rho_i} =& \frac{1+\sum_{k'} e^{-\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}} - e^{-\alpha f_k^*+\beta X_k+\gamma_{ij}+\xi_k}}{(1+\sum_{k'} e^{\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}})^2} \frac{\partial e^{-\alpha f_k^*+\beta X_k+\gamma_{ij}+\xi_k}}{\partial \rho_i} \\ =& \frac{1+\sum_{k'} e^{-\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}} - e^{-\alpha f_k^*+\beta X_k+\gamma_{ij}+\xi_k}}{(1+\sum_{k'} e^{\alpha f_{k'}+\beta X_{k'}+\gamma_{i'j'}+\xi_{k'}})^2} e^{-\alpha f_k^*+\beta X_k+\gamma_{ij}+\xi_k}(-\alpha)\frac{\partial f_k^*(\rho_i)}{\partial \rho_i} \\ =& -\alpha s_k^*(\rho_i)(1-s_k^*(\rho_i))\left(1+\frac{1}{\alpha(1-s_k^*(\rho_i))^2}\frac{\partial s_k^*(\rho_i)}{\partial \rho_i}\right), \end{aligned} \tag{A.2}$$

where the last equality uses (6). This equation implies

$$\frac{\partial s_k^*(\rho_i)}{\partial \rho_i} = -\alpha s_k^*(\rho_i)(1-s_k^*(\rho_i))^2. \tag{A.3}$$

Moreover, from (10) and by the envelope theorem of (5), we have

$$
\begin{aligned}
\frac{\partial q_{ki}(\rho_i)}{\partial \rho_i} =& \frac{\sum_{i'} \pi_{ki'}(\rho_{i'})^\sigma - \pi_{ki}(\rho_i)^\sigma}{(\sum_{i'} \pi_{ki'}(\rho_{i'})^\sigma)^2} \sigma \pi_{ki}(\rho_i)^{\sigma-1} \frac{\partial \pi_{ki}(\rho_i)}{\partial \rho_i} \\
=& \frac{\sum_{i'} \pi_{ki'}(\rho_{i'})^\sigma - \pi_{ki}(\rho_i)^\sigma}{(\sum_{i'} \pi_{ki'}(\rho_{i'})^\sigma)^2} \sigma \pi_{ki}(\rho_i)^{\sigma-1} (-L) s_k^*(\rho_i) \\
=& - \frac{L\sigma q_{ki}(\rho_i)(1 - q_{ki}(\rho_i)) s_k^*(\rho_i)}{\pi_{ki}(\rho_i)} \\
=& - \frac{\sigma q_{ki}(\rho_i)(1 - q_{ki}(\rho_i))}{f_k^*(\rho_i) - \rho_i - \tilde{c}_k} \\
=& - \alpha \sigma q_{ki}(\rho_i)(1 - q_{ki}(\rho_i))(1 - s_k^*(\rho_i)),
\end{aligned} \tag{A.4}
$$

where the last equality uses (6). Therefore, we have

$$
\begin{aligned}
&\frac{\partial q_{ki}(\rho_i)}{\partial \rho_i} s_k^*(\rho_i) + \frac{\partial s_k^*(\rho_i)}{\partial \rho_i} q_{ki}(\rho_i) \\
=& - \alpha q_{ki}(\rho_i) s_k^*(\rho_i)(1 - s_k^*(\rho_i)) \left(\sigma(1 - q_{ki}(\rho_i)) + 1 - s_k^*(\rho_i)\right).
\end{aligned} \tag{A.5}
$$

Plugging this equation into the F.O.C. of (11) gives equation (12).

## B Additional Empirical Results

In Table B.1, we replicate Table 3 but using total revenue (=AUM×management fee). The results are similar. In Table B.2 and Table B.3, we replicate Table 3 using the snapshot in December 2013 and in December 2016, respectively. The results are again similar, suggesting the stability of matching between index providers and ETF sponsors over time.

In Figure B.1 we report the market share of top twenty ETFs as of December 2019, which are used in the structural estimation of Section 5. The combined market share of top twenty ETFs is about 60%.

**Table B.1**
**Matching between index providers and ETF sponsors: total revenue**

In this table, we report matching between index providers and ETF sponsors. We use "others" to represent all index providers or ETF sponsors besides the top fives. Panel A reports the distribution of total revenue (=AUM×management fee) across various index providers from a given ETF sponsor's perspective. Panel B reports the distribution of total revenue across various ETF sponsors from a given index provider's perspective. We highlight cells that are over 50%. The sample period is December 2019.

**Panel A: From ETF sponsors' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | **55.0%** | 0.0% | 39.6% | 1.2% | 1.5% | 2.7% |
| Vanguard | 10.6% | **54.3%** | 4.7% | 19.8% | 10.7% | 0.0% |
| State Street | **100.0%** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Invesco | 22.6% | 0.0% | 4.8% | 0.0% | **67.2%** | 5.4% |
| Schwab | **99.0%** | 0.0% | 1.0% | 0.0% | 0.0% | 0.0% |
| Others | 7.6% | 0.0% | 0.0% | 1.1% | 17.6% | **73.7%** |

**Panel B: From index providers' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | 35.8% | 0.0% | **82.3%** | 48.3% | 10.5% | 8.4% |
| Vanguard | 3.5% | **100.0%** | 4.7% | 45.7% | 5.9% | 0.0% |
| State Street | 44.8% | 0.0% | 0.5% | 0.4% | 0.0% | 1.5% |
| Invesco | 9.1% | 0.0% | 5.3% | 0.0% | **51.7%** | 7.7% |
| Schwab | 2.2% | 0.0% | 4.4% | 0.0% | 0.0% | 0.1% |
| Others | 4.7% | 0.0% | 2.8% | 5.6% | 32.0% | **82.3%** |

## C Robustness Checks Using the Top Fifty ETFs

In this section, we present the structural estimation results using the top fifty ETFs as of December 2019, while taking remaining ETFs as an outside option.

In Table C.1, we present the results using the top fifty ETFs for investor demand parameters (columns (1) and (2)) and ETFs' cost parameters (columns (3) and (4)). The parameter estimates are similar to those using the top twenty ETFs, as reported in Table 8.

Table C.2 presents the main estimation results using the top fifty ETFs. Similar to the results based on the top twenty ETFs, ETF management fees declined from on average 16.4 basis points in 2012 to 12.9 basis points by the end of 2019. The decreases in management fees are driven mostly by marginal costs, which drop from 13.7 basis points to 10.2 basis points. The markup charged by ETFs, on the other hand, is rather stable, at around 2.7 basis points over time. As a result, the Lerner index (=markup/management fees) increased over time

**Table B.2**

## Matching between index providers and ETF sponsors: December 2013

In this table, we report matching between index providers and ETF sponsors. We use "others" to represent all index providers or ETF sponsors besides the top fives. Panel A reports the distribution of AUM across various index providers from a given ETF sponsor's perspective. Panel B reports the distribution of AUM across various ETF sponsors from a given index provider's perspective. We highlight cells that are over 50%. The sample period is December 2013.

**Panel A: From ETF sponsors' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | **57.1%** | 0.0% | 29.3% | 9.3% | 1.2% | 3.1% |
| Vanguard | 21.1% | **52.2%** | 5.8% | 14.9% | 6.0% | 0.0% |
| State Street | **97.7%** | 0.0% | 0.2% | 0.1% | 0.0% | 2.0% |
| Invesco | 33.2% | 0.0% | 5.1% | 0.0% | **58.1%** | 3.6% |
| Schwab | **88.4%** | 0.0% | 10.7% | 0.0% | 0.0% | 1.0% |
| Others | 11.3% | 0.0% | 3.9% | 10.1% | 18.9% | **55.8%** |

**Panel B: From index providers' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | 31.5% | 0.0% | **90.6%** | 9.0% | 5.7% | 23.6% |
| Vanguard | 3.6% | **100.0%** | 6.5% | **90.2%** | 23.8% | 0.0% |
| State Street | **58.9%** | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| Invesco | 3.3% | 0.0% | 2.8% | 0.0% | **64.0%** | 12.2% |
| Schwab | 2.3% | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% |
| Others | 0.4% | 0.0% | 0.0% | 0.8% | 6.5% | **64.1%** |

from about 17% in 2012 to 21% in 2019. In panel B, we see that the licensing fees charged by index providers decreased over time, from 5.0 basis points in 2012 to 4.0 basis points in 2019. The markup charged by index providers remained relatively stable at about 3.1 basis points. As a result, the Lerner index of index providers increased from 63% in 2012 to about 79% in 2019, with an average about 73%.

Table C.3 presents the counterfactual analysis results using the top fifty ETFs for December 2019. Similar to the results based on the top twenty ETFs, increasing the competitiveness of index providers reduces ETFs' marginal costs by about 3.1 basis points, which corresponds to a 31% decline. The lower marginal costs are passed on to investors, as management fees also decline by 3.1 basis points, from 12.9 to 9.8 basis points. This represents approximately a 24% decline.

**Table B.3**

## Matching between index providers and ETF sponsors: December 2016

In this table, we report matching between index providers and ETF sponsors. We use "others" to represent all index providers or ETF sponsors besides the top fives. Panel A reports the distribution of AUM across various index providers from a given ETF sponsor's perspective. Panel B reports the distribution of AUM across various ETF sponsors from a given index provider's perspective. We highlight cells that are over 50%. The sample period is December 2016.

**Panel A: From ETF sponsors' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | **53.7%** | 0.0% | 37.1% | 4.4% | 1.7% | 3.0% |
| Vanguard | 17.6% | **51.4%** | 5.8% | 19.2% | 6.0% | 0.0% |
| State Street | **99.6%** | 0.0% | 0.3% | 0.0% | 0.0% | 0.1% |
| Invesco | 36.1% | 0.0% | 6.9% | 0.0% | **52.0%** | 5.0% |
| Schwab | **91.4%** | 0.0% | 8.6% | 0.0% | 0.0% | 0.0% |
| Others | 10.7% | 0.0% | 1.5% | 5.7% | 20.4% | **61.8%** |

**Panel B: From index providers' perspective**

| | S&P Dow Jones | CRSP | FTSE Russell | MSCI | NASDAQ | Others |
|---|---|---|---|---|---|---|
| iShares | 32.0% | 0.0% | **84.2%** | 21.9% | 9.0% | 24.0% |
| Vanguard | 8.1% | **100.0%** | 10.1% | **74.2%** | 23.9% | 0.0% |
| State Street | **50.4%** | 0.0% | 0.6% | 0.0% | 0.0% | 0.9% |
| Invesco | 4.2% | 0.0% | 3.1% | 0.0% | **52.7%** | 7.8% |
| Schwab | 4.5% | 0.0% | 1.6% | 0.0% | 0.0% | 0.0% |
| Others | 0.9% | 0.0% | 0.5% | 3.9% | 14.5% | **67.3%** |

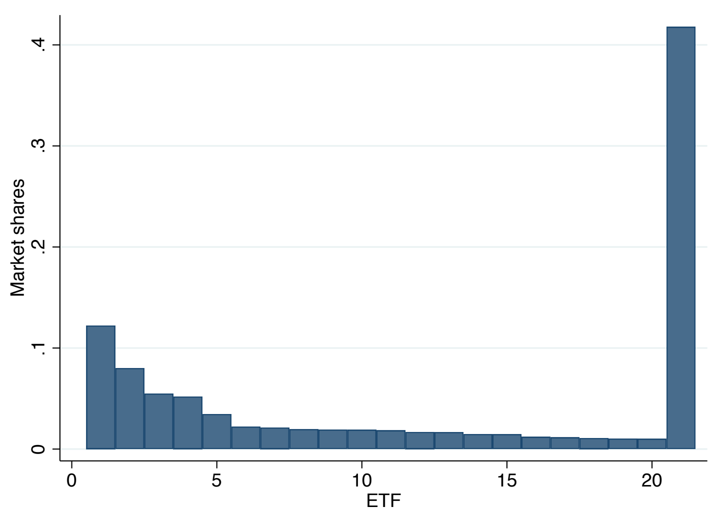



**Figure B.1. Market share of top twenty ETFs.** In this figure, we show the market share of top twenty ETFs as of December 2019, which are used in the structural estimation of Section 5. The x-axis shows the market share rank of each ETF. The combined market share of ETFs outside top twenty is about 42%, as also shown in the rightmost bar of the figure.

**Table C.1**

## Structural parameters: top fifty ETFs

This table reports the structural parameters for investor demand from equation (13) and ETF marginal costs from equation (15). We use the top fifty ETFs as of December 2019. In columns (1) and (2) the dependent variable is the (log) market share. For columns (3) and (4) the dependent variable is ETF marginal costs in basis points. Past return is the average of monthly returns in the past 12 months. The excluded dummy for index providers is for CRSP. All standard errors are clustered at the ETF sponsor level.

| | Investors demand parameters<br>Dep Var: Market share (log) | | ETF cost parameters<br>Dep Var: Marginal costs (bps) | |
|---|---|---|---|---|
| | (1)<br>OLS | (2)<br>IV | (3)<br>OLS | (4)<br>OLS |
| Management fees (bps) | -0.028*** | -0.379*** | | |
| | (0.007) | (0.054) | | |
| Past return (12 months) | 9.310*** | 2.078 | | |
| | (3.084) | (5.753) | | |
| FTSE Russell | | | 12.005*** | 5.279*** |
| | | | (1.613) | (1.960) |
| MSCI | | | 5.701*** | 3.311*** |
| | | | (1.294) | (0.642) |
| NASDAQ | | | 7.070* | -2.275 |
| | | | (4.000) | (2.926) |
| S&P Dow Jones | | | 8.958*** | 1.371 |
| | | | (1.773) | (1.733) |
| FE year-month | Yes | Yes | Yes | Yes |
| FE ETF sponsor × IP | Yes | Yes | No | No |
| FE ETF sponsor | | | No | Yes |
| Elasticity to fees | 0.41 | 5.63 | | |
| First-stage F stat | | 15.50 | | |
| Mean dep. var. | -4.56 | -4.56 | 10.85 | 10.85 |
| SD dep. var. | 0.88 | 0.88 | 8.39 | 8.39 |
| $R^2$ | 0.28 | -4.74 | 0.24 | 0.52 |
| Observations | 5,096 | 5,096 | 5,096 | 5,096 |

**Table C.2**

## Estimation results: based on the top fifty ETFs

This table reports several variables of interest for December 2012, 2016, and 2019. We use thetop fifty ETFs as of December 2019. The Lerner index is defined as the difference between price and marginal cost divided by price. We calculate management fees, licensing fees, marginal costs, markups, and the Lerner index for each ETF, and then report average across ETFs. The last two columns report differences in levels and in percentages between December 2019 and 2012.

| | 2012 | 2016 | 2019 | Change 2019-2012 | Change (%) 2019-2012 |
|---|---|---|---|---|---|
| **Panel A: ETFs** | | | | | |
| Management fees (bps) | 16.4 | 14.1 | 12.9 | -3.5 | -21.2 |
| Marginal costs (bps) | 13.7 | 11.4 | 10.2 | -3.5 | -25.4 |
| Markups (bps) | 2.7 | 2.7 | 2.7 | 0.0 | 0.1 |
| Lerner index (%) | 16.4 | 19.0 | 20.8 | 4.4 | 27.0 |
| **Panel B: Index providers** | | | | | |
| Licensing fees (bps) | 5.0 | 4.3 | 4.0 | -1.1 | -20.9 |
| Marginal costs (bps) | 1.9 | 1.0 | 0.8 | -1.1 | -55.9 |
| Markups (bps) | 3.1 | 3.3 | 3.1 | -0.0 | -0.0 |
| Lerner index (%) | 62.6 | 76.1 | 79.2 | 16.5 | 26.4 |

**Table C.3**

## Increasing competition among index providers: based on the top fifty ETFs

This table reports several variables of interest for December 2019 in the baseline case and the perfectly competitive index provider case. We use the top fifty ETFs as of December 2019. The Lerner index is defined as the difference between the price and marginal costs divided by the price. We calculate management fees, licensing fees, marginal costs, markups, and the Lerner index for each ETF, and then report average across ETFs. In the last two columns we report the differences in levels and percentages between the perfectly competitive scenario and the baseline scenario.

| Year 2019 | Baseline | Competitive IP | Change | Change (%) |
|---|---|---|---|---|
| **Panel A: ETFs** | | | | |
| Management fees (bps) | 12.9 | 9.8 | -3.1 | -24.3 |
| Marginal costs (bps) | 10.2 | 7.1 | -3.1 | -30.8 |
| Markups (bps) | 2.7 | 2.7 | 0.0 | 0.5 |
| Lerner index (%) | 20.8 | 27.6 | 6.8 | 32.7 |
| **Panel B: Index providers** | | | | |
| Licensing fees (bps) | 4.0 | 0.8 | -3.1 | -79.2 |
| Marginal costs (bps) | 0.8 | 0.8 | 0.0 | 0.0 |
| Markups (bps) | 3.1 | 0.0 | -3.1 | -100.0 |
| Lerner index (%) | 79.2 | 0.0 | -79.2 | -100.0 |