The Review of Financial Studies

# The Value of ETF Liquidity

**Marta Khomyn**
The University of Adelaide, Australia

**Tālis Putniņš**
Stockholm School of Economics in Riga, Latvia, and University of Technology Sydney, and Digital Finance CRC, Australia

**Marius Zoican**
Haskayne Business School, University of Calgary, Canada

We analyze how ETFs compete. Drawing on a new model and empirical analysis, we show that ETF secondary market liquidity plays a key role in determining fees. More liquid ETFs for a given index charge higher fees and attract short-horizon investors who are more sensitive to liquidity than to fees. Higher turnover from these investors sustains the ETF's high liquidity, allowing the ETF to extract a rent through its fee, and creating a first-mover advantage. Liquidity segmentation through clientele effects generates welfare losses. Our findings resolve the apparent paradox that higher-fee ETFs not only survive but also flourish in equilibrium. (*JEL* G11, G12, G14)

Received: August 1, 2022; Editorial decision: June 20, 2024
Editor: Itay Goldstein



---

We have greatly benefited from discussion with Wolfgang Bessler, David Cimon, Nicolae Gârleanu, Virginia Gianinazzi, Roger Edelen, Vyacheslav Fos, Raymond Kan, Pete Kyle, Nicolas Iñostroza, Charles Martineau, Shane Miller, Maureen O'Hara, Michael Pagano, Uday Rajan, Tom Smith, Damien Sherman, Elvira Sojli, Avanidhar Subrahmanyam, and Damien Wallace and seminar and conference participants at Rotman School of Management, Johns Hopkins Carey Business School, The Microstructure Exchange, 47th EFA Annual Meeting, 2020 NFA Conference, 4th SAFE Market Microstructure Conference, FIRS 2019 Conference, 8th Behavioral Finance and Capital Markets Conference, 31st Australasian Finance and Banking Conference, 9th ERMAS Conference, and 8th FIRN Annual Conference for helpful comments. Marta Khomyn gratefully acknowledges funding from Capital Markets CRC Limited and was a visiting researcher at Chi-X Australia during the time spent on this study. Tālis J. Putniņš gratefully acknowledges funding from the Australian Research Council [ARC DP200101445]. Marius Zoican gratefully acknowledges funding from the Social Sciences and Humanities Research Council [Insight Development Grant 430-2020-00014]. Marius Zoican acknowledges financial support from Quantitative Management Research Initiative (QMI) under the aegis of the Fondation du Risque, a joint initiative by Université Paris-Dauphine, l'École Nationale de la Statistique et de l'Administration ParisTech (ENSAE), and LFIS, France. The paper was completed while Marius Zoican was affiliated with the University of Toronto. Send correspondence to Marius Zoican, marius.zoican@haskayne.ucalgary.ca.

*The Review of Financial Studies* 37 (2024) 3092–3148

https://doi.org/10.1093/rfs/hhae041 Advance Access publication August 21, 2024

> 'Now you can trade the S&P 500 Index in real time' was the slogan in the newspaper for the first ETF. What kind of nut would do that?
>
> —John C. Bogle, founder of the Vanguard Group[1]

Exchange-traded funds (ETFs) have rapidly grown to $6 trillion in assets, or about one-quarter of global equity market capitalization. Some ETFs compete by selecting increasingly specialized, niche indices not offered by other issuers. Yet others fight for market share on the same indices as their competitors, like the S&P500 or Russell 2000. Such same-index ETFs account for approximately one-half of the total ETF market by dollar volume. Somewhat surprisingly, despite tracking an identical basket of stocks, these competing ETFs typically charge investors different fees. Even more surprisingly, the fee differentials are persistent and the higher fee ETFs not only survive but also flourish, often attracting more assets and trading activity.

How does competition in this market work, and what is it that investors are paying for when they choose a higher cost ETF over a cheaper competitor that tracks the same index? This paper shows that the answer is market liquidity. Consistent with our model, we show empirically that liquidity clienteles are instrumental in driving the fee-setting behavior of ETFs and explaining why competition emerges for some indices, but not others.

As an illustration of the central points of this paper, consider the fees of the largest three ETFs that track the S&P 500 index: State Street's SPY charges 9.45 basis points (bps) per annum, while BlackRock's IVV and Vanguard's VOO charge only 3 bps each.[2] Despite being more than twice as expensive, SPY has a greater market share (assets under management, AUM) than its nearest competitor (IVV). What distinguishes relatively expensive ETFs, such as SPY, from their competitors is the sheer amount of readily accessible liquidity in the secondary market.[3] SPY has 14 times greater daily trading volume than its nearest competitor, and also more than 10 times higher turnover. The situation is similar in most same-index ETFs: highly liquid first movers charge higher management fees compared to their competitors and fees show little sign of converging over time.

Our model of ETF competition predicts equilibria very much like the one illustrated by SPY and its competitors. The intuition is as follows. If faced with a choice of multiple ETFs tracking the same index, which ETF would

[1] Traders Magazine. 2014. Q&A: Vanguard's Bogle On the State of ETFs. March 26. https://www.tradersmagazine.com/departments/buyside/qa-vanguards-bogle-on-the-state-of-etfs/.

[2] There is a fourth ETF tracking S&P 500: State Street's SPLG also charges a 3-bps management fee. In contrast to SPY, also offered by State Street, SPLG is explicitly catered as to long- term investors and may not directly compete with the more liquid SPY fund.

[3] Arguably, the three ETFs in this example are different along other dimensions, such as index tracking performance. But SPY's higher fees are not a compensation for better index tracking performance, because SPY has a larger tracking error, compared to competitors.

an investor choose? Intuitively, and as we show formally, it depends on the investment horizon. Short-horizon investors ("high-turnover investors"), find it optimal to choose a more liquid ETF, even if it has a higher fee. Such investors would happily trade the SPY, because with short holding horizons, the fee differential becomes negligible, while the liquidity does not. This clientele, by trading more frequently, creates greater secondary market turnover and thereby reinforces the higher level of the ETF's liquidity. At the other end of the spectrum, long-horizon buy-and-hold investors are less concerned about liquidity, but more concerned about the fees that affect performance over the longer term. They will turn to one of SPY's lower-fee competitors. Their lower trading frequency reinforces the relatively lower level of liquidity in these competing ETFs.

The value of secondary market liquidity for an investor and the presence of liquidity clienteles affect the way ETFs set fees. A more liquid ETF, such as the SPY in this example, has no incentive to cut its fee to match its competitors because its high-turnover clientele is relatively insensitive to the fee. Short-horizon investors will continue to choose SPY due to its (self-perpetuating) high liquidity, which allows SPY to maintain its relatively high fee, despite competition from lower-cost funds. The attraction of short-horizon investors to more liquid ETFs creates a "liquidity begets liquidity" effect. Investors are caught in a form of prisoner's dilemma: while it would be optimal for all investors as a group to switch to a cheaper ETF, an individual short-horizon investor has no incentive to individually deviate from the more liquid, higher fee ETF. This effect allows issuers of highly liquid ETFs to extract a rent (via their fee) from the liquidity externalities of their clientele. Liquidity externalities also create a strong first-mover advantage among competing ETFs and lead to a less than perfectly competitive fee setting.

Using the model for guidance, we analyze the population of U.S. equity ETFs. We find that for ETFs that track the same index as at least one other ETF, the more liquid ETFs (measured by secondary market traded dollar volume and bid-ask spread) tend to charge higher fees, consistent with ETF issuers extracting a rent from the liquidity externalities of their investors. Within the same index, more liquid ETFs charge a fee premium of 10.95 bps p.a. compared to their direct competitors. More liquid ETFs maintain a 38% higher market share, and have 7.31 times greater trading volumes than their less liquid counterparts. We find that the liquidity channel dominates alternative explanations, such as capital gains lock-in and fund name recognition.

Our Shapley-Owen $R^2$ decomposition shows that liquidity is a major determinant of ETF market shares and management fees. It accounts 50 to 66% of the explained variation in market shares and 25 to 30% of the explained variation in fees. As alternative drivers, we include reputation-related factors like marketing expenses and brand recognition, the proportion of tax-sensitive investors, tracking error, and performance drag.

We find direct evidence for investor clienteles. We measure the share of transient investors (Bushee 1998) and the average holding duration (Cremers and Pareek 2016) of investors in each ETF. Using 13F data, we find that the share of short-term investors is 50% larger in high-fee ETFs relative to their competitors. Further, the holding duration of investors in high-fee funds is 3.64% lower than for investors in low-fee funds. These results support the mechanism proposed in the theory, by which high-fee ETFs capture the large, high-turnover clientele that allows them to charge a premium fee because of their high secondary market liquidity.[4]

Importantly, we find evidence of a strong first-mover advantage: 80% of high-fee funds are the first movers for their respective index, consistent with the first-mover advantage predicted by our model. A competitor ETF starts from a position of low liquidity and therefore draws in low-turnover investors that are more concerned about fees than liquidity, while the incumbent maintains the relatively fee-insensitive high-turnover investors. Recent anecdotal evidence from the Canadian market suggests that the first mover can establish a significant advantage even if competing ETFs are launched in rapid succession. The Purpose Bitcoin ETF (BTCC) was launched on February 18, 2021, with a management expense ratio (MER) of 1.5% of AUM. A virtually identical fund, the CI Galaxy Bitcoin ETF (BTCX), launched less than three weeks later on March 9, 2021, with a lower fee of 0.85%. By the end of 2021, BTCC accumulated US$1.96B in assets under management, about four times as much as its cheaper competitor (US$464M).

The theory highlights the importance of investor trading urgency in determining the relative advantages of competing ETFs. The intuition is that if investors in a given index have shorter holding periods (across all ETF competitors), the more liquid ETF can charge an even larger fee premium and capture a larger market share advantage. A one-standard deviation increase in investor trading urgency leads to a 3.35-bps wider gap in the fees of high- and low-fee ETFs tracking the same index, as well as a 2.48-bps larger difference in quoted spreads.

We also examine why some stock indices attract multiple competing ETFs, while others have a single monopolist ETF. A necessary condition for competition is sufficient total demand from investors so that a competing ETF could attract enough AUM to cover its fixed costs. Further, investor urgency should be high enough such that there is nontrivial demand for liquidity.



---

[4] Recent anecdotal evidence supports the intuition that ETF providers target different investor clienteles. In October 2020, the ETF provider Invesco launched the Invesco Nasdaq 100 ETF, tracking the same security basket as the highly traded Invesco QQQ Trust, but charging lower fees: 0.15% relative to 0.20%. In line with our model's intuition, John Feyerer (Invesco's senior director of equity ETF strategy) argues that "QQQM with its lower management fee may appeal to long-term buy-and-hold investors." See Michael Wursthorn. 2020. ETF Clones Multiply in Industry Fee War. Wall Street Journal. October 26. https://www.wsj.com/articles/etf-clones-multiply-in-industry-fee-war-11603704602.

We find empirical support for both of these drivers: ETF competition is more likely to emerge and be sustained when the aggregate index AUM and the share of short-term investors are higher. Additionally, we document that a higher share of tax-insensitive investors enhances ETF competition. The result is intuitive, considering that tax-sensitive investors may be more inclined to remain with an incumbent fund due to lock-in effects, and have weaker incentives to migrate to a competing fund even if it offers lower fees. Furthermore, the data suggest that ETFs tracking proprietary indices have some resistance to competition. In contrast, indices licensed from well-known entities like FTSE, Russell, S&P, NASDAQ, MSCI, or Dow Jones are more frequently targeted by multiple ETFs.

Liquidity fragmentation across same-index ETFs leads to welfare loss: by splitting investors across two (or more) competing ETFs, investors have shallower secondary market liquidity pools and higher trading costs. In addition to these welfare losses, the equilibrium involves transfers, as ETF issuers extract economic rents from their investors due to the value of liquidity.

Finally, our results help explain the striking concentration of liquidity in a handful of major funds: According to Balchunas (2016), 50% of ETF dollar volume is concentrated in the top 15 ETFs by traded volume (out of the total of almost 2,000 equity ETFs listed in the United States). This concentration of trading in a handful of ETFs persists despite no shortage of newcomers: a new ETF is launched on average every trading day.[5] Despite a "race to the bottom" in fee setting by more recent funds, the large incumbent funds are able to maintain their dominant market positions and retain relatively high fees: a phenomenon that we largely attribute to the value of secondary market liquidity.

Our paper sheds light on a key distinction between ETFs and mutual funds. Unlike mutual funds, ETFs benefit from continuous trading on secondary markets, meaning that a higher turnover enhances their liquidity and consolidates their market share. In contrast, increased turnover for mutual funds, as measured by daily flows, actually detracts from their performance, as documented by Johnson (2004) and Rakowski (2010). This is because investor flows force fund managers to make costly transactions in their portfolio. We further show that heterogeneity in investor horizons enables the coexistence of multiple ETFs tracking the same index, each catering to different market segments. In contrast, for mutual funds, which are not traded intraday, information and search frictions, among other factors, play a more pivotal role in product differentiation (as discussed in Hortaçsu and Syverson 2004).

This paper contributes to several strands of literature. The first is the growing body of studies of ETFs. While there has been much research on how ETFs affect the underlying markets, less attention has been paid to

[5] According to ETF.com, there were 150 ETF launches in the year 2021 (between January 1, and May 31, 2021), 318 in 2020, 255 in 2019, 269 in 2018, 271 in 2017, 247 in 2016, 284 in 2015, and 202 in 2014.

how ETFs compete.[6] A recent exception is the empirical study by Box, Davis, and Fuller (2020), which finds that fees of the incumbent ETF do not decrease as a result of competition. Our model provides an explanation for this tendency: Liquidity externalities prevent investors from switching to lower-fee competitors, allowing the incumbent ETF to maintain their higher fee and extract rents. In a similar vein, Evans et al. (2021) document that authorized participants engage in “operational shorting” by first selling ETF units and subsequently delaying the assembly of the basket and the creation of new ETF shares. Finally, Easley et al. (2021) study ETF fee and liquidity competition between indices and find that more active ETFs tend to charge higher fees and are less frequently traded by investors on secondary markets. The intuition is that active ETFs charge a fee premium over passive ETFs as they promise investors excess returns. In contrast, we focus on ETF competition within an index, where by construction there is no material difference in activeness, and shed light on the importance of network effects and liquidity clienteles. In this context, we show that more liquid ETFs within an index are able to sustain a higher fee.

A related strand of literature examines competition among index funds. In a case study of S&P 500 index funds, Hortaçsu and Syverson (2004) suggest that search costs and product differentiation play an important role in explaining fee dispersion between competing funds. Our study is different in that search costs and product differentiation are unlikely to be material for major index ETFs, whereas secondary market liquidity varies considerably among the competing ETFs and is a major draw-card of ETFs, particularly to institutional investors. For example, Agapova (2011) finds that institutional investors with higher trading needs prefer ETFs over conventional index funds.

Our paper is also related to studies of mutual funds and highlights an important difference between ETFs and traditional mutual funds. In mutual funds, high investor turnover is an undesirable feature of a fund, as redemptions and inflows create negative externalities on other investors. For example, Edelen (1999) shows that while open-ended fund managers may be informed, they often trade as uninformed traders simply because they facilitate inflows or redemptions, and this effect leads to fund underperformance. In contrast, we show that ETF investors benefit from higher investor turnover, because it leads to higher secondary market liquidity. Second, ETF managers set fees in

[6] Recent studies identify how ETFs transmit shocks to the underlying securities: Malamud (2016); Ben-David, Franzoni, and Moussawi (2018); Krause, Ehsani, and Lien (2014); Chinco and Fos (2021); Bhattacharya and O’Hara (2018); Huang, O’Hara, and Zhong (2020); Dannhauser and Hoseinzade (2022). Many studies examine the effects of ETFs on the efficiency of individual securities. Some find that ETFs improve price discovery (Dannhauser (2017); Ernst (2022); Madhavan (2016); Lettau and Madhavan (2018); Madhavan and Sobczyk (2016);Glosten, Nallareddy, and Zou (2021); Wermers and Xue (2015); Marshall, Nguyen, and Visaltanachoti (2013); Li and Zhu (2016)). However, others find that ETFs harm informational efficiency in underlying securities, increasing the proportion of market-wide information in prices but decreasing stock-specific information, and increasing trading costs (Hamm (2014); Da and Shive (2018); Israeli, Lee, and Sridharan (2017); Agarwal et al. (2018)).

response to the prevailing liquidity clienteles, rather than choosing the fees to entice a specific type of clientele (e.g., set high load fees to attract long-term clientele and minimize redemptions as in Chordia 1996). Therefore, ETFs and unlisted mutual funds operate and compete in rather different ways.

We also contribute to the liquidity clientele literature. Amihud and Mendelson (1986) show that long-horizon investors tend to concentrate in relatively less liquid stocks in which they can earn an illiquidity premium, while short-horizon investors are willing to sacrifice some returns in exchange for liquidity and thus hold more liquid stocks.[7] A similar mechanism is at play in our model: short-horizon investors hold highly liquid ETFs at the cost of a higher fee and long-horizon investors hold less liquid ETFs but enjoy a lower fee. However, our model differs in several important regards. Liquidity in our model is endogenous and a function of investor choices, whereas in Amihud and Mendelson (1986) the bid-ask spread is exogenously given. Importantly, in our model, ETF issuers are aware of the liquidity/fee trade-offs that investors face, and they set fees to extract rents from their secondary market liquidity. In a related model of vertical differentiation, Foucault and Parlour (2004) argue that exchanges can choose different trading technologies and listing fees to extract rents from IPO firms that are heterogeneous in their preferences for secondary market liquidity.

Finally, we build on a rich theoretical literature in industrial organization and finance examining network externalities, starting from the classical model of Katz and Shapiro (1985), who raise the possibility of multiple equilibria in markets with network effects. The concept of fulfilled expectation equilibria, where the equilibrium distribution of actions in a network game matches the initial beliefs, dates back to Green (1973). On financial markets, Pagano (1989) shows that if an asset is traded on two identical exchanges, in equilibrium market participants gravitate to a single exchange due to network effects. Market fragmentation is only sustainable with heterogeneous transaction costs, and only for knife-edge equilibria. In our model, ETFs are essentially "liquidity pools" where buyers and sellers meet to trade a particular security basket. Whereas in Pagano (1989) investors have different optimal trade sizes, in our setup traders have heterogeneous investment horizons. The distinction reinforces liquidity effects in equilibrium: The more liquid ETF attracts high-turnover traders, which helps it solidify its position as the more liquid ETF. In line with our documented first-mover advantage, Halaburda, Jullien, and Yehezkel (2020) highlight the path dependence in equilibrium selection: customers are likely to coordinate on platforms that already dominate the

[7] The broader literature on liquidity premiums is too vast to review here. See Holden, Jacobsen, and Subrahmanyam (2014) for a detailed review. The most closely related papers to ours are those that show that bond market yield spreads are related to secondary market liquidity, implying that investors are willing to sacrifice some return in exchange for liquidity (e.g., Amihud and Mendelson 1991; Krishnamurthy 2002; Longstaff, Mithal, and Neis 2005; Friewald, Jankowitsch, and Subrahmanyam 2013).

market, consolidating their market share. Finally, Markovich and Yehezkel (2022) examine how specific user groups can help maintain the dominance of a platform; In our setup, high-turnover investors allow the incumbent ETF to sustain better liquidity levels than its competitors.

There are parallels between the fragmentation of investors across different ETFs and the fragmentation of trading across markets. As noted earlier, trading venues compete on several dimensions, such as speed, fee structures, and transparency, which leads to clientele effects (Foucault, Kadan, and Kandel 2005; Yao and Ye 2018). In our model, investors only weigh network effects (liquidity) and fees when selecting an ETF. The two funds are otherwise identical. This contrasts, for example, with fragmentation between fast and slow exchanges (Pagnotta and Philippon 2018) or between lit and dark markets (Zhu 2014), where investors choose between two different market designs.

# 1. A Model of ETF Competition

## 1.1 Model primitives

**1.1.1 Asset** Consider a continuous-time economy where a single equity index can be traded. The index tracks a basket of stocks which has a common value $v_t$ at time $t > 0$. The index value follows a compound Poisson process where news arrives at rate $\eta > 1$. Upon the receipt of news, the common value $v_t$ changes by either $+\sigma$ (good news) or $-\sigma$ (bad news), with equal probabilities.

**1.1.2 Trading environment** Assets are traded on a continuous-time limit order market with time priority. A limit order is a price quote to either buy (a bid) or sell (an ask) a particular number of asset units. A "marketable" limit order (which is either a limit buy order with a price higher than the lowest prevailing ask or a limit sell order with a price below the highest prevailing bid) trades immediately against the corresponding bid or ask quote.

**1.1.3 Agents** Four types of agents are present in the economy:

(i) two ETF issuers that may choose to launch index funds;
(ii) a unit continuum of risk-neutral investors (**I**) who trade the index for liquidity reasons;
(iii) $m \geq 2$ competitive risk-neutral market makers (**M**) providing liquidity across both ETFs;
(iv) a fast arbitrageur (**A**) who can react to news before market makers.[8]

[8] A fifth agent, "authorized participants" (APs), play a role in the mechanics of ETFs by facilitating the creation or redemption of ETF units to keep the supply in line with investor demand so that the ETF price mirrors the value of the underlying basket of stocks. In the interest of maintaining a parsimonious model, we do not include APs as their role is rather mechanical and not directly related to liquidity given that market makers do not need to also be APs. Comerton-Forde and Marta (2021) document that most of the secondary market trading volume of French ETFs is generated by high-frequency traders who do not participate in the primary creation/redemption market.

There are infinitely many agents in (ii), such that individual investors do not affect the equilibrium outcomes. Table A1. in Appendix A summarizes the model's notation.

There are at most two risk-neutral exchange-traded *funds* (ETFs) that launch sequentially, that is, one ETF is a *leader* (L), and the other is a *follower* (F). The funds charge management fees $f_L$ and $f_F$, respectively, measured as a fraction of assets under management per unit time. Without loss of generality, we normalize the funds' marginal operating cost to zero.

Funds accumulate assets under management (AUM) from a unit mass of *investors*, who choose between the two ETFs. Investors are hit by (large enough) liquidity shocks at random times, exponentially distributed with rate $\lambda_i$ (as in, e.g., Pagnotta and Philippon 2018). Conditional on a liquidity shock, each investor $i$ is equally likely to buy or to sell a quantity $2Q \times \mathrm{d}i$ units of an ETF, where $Q > 0$ and $\mathrm{d}i$ is arbitrarily small. ETF shares are created and redeemed as needed, and the expected AUM in the ETFs tracking the index is $Q$.[9] Equivalently, each investor $i$ has an expected holding period of $\lambda_i^{-1}$. We assume that investors who are hit with a liquidity shock are impatient and only submit marketable orders. Finally, investors do not trade on news, but arbitrageurs do.

Importantly, investors have heterogeneous holding periods. In particular, we label a mass $\alpha$ of investors as *high-turnover* with arrival rate $\lambda_h$. Conversely, a mass $1-\alpha$ of investors are *low-turnover* and have the arrival rate $\lambda_\ell$, where $\lambda_h > \lambda_\ell > 0$. We write the average arrival rate for private value shocks as

$$\Lambda = \alpha \lambda_h + (1-\alpha)\lambda_\ell, \tag{1}$$

which we interpret as a measure of aggregate investor trading urgency.

**1.1.4 Timing** Figure 1 summarizes the sequence of events at each time $t$. First, at $t=-3$, the leader ETF posts a management fee $f_L$. Next, at $t=-2$, the follower ETF chooses whether to enter the market; if so, it posts a management fee $f_F$. At $t=-1$, each investor observes fees and chooses an ETF to trade. Trading starts at $t=0$: **M**s simultaneously post bid and ask quotes to buy or to sell the ETFs and investors may receive private value shocks. Investors do not receive an ETF endowment before $t=0$, and therefore they incur the half-spread on every trade they make, including when establishing their initial position.

At any point following $t=0$, there is continuous trading. In particular, a trade may be triggered by one of two possible events: either (a) the arrival of news, in which case the arbitrageur **A** trades against the market-maker, or (b) the arrival of a liquidity shock, in which case an investor **I** trades against **M**. Following a trade, market makers can update their quotes and the game continues.

9 An implicit assumption is that investors cannot replicate the underlying index directly without incurring prohibitive costs (such as price impact or time costs to manage a large portfolio).

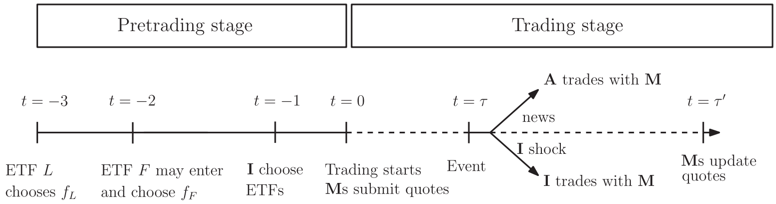


**Figure 1**
**Model timing**

## 1.2 Equilibrium

We solve the game by backward induction. First, we take ETF fees and investor choices as given, and derive the competitive bid-ask spread at $t \geq 0$. Next, we use the *attached consumers* solution concept in Biglaiser and Crémer (2020) to obtain the investors' equilibrium choice of ETFs at $t=-1$ given the management fees. Finally, we solve for equilibrium ETF management fees at $t=-2$ and $t=-3$, respectively.

**1.2.1 Liquidity provision and trading in the secondary market** The structure of the trading subgame starting at $t=0$ closely follows the literature on high-frequency trading and latency arbitrage: for example, the models in Budish, Cramton, and Shim (2015); Menkveld and Zoican (2017), or Brolley and Zoican (2022). We conjecture (and verify in the proof for Lemma 1) that at $t=0$ each **M** submits two quotes to trade ETF $k$: a bid quote to purchase $2Q\,\mathrm{d}i$ units of the ETF for $v-s_k$ and an ask quote to sell $2Q\,\mathrm{d}i$ units of the ETF for $v+s_k$, where $s_k \geq 0$. Therefore, $s_k$ denotes the half-spread for ETF $k$ (half of the distance between the bid and ask prices).

Consider first the case where news arrives at time $\tau$, and the value of index just before news is $v_\tau$. Upon the realization of good news, the asset value is $v_\tau+\sigma$. Since the arbitrageur is faster than the market maker, he can always "snipe" any stale quotes before **M** can update or remove them. The arbitrageur can buy ETF $k$ for $v_\tau+s_k$ and realize a positive profit of $\sigma-s_k$ as long as the half-spread is lower than the size of news $s \leq \sigma$, which we will show is always positive in equilibrium. The reasoning is symmetric for the case of bad news. Upon news, the arbitrageur optimally submits orders in both ETFs since they track the same index.

We turn next to the market makers' strategy. Let $\Lambda_k$ denote the arrival rate of an uninformed liquidity order for ETF $k$. Market-makers earn the half-spread $s_k$ when uninformed investors arrive to market, and incur a loss of $\sigma-s_k$ whenever their quotes are sniped by the arbitrageur upon the arrival of news. As in Menkveld and Zoican (2017), market-makers submit bid and ask quotes whenever the order book is empty, and arrive at the market in random order. The first market-maker to arrive at the exchange keeps her quotes in the order book, while the $m-1$ other market makers promptly cancel their orders given that they earn negative expected profits. The rationale is that an uninformed investor only trades against the first market-maker's quote, whereas all quotes in the order book are vulnerable to latency arbitrage losses upon a news event.

Since each market maker is equally likely to arrive first at the market with probability $\frac{1}{m}$, the expected profit for each **M** per unit of time for a quote size $2Q\,\mathrm{d}i$ is equal to[10]

$$\mathbb{E}\pi_{\mathbf{M}} = \frac{1}{m}\left[\frac{\Lambda_k}{\eta+\Lambda_k}s + \frac{\eta}{\eta+\Lambda_k}(s-\sigma)\right] \times 2Q\,\mathrm{d}i. \tag{2}$$

**Lemma 1.** *(Bid-ask spread)* The equilibrium bid-ask spread in ETF $k$ is pinned down by setting the expected profit in (2) equal to zero (otherwise there would be a profitable opportunity to undercut the spread in one ETF):

$$s^{\star}(\Lambda_k) = \frac{\eta}{\eta+\Lambda_k}\sigma. \tag{3}$$

The equilibrium bid-ask spread in equation 3 has natural properties. First, it increases in the arrival rate and size of news ($\eta$ and $\sigma$, respectively). Second, the ETF with a larger investor pool, that is, with a higher $\Lambda_k$, is more liquid. The rationale is that market makers are more likely to trade against uninformed investors on an ETF with a larger $\Lambda_k$, which reduces the average cost of providing liquidity.

The key findings of our model only require that the equilibrium bid-ask spread decreases in the mass of investors choosing a particular ETF. That is, the fee and liquidity rankings of the two ETFs do not depend on the specific functional form of the bid-ask spread. Our setup uses latency arbitrage and adverse selection costs as a mechanism to generate a positive spread that decreases in the mass of (uninformed) traders. This friction aligns with Ernst (2022), who documents that informed investors actively trade ETFs. Alternatively, the model could incorporate inventory costs as a driving force for the bid-ask spread. If two funds are not perfectly fungible, possibly due to clearing costs, market makers face higher inventory risks in the less traded fund, resulting in a wider spread.

**1.2.2 Investor ETF choice at $t=-1$** For an investor with arrival rate $\lambda_i$, the total cost of a round-trip in ETF $k$ can be decomposed into a *holding cost* (i.e., the management fee times the expected holding duration) and a *trading cost* (i.e., two times the half-spread in Lemma 1):

$$\mathrm{Cost}_i^k = \underbrace{\frac{1}{\lambda_i}f_k}_{\text{Holding cost}} + \underbrace{2s^{\star}(\Lambda_k)}_{\text{Trading cost}}. \tag{4}$$

[10] While market makers also incur management fees on long positions, they also earn fees from shorting ETFs. On average, since market-makers are equally likely to be long or short, their expected profits do not depend on fees.

The holding cost increases linearly in the expected holding duration because ETF fees accumulate over time. Meanwhile, the trading cost for each round-trip (i.e., consisting of one buy and one sell trade) does not depend on how long the position was held. As a result, trading costs make up a larger share of aggregate costs for high-turnover investors with shorter holding periods. Further, there are positive network externalities as investors prefer to be on the same platform as their peers to reduce trading costs.

We use the *attached consumers* (AC) solution concept first introduced in Biglaiser and Crémer (2020) to obtain the investors' ETF choice at $t=-1$. The AC solution is a Nash equilibrium refinement particularly suitable when there are coordination frictions between consumers in a setting in which firms compete sequentially on prices. In our setup, investors cannot easily coordinate on picking the lowest-fee fund to simultaneously minimize both holding and trading costs.

At the start of $t=-1$, all investors are allocated to the leader fund $L$ and may choose to migrate to the entrant fund, that is, the follower ETF $F$. The investor migration process proceeds as follows over an arbitrary number of steps. Investors evaluate their utility from staying with fund $L$ or migrating to $F$ under the assumption that no other investor moves. This assumption maps naturally to coordination problems. We rank investors by their gains from migrating, from high to low. If some investors *strictly* benefit from migrating to $F$, a small measure of those investors with the most to gain switch over to the follower fund. The process is repeated until no investor strictly benefits from moving.

Under the initial investor allocation, the aggregate investor arrival rate is equal to $\Lambda$ for the leader ETF and zero for the follower. Therefore, an investor with arrival rate $\lambda_i$ is better off by individually migrating at the beginning of $t=-1$ only if

$$\frac{1}{\lambda_i} f_F + 2\underbrace{s^\star(0)}_{=\sigma} < \frac{1}{\lambda_i} f_L + 2s^\star(\Lambda) \Leftrightarrow f_L - f_F > \lambda_i \times 2\sigma \frac{\Lambda}{\eta+\Lambda}, \tag{5}$$

where the last step follows from the bid-ask spread definition in Lemma 1.

Two immediate observations follow from Equation (5). First, in any two-fund equilibrium it must be that $f_F < f_L$: that is, the follower fund must post a lower fee to attract investors. Otherwise, no investor benefits from switching over to a fund with both a higher fee and lower liquidity. If $f_F \geq f_L$, then the only possible AC equilibrium is that all investors remain with the incumbent ETF.

Second, inequality (5) is more likely to be satisfied for low- than for high-turnover investors since $\lambda_\ell < \lambda_h$. This stems from the notion that low-turnover investors place a comparatively lower value on positive network externalities than their high-turnover counterparts because the holding costs holds a larger weight than the bid-ask spread for them. Consistent with our description of AC equilibrium mechanics, it is the low-turnover investors that migrate first to the follower fund since they stand to benefit relatively more from doing so.

The initial migration by a small fraction of low-turnover investors triggers a "snowballing" effect, ultimately compelling all low-turnover investors to transition to the follower fund $F$.[11] This is due to positive network externalities: any investor migration from $L$ to $F$ enhances the attractiveness of the follower fund over the leader by virtue of a reduced bid-ask spread.

If the entire mass $1-\alpha$ of low-turnover investors switches over to the follower fund, the resultant bid-ask spreads for funds $L$ and $F$ are equal to $s^\star(\alpha\lambda_h)$ and $s^\star((1-\alpha)\lambda_\ell)$, correspondingly. The migration process halts at this stage if no high-turnover investor finds it individually optimal to migrate from the leader to the entrant. That is, high-turnover investors stay with the leader if

$$\frac{1}{\lambda_h} f_L + 2s^\star(\alpha\lambda_h) \le \frac{1}{\lambda_h} f_F + 2s^\star((1-\alpha)\lambda_\ell) \Leftrightarrow$$

$$f_L - f_F \le \lambda_h \times 2\sigma \left( \frac{\eta}{\eta + (1-\alpha)\lambda_\ell} - \frac{\eta}{\eta + \alpha\lambda_h} \right). \qquad (6)$$

If inequality (6) is not met, then at least a small fraction of high-turnover investors also migrates to the entrant fund. This catalyzes a "snowballing" analogous to the one discussed before, leading to a unanimous shift of high-turnover investors from $L$ to $F$. Lemma 2 formalizes the result.

**Lemma 2.** *(Investors' equilibrium ETF choice at $t = -1$)*
Define $\Gamma_0 = \lambda_\ell \times 2\sigma \frac{\Lambda}{\eta+\Lambda}$ and $\Gamma_1 = \lambda_h \times 2\sigma \left( \frac{\eta}{\eta+(1-\alpha)\lambda_\ell} - \frac{\eta}{\eta+\alpha\lambda_h} \right)$. The unique AC equilibrium investor strategies in the game commencing at $t = -1$ are:

1. If $f_L - f_F \le \Gamma_0$, all investors select the leader fund $L$;
2. If $\Gamma_0 \le \Gamma_1$ and $f_L - f_F \in (\Gamma_0, \Gamma_1]$, low-turnover investors opt for the follower fund $F$ and high-turnover investors for $L$;
3. If $f_L - f_F > \max\{\Gamma_0, \Gamma_1\}$, all investors migrate to the follower fund $F$.

It is easy to verify that the AC equilibrium in Lemma 2 satisfies the conditions of a Nash equilibrium; no investor, regardless of type, can unilaterally improve their situation by selecting an alternative fund. This aligns with lemma 1 in Biglaiser and Crémer (2020, p. 237), which proves that all AC equilibria are Nash equilibria. However, as is common in games with network externalities (see, e.g., Dow 2004), the AC equilibrium is not the unique. The second, Pareto-optimal, Nash equilibrium emerges as a corner solution in which all investors choose the lowest-fee fund, which will therefore have the lowest feasible spread.

[11] Biglaiser and Crémer (2020) show that the mass of investors that migrate at once is irrelevant: the snowballing effect ensures that migration can occur in both "large" or "small" steps (see lemma 2 on p. 238).

The *attached consumers* solution concept therefore addresses two salient challenges inherent games with network externalities. First, is establishes a sharp criterion for equilibrium selection, resolving multiplicity. Second, it incorporates the notion of incumbency in the equilibrium analysis by emphasizing the role of initial investor allocation. As Biglaiser and Crémer (2020, p. 234) point out, “the concept of Nash equilibrium has no role for incumbency.” Finally, while the assumption of having only two investor types is a simplification of the ETF market, Biglaiser and Crémer (2020, p. 235) assert that the AC equilibrium concept and its results translate to an arbitrary finite number of types. Our model framework could therefore be extended to potentially represent a broader range of investor horizons.

**1.2.3 ETF fee setting** In this section, we solve for the optimal fee-setting strategy of ETF issuers at times $t=-2$ and $t=-3$ for the follower and leader funds, respectively. Both ETFs maximize their expected profit per unit of time, equal to the product between their management fee and the equilibrium mass of investors they attract. We proceed by backward induction, a standard approach in games with Stackelberg leadership. That is, we first solve for the follower's best response function at $t=-2$. Subsequently, we determine the leader's optimal management fee at $t=-3$. From Lemma 2, two funds may coexist with positive market shares solely if $\Gamma_0 \leq \Gamma_1$. We first study a scenario in which this inequality holds, facilitating the existence of a two-fund equilibrium, before considering the alternative scenario where $\Gamma_0 > \Gamma_1$.

**1.2.4 Two-fund equilibria** When $\Gamma_0 \leq \Gamma_1$, the follower's optimal response $f_F^\star$ to the leader's fee $f_L$ at $t=-2$ is determined as follows:

$$f_F^\star(f_L)=\begin{cases} 0 & \text{if } f_L \leq \Gamma_0 \text{ (do not enter the market)} \\ f_L-\Gamma_0 & \text{if } f_L \in \left(\Gamma_0, \frac{\Gamma_1+(1-\alpha)\Gamma_0}{\alpha}\right] \\ f_L-\Gamma_1 & \text{if } f_L > \frac{\Gamma_1+(1-\alpha)\Gamma_0}{\alpha}. \end{cases} \tag{7}$$

First, if the leader ETF $L$ sets a fee $f_L$ at or below $\Gamma_0$, then from Lemma 2, the follower ETF cannot set any positive fee $f_F$ allowing it to attract investors. Hence, ETF $F$ opts not to enter the market at $t=-2$.

Second, if $f_L$ falls within the range $(\Gamma_0, \Gamma_1]$, the follower's optimal strategy is to set $f_F = f_L - \Gamma_0$ to capture the low-turnover investors. Setting a higher fee would result in $f_L - f_F > \Gamma_0$, yielding zero market share for the follower. Conversely, offering a discount greater than $\Gamma_0$ on the leader's fee would unnecessarily reduce the follower's per-unit fee without expanding market share, thus directly diminishing profits.

Third, when $f_L > \Gamma_1$, the follower faces a choice: to either offer a deep fee discount at $f_F = f_L - \Gamma_1$, allowing it to capture the entire market, or offer a lower discount and set $f_F = f_L - \Gamma_0$, to serve only the $1-\alpha$ segment of low-turnover investors. The follower opts to target both high- and low-turnover investors if

$$1\times(f_L-\Gamma_1)>(1-\alpha)(f_L-\Gamma_0)\Leftrightarrow f_L>\frac{\Gamma_1+(1-\alpha)\Gamma_0}{\alpha}>\Gamma_1, \tag{8}$$

where the last inequality follows from $\Gamma_1>\Gamma_0$. Together, the three scenarios define the follower's reaction function as expressed in Equation (7).

Next, we consider the leader's optimal fee setting at $t=-3$. The maximum fee that allows the leader to monopolize the market by preventing entry is $f_L=\Gamma_0$; setting a fee above this threshold would enable the follower to undercut the leader's price, thereby attracting at least the low-turnover investor segment. Importantly, the leader would avoid setting $f_L>\frac{\Gamma_1+(1-\alpha)\Gamma_0}{\alpha}$ since it would incite the follower to significantly undercut the leader, leaving ETF $L$ with no customers. However, the leader can strategically permit entry by setting $f_L=\frac{\Gamma_1+(1-\alpha)\Gamma_0}{\alpha}$ at $t=-3$, catering exclusively to the high-turnover investor base $\alpha$. It turns out that accommodating the follower's entry is optimal if $\Gamma_1>\Gamma_0$, since

$$\underbrace{\alpha\times\frac{\Gamma_1+(1-\alpha)\Gamma_0}{\alpha}}_{\text{Leader profit if it accommodates entry}}=\Gamma_1+(1-\alpha)\Gamma_0>\underbrace{1\times\Gamma_0}_{\text{Leader profit if it deters entry}}. \tag{9}$$

**1.2.5 Single-fund equilibria** In the case where $\Gamma_0>\Gamma_1$, the ETFs are unable to segment their investor clientele. At $t=-2$, the follower's reaction function simplifies to:

$$f_F^\star(f_L)=\begin{cases}0, & \text{if } f_L\le\Gamma_0 \text{ (no market entry)},\\ f_L-\Gamma_1, & \text{if } f_L>\Gamma_0 \text{ (complete market capture)}.\end{cases} \tag{10}$$

As before, a leader fee below $\Gamma_0$ is enough to deter the follower fund from entering the market. However, should the leader set a fee above $\Gamma_0$, the follower can respond with a lower fee $(f_L-\Gamma_1>0)$ and sweep the entire market. Therefore, the optimal strategy for the leader is to set the fee at $f_L=\Gamma_0$, which is the maximum fee that prevents the follower's entry. Proposition 1 summarizes the equilibrium fees in all scenarios.

**Proposition 1.** *(Equilibrium fees and market structure)*
Let $\Gamma_0=\lambda_\ell\times2\sigma\frac{\Lambda}{\eta+\Lambda}$ and $\Gamma_1=\lambda_h\times2\sigma\left(\frac{\eta}{\eta+(1-\alpha)\lambda_\ell}-\frac{\eta}{\eta+\alpha\lambda_h}\right)$.

(i) If $\Gamma_0>\Gamma_1$, then the leader fund sets $f_L^\star=\Gamma_0$ at $t=-3$ and captures the entire market.
The follower fund does not enter the market at $t=-2$.

(ii) If $\Gamma_0 \le \Gamma_1$, then a two-fund equilibrium emerges where the management fees are:

$$f_L^\star = \frac{\Gamma_1 + (1-\alpha)\Gamma_0}{\alpha} \text{ and } f_F^\star = f_L^\star - \Gamma_0 = \frac{\Gamma_1 + \Gamma_0}{\alpha} - 2\Gamma_0. \tag{11}$$

At $t=-1$, all low-turnover investors choose the follower fund $F$ whereas the high-turnover investors choose the leader fund $L$.

We note that the attached consumers equilibrium concept can be extended to scenarios where three or more funds compete on the same index. This is made possible by the sequential nature of the game and the clear order in which investors make decisions. If a third fund were to enter the market tracking the same index, it would have to set its fees sufficiently low to initially attract low-turnover investors away from the second fund, thereby creating an even more cost-effective option for the investors with the longest holding horizon.

Our model assumes a mature investor base, with the total mass of investors fixed at one. However, in reality, the investor base could grow due to increasing interest in ETFs, either from diverse products or from new investors entering financial markets through gamified platforms. In a richer version of the model, this assumption could be relaxed. A larger mass of investors can deter entry by reinforcing the leader's initial liquidity, making it more challenging for new entrants to gain traction. Formally, with a greater mass of investors, there is a higher aggregate turnover ($\Lambda$) and a smaller spread differential, leading to a higher $\Gamma_0$ and a lower $\Gamma_1$ in the context of Proposition 1. While a comprehensive dynamic model is beyond the scope of our paper, we refer interested readers to Cabral (2011), who builds an overlapping generations model exploring competition between networks.

### 1.3 Comparative statics and predictions

We first focus on the two-fund equilibrium to obtain predictions about the cross-section of funds (Predictions 1 to 3). We consider how ETF fees relate to investor clienteles (holding horizons), market shares, liquidity, and profits. Further, our model generates a prediction about how changes in investor urgency affect the fee differentials of the competing ETFs (Prediction 4). Finally, we consider the first-mover advantage and conditions that lead to a two-fund equilibrium (Predictions 5 and 6). Figure 2 graphically illustrates the cross-sectional predictions by plotting comparative statics of equilibrium fees and bid-ask spreads with respect to the share of high-turnover investors.

#### 1.3.1 Cross-sectional predictions

**Prediction 1.** The high-fee ETF attracts shorter term investors than low-fee ETFs.

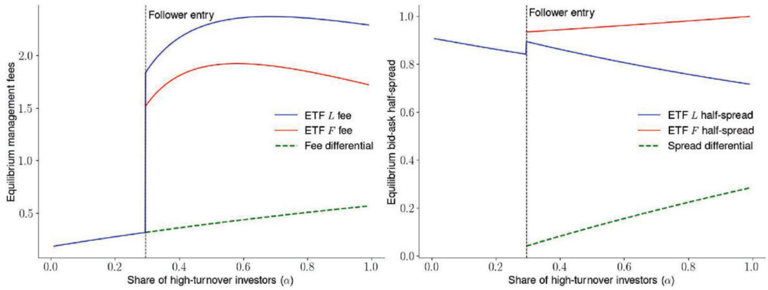


**Figure 2**
**Equilibrium comparative statics**
This figure illustrates the equilibrium outcomes as a function of the mass of high-turnover investors ($\alpha$). In particular, we plot the equilibrium ETF management fees (left), bid-ask spread (right), as well as the fee and spread differentials. Parameters: $\eta=10$, $\sigma=1$, $\lambda_\ell=1$, and $\lambda_H=4$.

The first prediction follows immediately from Proposition 1. In a two-fund equilibrium, the investors with high arrival rates (equal to $\lambda_h$) invest in the leader, high-fee ETF, while investors with low arrival rates (equal to $\lambda_\ell$) choose the follower, low-fee ETF.

**Prediction 2.** High-fee ETFs have higher liquidity (narrower bid-ask spreads) than low-fee ETFs.

The second prediction stems directly from Lemma 1, which states that the round-trip cost is negatively related to the aggregate investor arrival rate, $\Lambda_k$ for $k\in\{L,F\}$, as well as the fact that the combined arrival frequency of high-turnover investors exceeds that of their low-turnover counterparts (that is, $\alpha\lambda_h>(1-\alpha)\lambda_\ell$).

**Prediction 3.** ETFs with high fees have higher trading intensity (trades per unit time, or volume) and turnover (trades per unit investment) in the secondary market than ETFs with low fees.

Trading intensity in the two ETFs (trades per unit time) can be proxied by the aggregate investor arrival rate, $\Lambda_L=\alpha\lambda_h$ and $\Lambda_F=(1-\alpha)\lambda_\ell$. In a two-fund equilibrium, high-turnover investors find it optimal to choose the liquid ETF $L$, as they weigh liquidity savings more than the higher management fee. The presence of a high-turnover clientele contributes to the high trading intensity of the leader ETF.

### 1.3.2 Effects of investor urgency on and fee differentials

**Prediction 4.** The fee and bid-ask spread differential between the two ETFs increases in the share of high-turnover investors.

From Proposition 1 and since $\Gamma_0 = 2\sigma\lambda_\ell \frac{\Lambda}{\Lambda+\eta}$ and $\Lambda = \alpha\lambda_h + (1-\alpha)\lambda_\ell$, the fee differential in a two-fund equilibrium can be written as

$$f_L^\star - f_F^\star = 2\sigma\lambda_\ell \frac{\lambda_\ell + \alpha(\lambda_h - \lambda_\ell)}{\lambda_\ell + \alpha(\lambda_h - \lambda_\ell) + \eta}, \tag{12}$$

which increases in the share of high-turnover investors $\alpha$. Intuitively, a higher mass of high-turnover investors generates sufficient positive network externalities that allows the leader fund to charge a higher markup without risking to lose the high-turnover clientele.

Further, a larger mass of investors in the high-fee fund allows market makers to post narrower spreads in the leader ETF $L$, leading to a wider spread differential:

$$\Delta s \equiv s_F^\star - s_L^\star = 2\sigma\left(\frac{\eta}{\eta + (1-\alpha)\lambda_\ell} - \frac{\eta}{\eta + \alpha\lambda_h}\right) > 0, \tag{13}$$

To see that $\Delta s$ increases in $\alpha$, we note that

$$\frac{\partial \Delta s}{\partial \alpha} = 2\sigma\lambda_h\eta\left(\frac{\lambda_h}{(\eta + \alpha\lambda_h)^2} + \frac{\lambda_\ell}{(\eta + (1-\alpha)\lambda_\ell)^2}\right) > 0. \tag{14}$$

### 1.3.3 First-mover advantage and industry structure

**Prediction 5.** There is a first-mover advantage whereby the first ETF to launch for a given index has higher liquidity, higher trading intensity, and charges a higher fee than subsequent ETFs tracking the same index.

Prediction 5 arises from the dynamics of the "attached consumers" equilibrium. The follower ETF starts from a position of low liquidity and therefore draws in low-turnover investors that are more concerned about fees than liquidity or network externalities. Consequently, the first mover— the leader fund—secures the high-turnover investors, who exhibit a lower sensitivity to management fees. This strategic position endows the leader ETF with a competitive edge against newcomers. Since the higher liquidity of the leader fund appeals more to short-term investors, who prioritize trading costs over fees, the first mover is able to levy higher fees.

**Prediction 6.** A two-ETF economy obtains only if the share of high-turnover investors is large enough.

From Proposition 1, a two-ETF market structure obtains if and only if $\Gamma_1 - \Gamma_0 \geq 0$ or, equivalently,

$$\Delta\Gamma \equiv 2\sigma\left[\lambda_h\left(\frac{\eta}{\eta + (1-\alpha)\lambda_\ell} - \frac{\eta}{\eta + \alpha\lambda_h}\right) - \lambda_\ell\frac{\Lambda}{\eta + \Lambda}\right] \geq 0. \tag{15}$$

A simple derivation confirms that $\Delta\Gamma$ increases in $\alpha$ since

$$\frac{\partial \Delta\Gamma}{\partial \alpha} = 2\sigma \lambda_h \eta \left( \frac{\lambda_h}{(\eta+\alpha\lambda_h)^2} + \frac{\lambda_\ell}{(\eta+(1-\alpha)\lambda_\ell)^2} \right) > 0. \tag{16}$$

This implies that the follower ETF is more likely to enter the market for larger values of $\alpha$, the share for high-turnover investors.

Figure 2 illustrates this result. When the proportion of high-turnover investors ($\alpha$) is low, the leader ETF optimally sets a lower management fee to preempt market entry since the follower can capture the entire investor base. However, once $\alpha$ surpasses a critical threshold, the leader is able to elevate its fee and accommodate entry by ETF $F$. The leader's fee always exceeds that of the follower's. The fee differential increases in the leader's liquidity advantage, as quantified by the proportion of high-turnover investors.

The findings reported by the right panel echo those for bid-ask spreads. A higher share of high-turnover investors bolsters network effects within the leader fund, yielding tighter bid-ask spreads. At the same time, a lower mass of low-turnover investors results in wider spreads for the follower fund. Consequently, the differential between the leader's and follower's spreads increases in $\alpha$. We highlight that the leader's bid-ask spread experiences a discrete jump upon ETF $F$'s entry, since the low-turnover segment of the investor base immediately migrates to the follower.

**1.3.4 Welfare implications** A natural welfare measure in our model, since all investors trade whenever they receive a liquidity shock, is to subtract investor costs (both management fee and illiquidity) from the expected profits of the ETFs. We denote by $\omega_k$ the mass of investors who choose ETF $k$ in equilibrium (i.e., the market share). The aggregate cost for investors across both ETFs $L$ and $F$ follows from Equation (4):

$$\text{AggregateCost}_{\text{Investors}} = Q \times \left( \sum_{k\in\{L,F\}} \omega_k f_k + 2\sigma \frac{\eta}{\eta+\Lambda_k} \right). \tag{17}$$

Further, the expected profit of ETFs is

$$\mathbb{E}\text{Profit}_{ETF} = Q \sum_{k\in\{L,F\}} \omega_k f_k. \tag{18}$$

In a two-ETF economy, the expected welfare (net of the gains from trade, which investors always realize) can be written as the negative of liquidity costs driven by the asymmetric information and latency arbitrage.

$$\text{Welfare}_{\text{two ETF}} = -Q \times \left( 2\sigma \sum_{k\in\{L,F\}} \frac{\eta}{\eta+\Lambda_k} \right). \tag{19}$$

We note that management fees "wash out" in the welfare computation, as they represent transfers from investors to the fund. In a single ETF case, all investors choose the same ETF, which has a turnover of $\Lambda = \sum_k \Lambda_k$, which implies that the half-spread is $s_{\text{single-ETF}} = 2\sigma \frac{\eta}{\eta+\Lambda}$. We can therefore obtain a measure for the welfare loss in the oligopolistic equilibrium:

$$\begin{aligned}\Delta\,\text{Welfare} &= \text{Welfare}_{\text{one ETF}} - \text{Welfare}_{\text{two ETF}} \\ &= 2Q\sigma \underbrace{\left( \sum_{k\in\{L,F\}} \frac{\eta}{\eta+\Lambda_k} - \frac{\eta}{\eta+\Lambda} \right)}_{>0}. \end{aligned} \tag{20}$$

Splitting up liquidity generates network inefficiencies, which reduces welfare in equilibrium. Each fund aggregates only a fraction of investors, leading to a higher adverse selection cost since arbitrageurs can trade in both ETFs upon the realization of news.

## 2. Empirical Analysis

### 2.1 Data and descriptive statistics

We obtain daily data from ETF Global on the universe of U.S.-domiciled ETFs that track identical or very similar indices. Since the ETF Global coverage is only complete starting from 2016, our sample runs from January 2016 through December 2020. We focus on "plain-vanilla" funds, and consequently impose a number of simple filters on the daily ETF Global data set: (a) the asset class is "Equity," (b) funds are not levered (i.e., the *Leverage* indicator is zero), (c) funds are passive (i.e., the *Active* indicator is zero), (d) funds are not inverse (i.e., the *Description* field does not include the words "short" or "bear,"), and (e) the fund is not an ETN (exchange traded note). The next set of filters are applied on quarterly aggregated data. We follow Broman and Shum (2018) and apply two additional filters: we focus on ETFs that are active for at least 10 quarters in our sample and exclude from our sample the first two quarters of an ETF's existence. To closely track the model predictions, we retain for each index and quarter the largest-two ETFs by assets under management (i.e., removing the third-largest competitors). To be able to sharply define low- and high-fee funds, we focus on ETF pairs charging different fees to track the same index. Appendix C and Appendix D report some further (minor) data cleaning procedures and provide full details on matching competitor ETFs. In particular, we exclude two indices (SPLG and SPSM) which change their benchmark indices throughout our sample.

We use two approaches to identify the subsample of ETFs that compete in tracking the same or effectively the same index as other ETFs, which we refer to as "same-index" ETFs. First, any ETF that tracks exactly the same index as another ETF (using data from ETF Global) is added to the list of same-index ETFs. There are 49 such ETFs. Second, we identify ETFs that effectively track the same or very similar index and can therefore be viewed as direct competitors.[12] We do this by focusing on the ETFs for which we are most likely to find direct competitors[13], then manually identifying close competitors using FactSet data (FactSet provides "Comps", meaning comparable ETFs when an ETF has a close competitor)[14], and then empirically verifying any identified competitor ETFs by requiring that the monthly return correlation of an ETF and its identified competitor(s) is above 0.90 over the period 2016–2020. This process yields a further 56 ETFs that face direct competition in effectively tracking the same or very similar index.

These basic screens give us our baseline sample of 105 equity ETFs that have a combined AUM of $1.71 trillion and combined average daily dollar volume of $45.24 billion in the secondary market.

We use daily data from ETF Global to compute quarterly averages for each ETF's management fee (net expense ratio or management expense ratio [MER]), other fund expenses (which would include, among others, index licensing costs and distribution costs), fee waivers, assets under management (AUM), and number of constituents. We record the issuer, underlying index, and inception date for each ETF. We obtain data on time-weighted quoted, effective, and realized spreads from TAQ Intraday Indicators, measured in both basis points and dollars. We add data daily ETF returns, prices, and traded volume from the Center for Research in Security Prices (CRSP) database. The data on the underlying index returns and the ETF legal structure (unit investment trust or open-ended fund) are from FactSet. The outstanding daily ETF shares are from Morningstar Direct.

Furthermore, we collect data on fees allocated to marketing (i.e., 12B-1 fees, as in Barber, Odean, and Zheng 2005) from the CRSP Mutual funds database. We also parse ETF annual reports and hand-collect data on security lending revenue. To proxy for the ETF's name recognition, we collect data on the average daily number of times each ETF ticker is mentioned on the social media Stocktwits.[15]

Finally, we construct empirical proxies for investors' trading horizon. To do so, we use quarterly holdings for each ETF in the sample from 13F filings data and the definition of transient and tax-sensitive investors in Bushee (1998) to compute the AUM share of short-term and tax-insensitive investors

[12] For example, both VTI (Vanguard Total Stock Market ETF) and ITOT (iShares Core S&P Total U.S. Stock Market ETF) offer total U.S. stock market exposure and have very similar percentage holdings in Technology, Consumer Cyclicals, Financial and other sectors. These ETFs' underlying index returns have a correlation coefficient of 0.99 too. The ETFs effectively track the same underlying portfolio, and their underlying indices differ only in brand name (CRSP Total Market Index vs. S&P Total Market Index).

[13] Following the guidance from the theory model, we focus on ETFs that have above-median AUM and are in the 90th percentile by dollar volume in 2019 (to avoid the market turbulence of March-April 2020).

[14] In identifying comparable ETFs, FactSet considers the share class type, the legal structure, the country of the primary exchange, whether the ETFs follow the same benchmark, and whether the ETFs have the same selection strategy as the specified fund share class.

[15] We thank Edna Lopez Avila and Charles Martineau for sharing this data set with us.

**Table 1**
**Summary statistics**

| | (1) # observations | (2) Mean | (3) Standard deviation | (4) p25 | (5) Median | (6) p75 |
|---|---|---|---|---|---|---|
| AUM (US$ bn) | 1,752 | 13.95 | 33.21 | 1.11 | 4.21 | 12.90 |
| Fee (bps) | 1,752 | 25.35 | 18.82 | 12.00 | 18.00 | 35.00 |
| Quoted spread (bps) | 1,752 | 7.61 | 9.08 | 2.44 | 4.56 | 8.61 |
| Effective spread (bps) | 1,752 | 4.68 | 6.33 | 1.61 | 2.64 | 4.78 |
| Realized spread (bps) | 1,752 | 3.19 | 5.84 | 0.36 | 1.23 | 3.32 |
| ETF turnover ratio | 1,750 | 5.89 | 8.60 | 1.26 | 2.17 | 5.81 |
| Transient investors (% of AUM) | 1,673 | 0.23 | 0.21 | 0.08 | 0.16 | 0.29 |
| Tax-insensitive investors (% of AUM) | 1,703 | 81.18 | 17.63 | 74.67 | 86.58 | 93.29 |
| Manager duration (quarters) | 1,719 | 5.27 | 0.92 | 4.87 | 5.49 | 5.85 |
| Time since inception (quarters) | 1,752 | 13.73 | 5.59 | 9.50 | 14.00 | 18.25 |
| Lending income (bps of AUM) | 1,752 | 3.14 | 6.07 | 0.12 | 0.92 | 3.26 |
| Marketing expense (bps) | 1,752 | 3.29 | 7.74 | 0.00 | 0.00 | 0.60 |
| Stock tweets (daily count) | 1,752 | 41.93 | 449.86 | 1.07 | 1.36 | 2.02 |
| Other expenses (%) | 1,752 | 3.23 | 5.09 | 0.00 | 0.00 | 4.00 |
| Fee waivers (%) | 1,752 | −0.61 | 2.62 | 0.00 | 0.00 | 0.00 |
| Tracking error (bps) | 1,752 | 28.78 | 36.55 | 10.04 | 15.25 | 31.97 |
| Performance drag (bps) | 1,752 | −1.87 | 6.18 | −2.15 | −1.17 | −0.48 |

This table reports summary statistics of our main panel consisting of ETF that compete with each other in tracking the same index. The unit of observation is ETF-quarter.

for each ETF-quarter. In the United States, Form 13F is filed quarterly by all institutional investment managers with at least $100 million in assets under management. Bushee (1998) uses k-means cluster analysis to categorize investors into quasi-indexers, transients, and dedicated investors. This analysis considers several factors: portfolio turnover (measured as the average absolute change in the institution's ownership positions over a quarter, adjusted for total equity changes), block ownership, diversification, and momentum strategy usage. Quasi-indexers correspond to long-term investors in our model due to their low turnover and high diversification. At the same time, transient investors, characterized by high turnover, correspond to short-term investors. As an additional measure of investor urgency, we also compute the average stock holding duration for each ETF using the methodology of Cremers and Pareek (2016).

Table 1 reports summary statistics. The average ETF in our sample is 3.5 years old (13.73 quarters), it has US$13.95 billion in assets under management, and it charges a fee of 25.35 basis points. Most assets under management are held by tax-insensitive investors (81.18% on average), and the proportion of shares held by transient investors is 23%. The average ETF is held in 13F portfolios for 5.27 quarters, a result that is quantitatively similar to the findings of Broman and Shum (2018).

## 2.2 Investor clienteles in same-index ETFs

The first prediction of our model states that high-fee ETFs attract investors with a shorter holding horizon. To test this conjecture empirically, we follow Broman and Shum (2018) and compute the fraction of short-term owners in each ETF-quarter, using the 13F institutional holding data and the definition of transient investors of Bushee (1998).[16] Since capital gains tax may reduce traders' propensity to switch between funds and therefore affect the strength of clientele effects, we also use Bushee's Institutional Investor Classification Data to label investors as tax sensitive or tax insensitive (Blouin, Bushee, and Sikes 2017).

To provide additional evidence, we also compute the Cremers and Pareek (2016) measure of investor holding duration from 13-F data. For each investor $j$ in our sample and quarter $T$, we determine the average holding time for a particular stock $i$, across all stocks in the investor portfolio, adjusted for buys and sells:

$$\text{Duration}_{i,j,T-1} = \frac{(W-1)H_{i,j}}{H_{i,j}+B_j} + \sum_{t=T-W+1}^{T-1} \frac{(T-t-1)\alpha_{i,j,t}}{H_{i,j}+B_{i,j}}, \tag{21}$$

where $t$ and $T$ are measured in quarters, $B_{i,j}$ represents the total percentage of shares bought by investor $j$ in stock $i$ between $T-W$ and $t-1$, $H_{i,j}$ is the percentage of shares outstanding of $i$ held by $j$ at $T-W$, and $\alpha_{i,j,t}$ is the percentage of shares outstanding of $i$ purchased ($\alpha > 0$) or sold ($\alpha < 0$) by $j$ between $t-1$ and $t$. Consistent with the literature, we choose a maximum horizon $W=12$ since any informational effects beyond that horizon are assumed to be small. Finally, we compute a weighted average of investor duration for each ETF-quarter in our sample, using weights proportional to the ETF ownership of each investor as stated in the 13F reports.

Figure 3 illustrates the average quarterly holdings for different investor types. First, tax-related behavior among ETF holders appears to be limited as, on average, 81.18% of AUM is held by tax-insensitive investors. Second, in line with our first empirical prediction, high-fee funds are 28% more likely to be held by transient investors than low-fee funds (25.25% of AUM and 19.75% of AUM for high- and low-fee ETFs, respectively). The AUM share of transient investors is a lower bound for their share of trading volume since short-term investors trade more frequently.

The bottom-two panels focus on the (continuous) measure of investor duration for different investor types. Transient investors have a lower holding duration than quasi-indexers, which points to the two alternative measures being consistent with each other. Tax-insensitive investors tend to have shorter holding horizons than tax sensitive investors, likely because they are able to adjust their portfolio more frequently without concerns about tax implications.

We formally test Prediction 1 by regressing the ratio of AUM held by transient investors on a dummy variable for the ETF charging the highest fee for a particular index, while controlling for tracking error, lending revenue,

[16] We download yearly investor classifications from https://accounting-faculty.wharton.upenn.edu/bushee/ and match them with quarterly 13F holdings data.

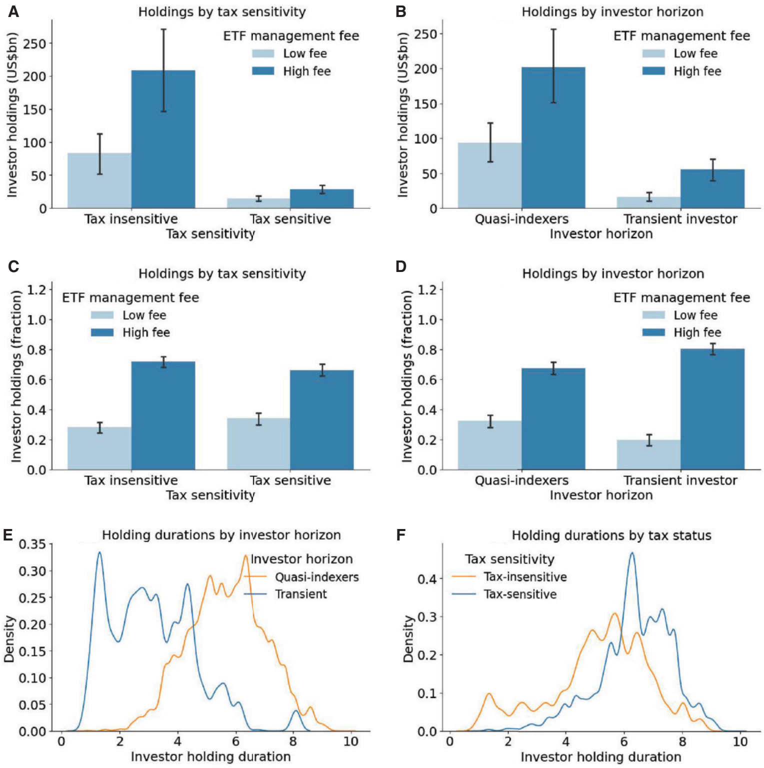


**Figure 3**
**Types of institutional investors in ETFs**
The top-two panels in this figure show the breakdown of ETF holdings between tax sensitive investors (TSI) and tax insensitive investors (TII), as well as between short-term (transient) and long-term (quasi-indexer) investors. We measure ETF holdings separately for high- and low-fee funds (High fee, Low fee) in a given index. The middle-two panels show the relative holding share of investor groups in high- and low-fee funds. The bottom-two panels illustrate the distribution of holding duration, computed as in Cremers and Pareek (2016), for the different investor types. The error bars correspond to 95% confidence intervals.

marketing fees and other expenses, the AUM, and age of the fund, as well as name recognition as proxied by the average daily count of social media investing platform StockTwits posts in a given quarter (following, e.g., Cookson and Niessner 2020; Lopez Avila, Martineau, and Mondria 2023). Since ETFs are a relatively new investment vehicle, many ETFs were introduced throughout our sample. We follow Broman and Shum (2018) and control for the ETF age and the average age of the investor base.

The results in Table 2 support Prediction 1. Using the estimate in the second column, which includes all control variables, the share of transient investors is 9-percentage-points larger for the high-fee ETF than for the low-fee ETF for a given index. The effect is statistically significant and economically large, representing a 50% difference in the holdings of short-term investors between high- and low-fee funds.

**Table 2**
**Transient investors and ETF fees**

| | AUM share of transient investors | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High fee | 0.05*** | 0.08*** | 0.05*** | 0.06*** | 0.05*** | 0.09*** |
| | (2.97) | (4.11) | (2.89) | (3.70) | (2.87) | (4.61) |
| Name recognition | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | | (0.79) | (0.80) | (0.66) | (0.51) | (0.37) |
| Log index AUM | | −0.10 | −0.10 | −0.10 | −0.11 | −0.09 |
| | | (−1.43) | (−1.44) | (−1.49) | (−1.56) | (−1.36) |
| Lending income | | −0.02** | | −0.03** | −0.03** | −0.02** |
| (bps of AUM) | | (−2.13) | | (−2.80) | (−2.70) | (−2.15) |
| Marketing expense (bps) | | 0.08*** | | 0.03** | 0.08*** | 0.03** |
| | | (3.68) | | (2.36) | (3.55) | (2.24) |
| Other net expenses | | 0.06** | | | 0.06*** | |
| | | (2.83) | | | (2.99) | |
| Tracking error (bps) | | 0.00 | | | | 0.00 |
| | | (0.19) | | | | (0.23) |
| Performance drag (bps) | | −0.00 | | | | −0.00 |
| | | (−0.81) | | | | (−0.92) |
| Unit investment trust | | −0.05 | | | | 0.02 |
| | | (−0.82) | | | | (0.44) |
| ETF age (quarters) | | −0.04* | | | | −0.04** |
| | | (−1.97) | | | | (−2.14) |
| Time since first position | | 0.03 | | | | 0.04 |
| | | (0.75) | | | | (0.87) |
| Constant | 0.20*** | 0.19*** | 0.20*** | 0.19*** | 0.20*** | 0.18*** |
| | (16.83) | (16.96) | (16.80) | (17.60) | (19.34) | (15.25) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,673 | 1,673 | 1,673 | 1,673 | 1,673 | 1,673 |
| Adjusted $R$-squared | .50 | .54 | .51 | .52 | .54 | .53 |

The unit of observation is ETF-quarter. The dependent variable is the share of assets under management (AUM) held by transient investors. *High fee* is a dummy variable that is one if the ETF has the highest management fee among competing same-index ETFs. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Log index AUM* is the logarithm of the aggregate AUM for all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs, such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average difference in daily returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. *ETF age* is the number of quarters since inception. *Time since first position* is the time (in quarters) since the ETF was first purchased by an investor, averaged across all investors with weights given by their 13F dollar position. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Figure 4 graphically illustrates the clientele effects documented in Table 2. After accounting for all control variables, high-fee ETFs attract a higher share of transient investors than low-fee ETFs for a given index.

We repeat our analysis using the average investor duration measure defined in Equation (21) as a proxy for the investor horizon. A drawback of investor duration for our purposes is that, since it is based on quarterly 13F reports, it does not take into account any trading activity within a quarter and therefore might not accurately proxy for the short-term liquidity needs of higher-frequency traders. On the other hand, the duration measure enables us to further identify the clientele effect within different investor groups. We separately

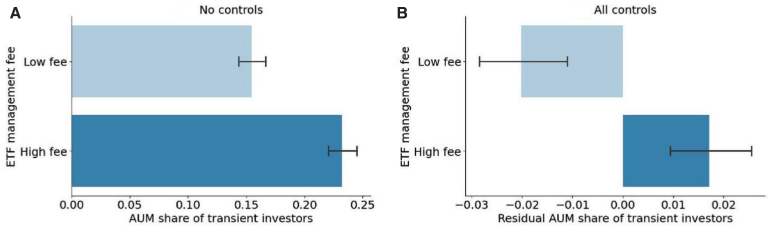


**Figure 4**
**ETF investor clienteles**
This figure plots the average residual AUM share of transient investors (panel A) and investor holding duration (panel B) between 2016Q1 and 2020Q4 for low- and high-fee ETFs. The residuals are obtained after controlling for index and quarter fixed effects, log AUM for the index, lending income, marketing and other expenses, and name recognition as proxied by StockTwits message count, tracking error, performance drag, ETF age, and the age of the investor base. The error bars correspond to 95% confidence intervals.

compute the average holding duration for tax-insensitive and tax-sensitive investors. Our model in Section 1 suggests that the liquidity channel should be driven by tax-insensitive investors, since they are able to switch in and out of positions without triggering a tax event.

Table 3 presents the results. Consistent with our prediction, tax-insensitive investors in high-fee ETFs have a shorter investment horizon. Using the estimates in column 3, the holding duration of investors in high-fee ETFs is 3.64% (0.27/5.49) lower than for investors in low-fee ETFs, after including all controls. If the clientele effect is primarily driven by turnover differences, then controlling for turnover should eliminate or substantially diminish the clientele effect. In column 3, we include turnover as a control variable and find that the difference in investor duration across funds is halved (coefficient drops from −0.27 to −0.13) and is no longer statistically significant. This finding is consistent with our proposed channel that liquidity differences across funds drive investor clienteles.

Columns 5–7 and 8–11 in Table 3 repeat the analysis above for the subsample of tax-sensitive investors and the entire investor pool, respectively. Tax-sensitive investors in the high fee ETFs maintain their position for 0.23 quarters *longer* than their peers in low fee ETFs. Since high-fee ETFs tend to be older than low-fee ETFs for a given index, the results are consistent with a capital lock-in effect, where tax sensitive investors are discouraged to switch to low-fee ETFs as doing so would trigger a costly tax event. However, since tax-sensitive investors hold on average less than one-fifth of AUM for ETFs in our sample, the liquidity channel dominates the capital lock-in effect. Across the aggregate pool of investors, high-fee ETFs attract shorter-term investors, but the effect is not statistically significant in all regression specifications.

## 2.3 Fees and liquidity in the cross-section of ETFs

The model of ETF competition in Section 1 implies that for a given index, the ETF charging a higher fee enjoys better liquidity relative to the low-fee

**Table 3**
Investor holding duration and ETF fees

| | Investor holding duration | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Tax-insensitive investors | | | | Tax-sensitive investors | | | All investors | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| High fee | −0.37** | −0.38** | −0.27** | −0.13 | 0.23** | 0.19** | 0.16* | −0.18 | −0.17* | −0.15 | −0.04 |
| | (−2.62) | (−2.79) | (−2.32) | (−1.36) | (2.64) | (2.35) | (1.79) | (−1.56) | (−1.79) | (−1.56) | (−0.56) |
| ETF turnover | | | | −0.36*** | | | | | | | −0.24*** |
| | | | | (−3.70) | | | | | | | (−3.56) |
| Name recognition | | −0.06** | −0.05 | 0.02* | | −0.03*** | −0.04 | | −0.06** | −0.04 | 0.00 |
| | | (−2.61) | (−1.67) | (1.89) | | (−2.94) | (−1.65) | | (−2.40) | (−1.43) | (0.27) |
| Log index AUM | | 0.33 | 0.33 | 0.20 | | −0.13 | −0.12 | | 0.08 | 0.08 | −0.01 |
| | | (0.94) | (0.94) | (0.53) | | (−0.45) | (−0.44) | | (0.27) | (0.29) | (−0.05) |
| Lending income | | 0.02 | 0.01 | −0.01 | | −0.02 | −0.02 | | 0.05 | 0.05 | 0.03 |
| (bps of AUM) | | (0.38) | (0.26) | (−0.16) | | (−0.51) | (−0.46) | | (1.20) | (1.18) | (0.80) |
| Marketing | | −0.29*** | −0.55*** | −0.38*** | | −0.27*** | −0.20 | | −0.32*** | −0.38*** | −0.27*** |
| expense (bps) | | (−3.92) | (−4.20) | (−3.54) | | (−4.86) | (−1.69) | | (−6.35) | (−4.27) | (−3.28) |
| Other net expenses | | | −0.29* | −0.16 | | | 0.07 | | | −0.06 | 0.02 |
| | | | (−2.02) | (−1.42) | | | (0.60) | | | (−0.69) | (0.32) |
| Tracking error | | | −0.08 | 0.02 | | | 0.07* | | | −0.04 | 0.04 |
| (bps) | | | (−0.94) | (0.39) | | | (2.00) | | | (−0.74) | (0.92) |
| Performance drag | | | −0.01 | −0.01 | | | 0.02 | | | −0.01 | −0.00 |
| (bps) | | | (−0.38) | (−0.33) | | | (1.37) | | | (−0.47) | (−0.14) |
| Unit investment | | | −0.20 | −0.27 | | | 0.15 | | | −0.21 | −0.26 |
| trust | | | (−0.66) | (−1.37) | | | (0.58) | | | (−0.69) | (−1.14) |

(continued)

**Table 3**
Continued

| | Investor holding duration | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Tax-insensitive investors | | | | Tax-sensitive investors | | | All investors | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| ETF age (quarters) | 0.42*** | 0.39*** | 0.39*** | 0.48*** | 0.03 | 0.05 | 0.04 | 0.32*** | 0.29*** | 0.29*** | 0.35*** |
| | (3.75) | (3.70) | (3.64) | (4.61) | (0.45) | (0.73) | (0.63) | (3.57) | (3.63) | (3.60) | (4.36) |
| Time since first position | −0.08 | −0.39 | −0.37 | −0.32 | 0.29** | 0.01 | 0.01 | −0.12 | −0.46* | −0.46* | −0.42* |
| | (−0.34) | (−1.40) | (−1.32) | (−1.20) | (2.67) | (0.09) | (0.12) | (−0.58) | (−2.06) | (−2.05) | (−1.94) |
| Constant | 5.53*** | 5.54*** | 5.49*** | 5.42*** | 5.76*** | 5.78*** | 5.80*** | 5.35*** | 5.35*** | 5.34*** | 5.30*** |
| | (67.09) | (74.57) | (82.84) | (103.08) | (123.99) | (112.03) | (108.41) | (82.38) | (101.38) | (105.14) | (132.05) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,703 | 1,703 | 1,703 | 1,701 | 1,596 | 1,596 | 1,596 | 1,719 | 1,719 | 1,719 | 1,717 |
| Adjusted $R$-squared | .31 | .34 | .36 | .40 | .41 | .45 | .45 | .41 | .47 | .47 | .49 |

The unit of observation is ETF-quarter. The dependent variable is the investor holding duration, defined in Cremers and Pareek (2016) and Equation (21), and averaged across all investors holding the given ETF in a given quarter. *High fee* is a dummy variable that is one if the ETF has the highest fee among competing same-index ETFs. *ETF turnover* is the annualized secondary market traded dollar volume divided by the ETF's AUM. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Log index AUM* is the logarithm of the aggregate AUM for all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs, such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average difference in daily returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. *ETF age* is the number of quarters since inception. *Time since first position* is the time (in quarters) since the ETF was first purchased by an investor, averaged across all investors with weights given by their 13F dollar position. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. * $p$ <.1; ** $p$ <.05; *** $p$ <.01.

ETF (Prediction 2). Moreover, high-fee ETFs are also expected to have higher secondary market dollar volume and turnover (dollar volume divided by AUM) as they attract the higher-turnover clientele who care more about liquidity than fees (Prediction 3). Finally, if the aggregate mass of short-term investors is larger than the mass of long-term investors, high-fee funds are expected to command both larger market shares and higher profits.

To formally test the predictions and quantify the liquidity and volume gap between high- and low-fee ETFs, we estimate OLS regressions of ETF liquidity, turnover, volume, market share, and profits on a dummy variable for the ETFs that have the highest fee among the same-index competitors. We control for the total AUM of ETFs tracking the index, marketing expenses, name recognition, security lending income, tracking error, performance drag, the AUM share of tax-insensitive investors, and past performance of the fund (lagged ETF return):

$$y_{i,t} = d_{\text{High fee},i,t} + \text{Controls}_{i,t} + \text{error}_{i,t}, \tag{22}$$

where $i$ runs over ETFs and $t$ indexes quarters. The dependent variable $y$ stands for, in turn, fund management fee, quoted spread, log dollar volume, secondary market turnover, log profit, and AUM market share.

The regression results in Table 4 support Predictions (2) and (3), as well as documenting that high-fee funds enjoy larger assets under management and profits. Model (1) shows that fee differences between funds tracking the same index are substantial: high-fee ETFs charge 54.75% more than their low-fee counterparts (an 10.95-bps premium over an average low fee of 20.20 bps).

In line with Prediction 2, high-fee ETFs have a 3.02-bps narrower bid-ask spread (i.e., 42.65% lower) than low-fee ETFs, keeping the index and year-month fixed and controlling for other variables (including tracking error, ETF lending income, and marketing expenses). This result is statistically significant at the 1% level and economically meaningful. It suggests that investors value ETF liquidity and are willing to pay for it. The results in Table 4 also suggest that high-fee ETFs tend to have larger trading volumes (Model 3) and turnover (Model 4) that empirically support Prediction 3. Further, we document that high-fee ETFs enjoy a 38-percentage-points larger market share than low-fee ETFs tracking the same index (Model 5). The result implies a 69%–31% market share split, which translates into higher profits for high-fee funds (Model 6).

Figure 5 illustrates that the cross-sectional predictions are supported by the data. The figure plots the average bid-ask spread, market share, log trading volume, and turnover ratio for low- and high-fee same-index ETFs. The top panels show raw differences (i.e., without controls), whereas the bottom panels plot residuals from a regression model that accounts for all controls in Table 4, including index and quarter fixed effects. The figure provides salient evidence that high-fee fund enjoy higher market shares, have lower bid-ask spreads and higher secondary market turnover and trading volumes.

**Table 4**
**Fees and liquidity for competing ETFs**

| | Fee (1) | Quoted spread (2) | log dollar volume (3) | Turnover (4) | Market share (5) | log profit (6) |
|---|---|---|---|---|---|---|
| High fee | 10.95*** | −3.02*** | 1.99*** | 5.45*** | 0.38*** | 1.71*** |
| | (8.55) | (−4.15) | (6.63) | (5.79) | (5.48) | (7.52) |
| Name recognition | −0.45 | −0.17** | 0.09 | 1.70*** | 0.02 | 0.01 |
| | (−1.24) | (−2.22) | (1.34) | (4.32) | (1.01) | (0.30) |
| Log index AUM | −0.05 | −6.09*** | 1.57*** | −2.69* | 0.03*** | 1.60*** |
| | (−0.28) | (−3.40) | (36.94) | (−2.03) | (3.53) | (19.56) |
| Lending income (bps of AUM) | 1.45** | −2.23*** | 0.52*** | 0.23 | 0.12*** | 0.56*** |
| | (2.26) | (−6.21) | (3.47) | (0.86) | (3.80) | (5.30) |
| Marketing expense (bps) | −3.35** | 0.20 | −0.16 | 3.13** | −0.09 | −0.78*** |
| | (−2.56) | (0.25) | (−0.45) | (2.77) | (−1.20) | (−2.97) |
| Other net expenses | −0.58 | −1.88** | 0.15 | 2.71** | 0.01 | −0.24 |
| | (−0.40) | (−2.38) | (0.40) | (2.34) | (0.14) | (−0.78) |
| Tracking error (bps) | 4.51*** | −1.57** | 0.27* | 2.68*** | −0.03 | 0.16 |
| | (3.59) | (−2.42) | (1.77) | (3.97) | (−1.15) | (1.53) |
| Performance drag (bps) | 1.57** | −0.21 | 0.05 | 0.17 | −0.02* | 0.03 |
| | (2.32) | (−0.66) | (0.90) | (1.00) | (−1.83) | (0.74) |
| Unit investment trust | 2.47 | 4.63*** | −0.59 | −0.31 | −0.51** | −0.49 |
| | (0.64) | (4.11) | (−0.77) | (−0.08) | (−2.64) | (−1.41) |
| Tax-insensitive investors (TII) | −0.00 | 0.02 | 0.00 | 0.05*** | −0.00 | −0.00 |
| | (−0.23) | (1.32) | (0.11) | (3.29) | (−0.70) | (−1.11) |
| Lagged ETF return | 1.83 | 2.50** | −1.45 | −7.06 | 0.27*** | 1.40 |
| | (0.35) | (2.59) | (−0.91) | (−1.29) | (8.76) | (1.48) |
| TII × Lagged ETF return | −0.01 | −0.03 | 0.01 | 0.06 | −0.00*** | −0.02* |
| | (−0.14) | (−1.16) | (0.41) | (0.63) | (−8.80) | (−1.92) |
| Constant | 20.20*** | 7.08*** | 16.44*** | −0.54 | 0.37*** | 24.37*** |
| | (12.97) | (5.47) | (37.03) | (−0.41) | (4.23) | (70.05) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,703 | 1,703 | 1,703 | 1,701 | 1,703 | 1,701 |
| Adjusted $R$-squared | .89 | .65 | .72 | .61 | .41 | .77 |

The unit of observation is an ETF-quarter. The dependent variables in the regressions are (1) ETF fee (management expense ratio) in bps, (2) Quoted bid-ask spread (in bps), (3) Log dollar volume, (4) Turnover as a fraction (annualized secondary market traded dollar volume divided by the ETF's AUM), (5) Market share as a fraction (AUM of the ETF divided by the total AUM of all ETFs tracking the given index), and (6) log profitability (AUM times fee). *Quoted spread* (in bps) is computed as the time-weighted bid-ask spread in the ETF secondary market divided by the ETF's midpoint price. *High fee* is a dummy variable that is one if the ETF has the highest management fee among competing same-index ETFs. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Log index AUM* is the logarithm of the aggregate AUM for all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs, such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average difference in daily returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts.*Tax-insensitive investors* is the share of AUM held by tax-insensitive investors for a given index and quarter. *Lagged ETF return* is the percentage change in ETF price from the previous quarter. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. *$p<.1$; **$p<.05$; ***$p<.01$.

We include three robustness checks in Appendix E. First, Table E3. replicates the Table 4 analysis, but exclusively for the subsample of ETFs tracking the same benchmark. Second, Table E4. shifts focus to cross-sectional regressions, employing ETF-level averages across time for all dependent and control variables in Table 4, as well as the ratio of transient investors to replicate the

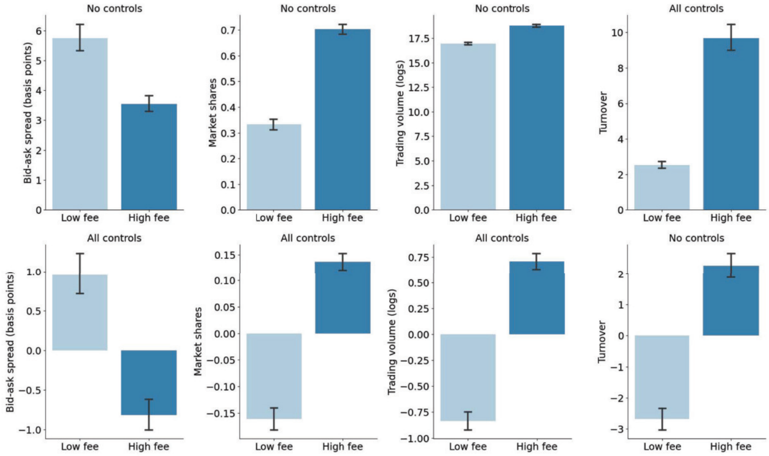


**Figure 5**
**Liquidity in the cross-section of competing ETFs**
This figure plots the average bid-ask spread, market share, trading volume, and turnover for high- and low-fee same-index ETFs. The top panels illustrate raw averages, that is, without including control variables. The bottom panels plot residuals that are obtained after controlling for index and quarter fixed effects, as well as all control variables in Table 4. The error bars correspond to 95% confidence intervals.

clientele effect from Table 2. Lastly, Table E5. reevaluates the regression model in Table 4 using a broader range of liquidity measures, including quoted spreads measured in dollars, as well as effective and realized spreads, measured in both basis points and U.S. dollars.

**2.3.1 The contribution of liquidity to ETF competition: A Shapley-Owen $R^2$ decomposition** To pinpoint the contribution of the liquidity channel to explaining differences between competing ETFs, we regress their market shares and fees on liquidity and nonliquidity measures and use the Shapley-Owen methodology (Owen 1977) to decompose the $R^2$ of the regression models. The Shapley-Owen method can be implemented on a model with $k$ independent variables $\Omega=\{x_1,x_2,\ldots,x_k\}$ by estimating $2^k$ regressions, one for each possible subset of independent variables. The marginal contribution of independent variable $j$ (i.e., its Shapley value) to the regression $R^2$ can then be computed as:

$$R_j^2=\sum_{S\subseteq\Omega-\{x_j\}}\frac{R^2\left(S\cup\{x_j\}\right)-R^2(S)}{k\binom{k-1}{|S|}}, \tag{23}$$

where $R^2(S)$ is the $R^2$ of a regression on the variables in $S$, and $|S|$ is the element count of $S$.

We partition the independent variables into four sets: (1) liquidity-related variables (quoted, effective, and realized spreads, respectively, turnover), (2) "reputation" related variables (marketing expenses, other expenses including distribution costs, net of fee waivers used to retain investors, and name recognition variables proxied by the ETF's social media footprint), (3) tax-related variables including the share of tax-insensitive investors, and (4) other miscellaneous variables including lending income, tracking error, and performance drag.

To report a conservative estimate for the contribution of the liquidity channel, we only include one liquidity variable at a time in the regression, whereas we always add all available controls. Doing so will tend to understate the role of liquidity compared to the alternative channels, but prevents incorrectly overstating the importance of liquidity.

Table 5 reports the decomposition results. We find that liquidity is the most important determinant of market share differentials, accounting for two-thirds (67.65% in column 1) of the explained variance in market shares, followed by reputation-related variables (19.95%). At the same time, liquidity (as a lower bound) is the second-most important determinant of ETF fees, accounting for 29.91% (column 5) of the explained variance. Interestingly, the fraction of tax-insensitive investors does not explain a large fraction of either market shares or fees (between 0.75% and 1.96% across all specifications); this is potentially because the share of tax-sensitive investors across ETFs is rather low.

The liquidity variables still account for a smaller portion of the variance in fees than do market shares. This result can be attributed to the nature of market shares, which are bounded between zero and one and primarily reflect investor decisions for one ETF or another. Conversely, management fees are unbounded and influenced by a broader array of factors. For instance, ETFs on indices that are harder to track often command higher fees, contributing to greater variation in fees. In fact, Table 5 shows that tracking error accounts for approximately 30% of the $R^2$ in the fee regressions.

In Table E6. from Appendix E, we conduct a robustness check by replicating Table 5 using quoted, effective, and realized spreads in US$ cents instead of basis points. The findings mirror the main analysis in terms of coefficient sign, magnitude, and significance, but reveal that liquidity measures in cents explain less of the market share and fee variability. However, spread measured in dollars are a particularly noisy measure of liquidity in a cross-section of assets with different price levels. This issue is also pertinent for ETFs tracking the same index, since they may implement different NAV scaling factors. Additionally, stock splits or reverse splits, among other practices, can further skew dollar spread measurements and complicate cross-asset liquidity comparisons.

**2.3.2 Investor urgency and ETF competition** Prediction 4 states that the spread and fee differentials increase in the share of high-turnover investors in an index. To test this prediction, we compute an index-level investor urgency measure as the share of transient investors for the given index each quarter

**Table 5**
Liquidity channel versus alternatives

| | Market share | | | | Fee | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Liquidity $R^2$ contribution** | **67.65%** | **64.84%** | **61.40%** | **44.59%** | **29.91%** | **27.44%** | **26.71%** | **21.35%** |
| Quoted spread (bps) | −0.03*** | | | | −0.41*** | | | |
| | (−5.34) | | | | (−3.85) | | | |
| Effective spread (bps) | | −0.03*** | | | | −0.53*** | | |
| | | (−4.61) | | | | (−3.49) | | |
| Realized spread (bps) | | | −0.03*** | | | | −0.54*** | |
| | | | (−3.81) | | | | (−4.46) | |
| ETF turnover | | | | 0.02*** | | | | 0.32* |
| | | | | (5.23) | | | | (1.99) |
| **Reputation $R^2$ contribution** | **19.95%** | **22.28%** | **24.40%** | **33.45%** | **27.98%** | **29.22%** | **29.28%** | **35.88%** |
| Name recognition | 0.02 | 0.02 | 0.02 | −0.01 | −0.29 | −0.24 | −0.21 | −0.71* |
| | (1.09) | (1.24) | (1.36) | (−0.79) | (−1.32) | (−1.15) | (−1.02) | (−1.69) |
| Marketing expense (bps) | −0.00 | −0.00 | −0.00 | −0.01 | −0.19 | −0.14 | −0.14 | −0.34 |
| | (−0.37) | (−0.00) | (−0.05) | (−1.07) | (−0.75) | (−0.51) | (−0.53) | (−1.03) |
| Other net expenses | 0.68 | 1.09 | 1.11 | 0.51 | 20.97 | 27.38 | 27.03 | 18.20 |
| | (0.67) | (1.07) | (1.07) | (0.42) | (0.79) | (1.01) | (0.98) | (0.53) |
| **TII $R^2$ contribution** | **0.75%** | **0.81%** | **0.87%** | **1.01%** | **1.91%** | **1.95%** | **1.96%** | **1.68%** |
| Tax-insensitive investors (TII) | 0.00 | 0.00 | 0.00 | −0.00 | 0.02 | 0.02 | 0.02 | −0.00 |
| | (0.34) | (0.26) | (0.24) | (−1.17) | (0.93) | (0.90) | (0.90) | (−0.13) |

(continued)

**Table 5**
*Continued*

| | Market share | | | | Fee | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Other $R^2$ contribution** | **11.63%** | **12.05%** | **13.32%** | **20.93%** | **40.18%** | **41.37%** | **42.04%** | **41.07%** |
| Lending income (bps of AUM) | 0.01** | 0.01* | 0.01* | 0.02*** | 0.11 | 0.10 | 0.11 | 0.24 |
| | (2.23) | (1.78) | (1.97) | (3.56) | (0.36) | (0.30) | (0.34) | (0.85) |
| Tracking error (bps) | −0.00 | −0.00 | −0.00 | −0.00 | 0.14*** | 0.14*** | 0.15*** | 0.13*** |
| | (−1.00) | (−0.78) | (−0.42) | (−1.30) | (2.73) | (2.75) | (2.91) | (3.63) |
| Performance drag (bps) | −0.00 | −0.00 | −0.00 | −0.00 | 0.29** | 0.30** | 0.30** | 0.29** |
| | (−1.38) | (−0.99) | (−1.02) | (−0.70) | (2.15) | (2.27) | (2.29) | (2.48) |
| Unit investment trust | −0.21 | −0.26 | −0.27 | −0.35* | 10.11** | 9.33** | 9.19** | 7.93 |
| | (−0.75) | (−0.89) | (−0.93) | (−1.73) | (2.25) | (2.14) | (2.08) | (1.40) |
| Constant | 0.00 | 0.00 | 0.00 | 0.00** | 0.01 | 0.00 | 0.00 | 0.04 |
| | (0.34) | (0.04) | (0.25) | (2.45) | (0.24) | (0.01) | (0.07) | (1.36) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,703 | 1,703 | 1,703 | 1,701 | 1,703 | 1,703 | 1,703 | 1,701 |
| Adjusted $R$-squared | 0.35 | 0.33 | 0.31 | 0.30 | 0.38 | 0.38 | 0.38 | 0.37 |

This table reports the Shapley-Owen (Owen 1977) relative contributions to ETF market shares and fees, computed as in Equation (23). The unit of observation is an ETF-quarter. The dependent variables in the regressions are (1) Market share (as a fraction) and (2) ETF fee (management expense ratio) in bps. *Quoted spread* (in bps) is the time-weighted bid-ask spread in the ETF secondary market divided by the ETF's midpoint price. *Effective spread* (in bps), computed as the time-weighted average of the signed difference between trade price and midpoint. *Realized spread* (in bps), computed as the time-weighted average of the signed difference between trade price and the midpoint five minutes after the trade. *ETF turnover* is the annualized daily dollar volume in the ETF's secondary market divided by AUM. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* are taken from ETF Global and capture ETF costs, such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tax-insensitive investors* is the AUM share held by tax-insensitive investors for a given index and quarter. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average daily difference in returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. Standard errors are clustered by quarter. $t$-statistics are reported in parentheses. *$p$ <.1; **$p$ <.05; ***$p$ <.01.

(which we label $\mathrm{TRA_{ix}}$).[17] Next, we regress quoted bid-ask spreads and fees on (a) a high-fee ETF dummy, (b) the dummy's interaction with the investor urgency measure, and (c) a standard set of control variables (same as in previous regressions), where $i$ runs over ETFs and $t$ indexes quarters:

$$y_{i,t} = \beta_1 d_{\text{High fee},i,t} + \beta_2 d_{\text{High fee},i,t} \mathrm{TRA_{ix}} + \beta_3 \mathrm{TRA_{ix}} + \mathrm{Controls}_{i,t} + \mathrm{error}_{i,t}, \quad (24)$$

The results in Table 6 are in line with our theoretical prediction. First, from Model (3), a one-standard-deviation increase in the share of transient investors increases the difference between high- and low-fee ETF bid-ask spreads increases from 3.21 bps to 5.69 bps. Second, a one-standard-deviation increase in investor urgency leads to a 3.35 bps, or 30%, wider difference in the fees of high- and low-fee ETFs (Model 6).

### 2.4 What drives investor coordination?

In this section, we explore why certain ETFs become "high-fee" liquid funds within their indices. Prediction 5 suggests that coordination failures confer a first-mover advantage on the incumbent ETF. The oldest fund in an index may charge higher fees because short-term investors cannot coordinate to collectively switch over to a less expensive fund. Alternatively, a high-fee ETF may have a superior reputation, driven, for example, by increased marketing efforts or stronger brand recognition. We also examine product quality differences. The higher fees could reflect better index tracking (a lower tracking error or performance drag) or additional returns from securities lending.

We estimate a linear probability model, in which we regress a dummy variable indicating whether an ETF is the high-fee fund for a given index on a first-mover dummy variable that takes the value one for the oldest ETF tracking the index. We include controls for marketing expenses, name recognition, lending income, the AUM share of tax-insensitive investors, tracking error, and performance drag.

Table 7 reports the results. We find that being the first ETF to track a particular index is by far the strongest predictor of the fund charging a high fee. Approximately 80% of high-fee ETFs are the first movers for their respective index consistent with incumbency playing a key role in the competitive dynamics of ETFs, consistent with the model.

By itself, the first-mover advantage explains 35% of the variation between high- and low-fee ETFs. Once we account for the other controls, the model $R^2$ only increases to 45%. We find a positive relation between an ETF's fee and their name recognition and other expenses, which can include marketing and distribution costs not captured by 12b-1 fees. At the same time, high-fee funds tend to have larger tracking errors and creation fees.

[17] From Table 3, investor holding duration is strongly correlated with the ETF age. Since competing same-index ETFs launch at different times, it is not straightforward to aggregate holding duration at the index level.

**Table 6**
**Investor urgency and differences in liquidity, market shares, and fees**

| | Quoted spread | | | Fee | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High fee | −4.21*** | −3.46*** | −3.21*** | 12.50*** | 11.83*** | 11.19*** |
| | (−5.62) | (−5.05) | (−4.44) | (8.59) | (7.97) | (8.82) |
| High fee × $TRA_{ix}$ | −2.95** | −2.77** | −2.48** | 3.60* | 3.94** | 3.35** |
| | (−2.63) | (−2.62) | (−2.29) | (2.06) | (2.73) | (2.14) |
| Index AUM share of transient investors ($TRA_{ix}$) | 2.62*** | 1.91** | 1.85** | −2.03** | −1.80** | −1.80** |
| | (3.14) | (2.62) | (2.39) | (−2.42) | (−2.48) | (−2.17) |
| Name recognition | | 0.16* | −0.05 | | −0.40** | −0.54 |
| | | (2.06) | (−0.81) | | (−2.44) | (−1.36) |
| log index AUM | | −7.03*** | −6.01*** | | 0.31* | 0.19 |
| | | (−4.14) | (−3.24) | | (2.00) | (0.80) |
| Lending income (bps of AUM) | | −2.00*** | −2.02*** | | 1.31 | 1.17* |
| | | (−6.92) | (−6.42) | | (1.52) | (1.78) |
| Marketing expense (bps) | | 1.14 | 0.80 | | −4.61*** | −4.20*** |
| | | (1.30) | (0.93) | | (−3.07) | (−3.11) |
| Other net expenses | | −1.02 | −1.36 | | −1.03 | −1.27 |
| | | (−1.23) | (−1.57) | | (−0.65) | (−0.84) |
| Tracking error (bps) | | | −1.09* | | | 3.85*** |
| | | | (−1.98) | | | (3.04) |
| Performance drag (bps) | | | −0.13 | | | 1.48** |
| | | | (−0.44) | | | (2.31) |
| Unit investment trust | | | 3.18** | | | 4.44 |
| | | | (2.76) | | | (0.98) |
| Tax-insensitive investors | | | 0.03* | | | −0.02 |
| | | | (1.95) | | | (−0.93) |
| Lagged ETF return | | | 3.77** | | | 0.50 |
| | | | (2.64) | | | (0.16) |
| TII × Lagged ETF return | | | −0.05 | | | 0.00 |
| | | | (−1.66) | | | (0.02) |
| Constant | 9.72*** | 9.33*** | 6.50*** | 19.09*** | 19.44*** | 21.01*** |
| | (19.92) | (21.24) | (5.20) | (19.31) | (20.51) | (12.99) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,752 | 1,752 | 1,703 | 1,752 | 1,752 | 1,703 |
| Adjusted $R$-squared | .62 | .68 | .67 | .86 | .89 | .90 |

The unit of observation is an ETF-quarter. The dependent variables in the regressions are (1) the quoted spread (in bps), computed as the time-weighted bid-ask spread in the ETF secondary market divided by the ETF's midpoint price, and (2) ETF fee (management expense ratio) in bps. *High fee* is a dummy variable for the ETF that has the highest fee among same-index competitors. $TRA_{ix}$ is the share of AUM held by transient investors across all ETFs tracking a particular index. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Log index AUM* is the logarithm of the aggregate AUM of all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs, such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the difference in average daily returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. *Tax-insensitive investors* is the AUM share held by tax-insensitive investors for a given index and quarter. *Lagged ETF return* is the percentage change in ETF price from the previous quarter. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. *$p$ <.1; **$p$ <.05; ***$p$ <.01.

We do not find a significant difference between the share of tax-sensitive investors in high- and low-fee funds. In particular, we do not find support for a channel in which the high-fee ETF is able to charge a higher fee due to a capital lock-in effect where long-term investors do not switch to the cheaper ETF to avoid a taxable event. Furthermore, such a channel would be inconsistent

**Table 7**
**First-mover advantage: Linear probability model**

| | High fee fund | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| First mover | 0.62*** | 0.57*** | 0.60*** | 0.53** | 0.42** |
| | (7.63) | (5.49) | (4.50) | (2.58) | (2.31) |
| First mover × Different benchmarks | | | −0.06 | | |
| | | | (−0.36) | | |
| Marketing expense (bps) | | 0.14 | 0.15 | 0.15 | 0.14 |
| | | (1.33) | (1.40) | (0.98) | (1.37) |
| Other net expenses | | 0.15 | 0.16 | 0.16 | 0.14 |
| | | (1.38) | (1.52) | (1.11) | (1.39) |
| Name recognition (Twitter msg.) | | 0.02* | 0.01 | 0.01 | 0.02* |
| | | (2.09) | (1.57) | (1.38) | (2.06) |
| Lending income (bps of AUM) | | −0.06 | −0.06 | −0.06 | −0.05 |
| | | (−0.89) | (−0.83) | (−0.97) | (−0.64) |
| Tracking error (bps) | | 0.07** | 0.07** | 0.06** | 0.07** |
| | | (2.57) | (2.33) | (2.26) | (2.67) |
| Creation fee | | 0.24** | 0.24** | 0.25** | 0.22** |
| | | (2.76) | (2.79) | (2.71) | (2.57) |
| Performance drag (bps) | | 0.01 | 0.01* | 0.01 | 0.01 |
| | | (0.95) | (1.93) | (0.81) | (0.94) |
| Different benchmarks | | −0.40*** | −0.38*** | −0.40*** | −0.37*** |
| | | (−8.32) | (−7.12) | (−8.07) | (−5.41) |
| Different lead market-maker | | 0.00 | 0.00 | 0.00 | −0.09 |
| | | (0.15) | (0.11) | (0.16) | (−1.00) |
| Tax-insensitive investors | | 0.02 | 0.02 | 0.02 | 0.02 |
| | | (0.88) | (0.91) | (0.90) | (0.95) |
| First mover × Major brand index | | | | 0.05 | |
| | | | | (0.19) | |
| First mover × Different lead market-maker | | | | | 0.18 |
| | | | | | (1.07) |
| Constant | 0.20*** | 0.45*** | 0.44*** | 0.45*** | 0.50*** |
| | (3.62) | (12.37) | (12.40) | (12.13) | (7.86) |
| Index FE | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,752 | 1,703 | 1,703 | 1,703 | 1,703 |
| Adjusted $R$-squared | .35 | .45 | .45 | .45 | .45 |

The unit of observation is an ETF-quarter.The dependent variable is a dummy variable for the highest fee ETF among same-index competitors. *First mover* is a dummy variable indicating whether the ETF was the first to track the given index. *Different benchmarks* is a dummy variable that is one if the competing ETFs track different benchmarks closely related to the same index. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* are taken from ETF Global and capture ETF costs, such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Creation fee* is the ETF creation-redemption fee, in bps. *Performance drag* (in bps) is the difference in average daily returns of the ETF and its underlying index. *Different lead market-maker* is a dummy variable for whether the two ETFs employ a different lead market-maker. *Tax-insensitive investors* is the AUM share held by tax-insensitive investors for a given index and quarter. *Major brand index* is a dummy variable that is one if the tracked index is licensed by FTSE, Russell, S&P, NASDAQ, MSCI, or Dow Jones. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. *$p<.1$; **$p<.05$; ***$p<.01$.

with the fact that high-fee ETFs exhibit high trading activity and significant turnover, which would imply a large number of taxable events.

We emphasize two suggestive additional cross-sectional findings, corresponding to Models 3 and 4 in Table 7, that relate to the magnitude of the incumbents' advantage. First, the first mover is 6-percentage-points more likely

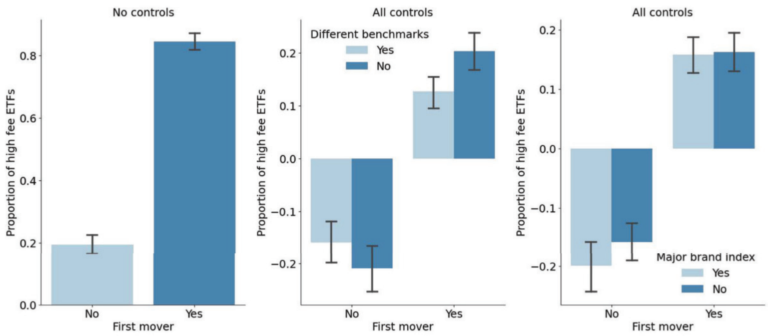


**Figure 6**
**First-mover advantage**
This figure illustrates the proportion of high-fee ETFs that are first to track a given index versus subsequent ETFs tracking the index. The left panel shows the unconditional share of high-fee ETFs, whereas the center and right panels plot the residual share of high-fee ETFs after controlling for all variables in Table 7. Additionally, the center (right) panel separates ETFs that track the same versus slightly different benchmark (track a major brand index or not). A “major brand index” is one licensed by FTSE, Russell, S&P, NASDAQ, MSCI, or Dow Jones. The error bars correspond to 95% confidence intervals.

to become the high-fee fund if the two ETFs track the same benchmark. The result is not statistically significant, however the sign of the coefficient is intuitive: if the entrant ETF chooses a similar, yet distinct benchmark, it might be able to distinguish itself from its competitors and attract more investors. In Model (4) of Table 7, we find a similar (yet statistically not significant) effect if the two ETFs track a “major brand index,” that is, one licensed by FTSE, Russell, S&P, NASDAQ, MSCI, or Dow Jones. To the extent tracking a major index translates to the two ETFs being more similar to each other, this should enhance the first-mover advantage.

Figure 6 visualizes the magnitudes of the results in Table 7. It illustrates (a) a strong first-mover advantage in that 80% of high-fee funds are incumbents in tracking a particular index, and (b) the first-mover advantage is somewhat stronger for ETFs tracking the exact same benchmark and for those tracking an index licensed by a major brand.

Different sponsors appear to “specialize” as either high-fee first movers or low-fee entrants, with BlackRock often taking on the incumbent role, while Vanguard tends to enter later as a cheaper competitor. In our sample, Vanguard funds (30 ETFs) are typically cheaper, with only 9.6% being labeled as high-fee category, whereas State Street (22 ETFs) and BlackRock (35 ETFs) have 54.5% and 80.55% high-fee funds, respectively. At the same time, just 12.9% of Vanguard ETFs are first movers for a given benchmark index, compared to 59.09% for State Street and 71% for BlackRock.

## 2.5 When does ETF competition emerge?

Why do some indices attract competing ETFs, while others do not? If we extend our model to account for the costs of launching and managing an ETF, Prediction 6 conjectures that a two-ETF economy arises when the share of high-turnover investors is large enough.

We extend our sample to include ETFs that do not face competition within an index so that we can test the factors that lead to the emergence of competition. We obtain a total sample of 1,071 indices tracked by ETFs, out of which only 65 indices are tracked by multiple funds. Although there are fewer competing ETFs than monopolist funds, competing ETFs manage 61% of total AUM as they tend to be much larger than monopolists. That is, indices tracked by multiple ETFs amass US$bn 1509.5 in AUM, compared to US$bn 959.27 for indices tracked by a single fund.

To formally test Prediction 6 and identify factors that lead to competition in a given ETF, we estimate Probit models using the cross-section of indices. The dependent variable is a dummy variable indicating whether the index is tracked by multiple ETFs. As in the previous sections, we measure investor urgency as the share of AUM held by transient investors. Further, we measure investor demand for a given index by summing the AUM of the ETFs tracking the index.

Table 8 reports the results of the Probit models. Consistent with Prediction 6, greater investor urgency, as proxied by the share of transient investors, increases the likelihood of a multiple-fund equilibrium.

Further, a high investor demand (AUM of ETFs for a given index) is a strong predictor of the emergence of same-index competition. Large demand from investors increases the likelihood that the low-fee ETF achieves sufficient economies of scale to break even and survive in equilibrium. The higher the total AUM tracking an index, the easier it is for an ETF to cover its fixed costs, even if the ETF has a lower market share than their competitor. A larger share of tax insensitive investors is also positively correlated with ETF competition: this is a natural result, since tax sensitive investors might be "locked in" the incumbent fund and unlikely to switch to a competitor even it charges a lower fee. Finally, it appears that creating a proprietary index offers some degree of insulation against competition, since indices licensed by major brands, such as FTSE, Russell, S&P, NASDAQ, MSCI, or Dow Jones, are more likely to be tracked by multiple ETFs.

### 2.6 The dynamics of fees around follower entry

In this section, we analyze the behavior of monopolistic leader ETFs in response to entry or potential entry of follower ETFs. Our focus is on events within our 2016 to 2020 sample, where a new ETF enters the market and competes with an incumbent monopolist ETF tracking the same index. To pinpoint fee changes, we collect daily data on management fees and assets under management (AUM) in a 2-year window around the entry data. Our data set includes 17 competitor entry events.

Figure 7 shows the results. On average, fees decrease by a small amount from 48.88 basis points before entry to 48.00 basis points post-entry. The result

**Table 8**
**Determinants of ETF competition**

| | Multiple ETFs tracking the index | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| log AUM index | 0.48*** | 0.43*** | 0.43*** | 0.50*** | 0.56*** | 0.58*** |
| | (7.96) | (7.13) | (6.87) | (7.23) | (6.30) | (7.16) |
| Transient investors (% AUM) | | | | | 0.43*** | 0.39*** |
| | | | | | (2.81) | (2.65) |
| Tax-insensitive investors (% AUM) | | | | | | 0.48*** |
| | | | | | | (3.33) |
| Top-3 ETF issuer | | 0.37** | 0.37** | 0.36* | 0.34* | 0.36* |
| | | (2.08) | (2.00) | (1.93) | (1.82) | (1.83) |
| Major brand index | | 0.32* | 0.31* | 0.33* | 0.37** | 0.37** |
| | | (1.86) | (1.81) | (1.89) | (2.22) | (2.18) |
| Number constituents (000s) | | | 0.03 | 0.01 | 0.07 | 0.16 |
| | | | (0.28) | (0.09) | (0.65) | (1.46) |
| Quoted spread (bps) | | | | 0.34*** | 0.28*** | 0.26*** |
| | | | | (5.33) | (4.47) | (4.15) |
| Constant | −11.31*** | −10.72*** | −10.68*** | −11.90*** | −13.30*** | −13.85*** |
| | (−8.72) | (−8.42) | (−8.17) | (−8.45) | (−7.19) | (−8.26) |
| Observations | 1,068 | 1,068 | 1,068 | 1,068 | 1,019 | 1,019 |
| Pseudo-$R^2$ | .384 | .397 | .397 | .409 | .426 | .448 |

This table reports results of Probit regressions in which the units of observation are indices. The dependent variable is the probability of observing multiple ETFs competing in tracking the given index. The independent variables are listed in the first column. All variables are computed as the average per index for 2016-2020. *Transient investors* is the share of AUM held by transient investors across all ETFs tracking a particular index. *Tax-insensitive investors* is the share of AUM held by tax-insensitive investors across all ETFs tracking a particular index. The *Top-3 issuer* dummy variable is one if a given index has ETFs issued by Vanguard, BlackRock, or State Street. *Major brand index* is a dummy variable that is one if the tracked index is licensed by FTSE, Russell, S&P, NASDAQ, MSCI, or Dow Jones. *Number of constituents* (in thousands) is the average number of constituents in the index. *Quoted spread* (in bps) is the bid-ask spread in the ETF secondary market divided by the ETF's midpoint price. AUM (in USD billion) is assets under management in the ETF(s). Chi-squared statistics are reported in parentheses. *$p<.1$; **$p<.05$; ***$p<.01$.

mirrors that of Box, Davis, and Fuller (2020), who find that incumbent ETF fees remain relatively flat following the launch of a new fund tracking the same index. At the same time, the average entrant fee is only 20 basis points, that is less than half of the average fee for the incumbent.

A key observation in Figure 7 is that monopolistic ETFs adjust their fees in the year prior to entry: eight funds adjusted their fees once, and three funds do so twice in the year leading up to new entrants' arrival. This observation suggests a proactive response to competitive threats.

In contrast, none of the ETFs in our sample alters their fee in the year following a competitor's entry. This lack of post-entry fee changes is consistent with the funds having reached a "steady-state" equilibrium prior to the actual launch date.

Further, in line with Box, Davis, and Fuller (2020), we document a diverse range of responses across incumbent ETFs. In our sample, leader funds reduce fees before entry in six cases, with an average decrease of 5.42%, while in five other instances, fees increase before entry, with an average rise of 0.70%. In the remaining six cases, fees remained unchanged throughout the 2-year period.

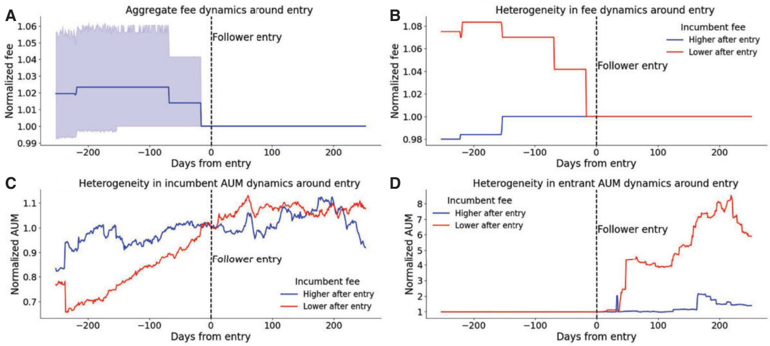


**Figure 7**
**Leader ETF fee dynamics around follower entry**
This figure plots the management fee and assets under management (AUM) for first-mover ETFs 1 year before and after the entry of their initial competitors. The daily data, sourced from ETF Global, encompasses 17 entry events between 2016 and 2020. All variables are normalized to a value of one on the entry day of the competitor. Panels B and C plot the average fee and AUM across two distinct groups of ETFs: those that increased their fees and those that decreased them prior to the entry of competitors. Panel D plots the AUM growth for entrant funds, separately for indices where the incumbent raised or decreases their fees upon entry.

The bottom-right panel in Figure 7 shows that ETFs that reduce their fees prior to new entrants' arrival experience rapid AUM growth in the year before entry. In contrast, the funds that raised their fees prior to entry display a more stable AUM trajectory.

Our model in Section 1 provides insights to explain why leader funds may increase fees prior to entry. The rationale is that, in the context of a stable AUM (i.e., a fixed mass of investors), the threat of entry incentivizes the leader ETF to maintain lower fees and preserve their monopoly position. A shift in investor preferences—in our setting, a change in the distribution of investor horizons—may allow the leader fund to accommodate competitor entry and raise their fee as illustrated in Figure 2.

A richer model that accounts for fixed costs could provide a rationale for leader funds reducing fees prior to entry. Consider a scenario in which potential entrants are deterred by significant fixed costs, such that entry becomes optimal only if demand for the index is sufficient; that is, the index's AUM exceeds a specific threshold. Under this scenario, when the AUM is relatively small, leader funds can charge monopoly-level fees without the threat of entry and regardless of the investor horizon distribution. However, as AUM increases, leader funds anticipate that entry has become viable and consequently reduce fees to competitive levels. Panel D in Figure 7 documents that the assets under management (AUM) for new entrants grow much faster for those indices for which incumbents reduced fees upon the entry of competitor. This is consistent with entry being driven by a growing interest in the benchmark index. It is a "rising tide lifts all boats" scenario.

Another potential reason ETFs may hesitate to raise fees after a new entrant, is the risk of legal actions, including antitrust or consumer protection claims, that may restrict their equilibrium strategies. This concern is heightened if regulators view new competitors as benchmarks for assessing fair fee levels.

## 3. Conclusion

This paper sheds light on how ETFs compete with one another, highlighting the important role of liquidity in this process. While some ETFs compete in product differentiation (selecting unique indices) like in product markets and active mutual funds, we show that many of the largest and most actively traded ETFs engage in a second form of competition involving a trade-off between liquidity and fees. High-turnover investors value liquidity and are willing to pay higher fees to access it. ETF issuers are aware of this value of liquidity and therefore more liquid ETFs charge higher fees and extract a liquidity rent.

If the demand to trade a particular index is large enough, liquidity clienteles emerge in equilibrium: highly liquid high-fee ETFs serve high-turnover investors, while relatively less liquid low-fee ETFs serve low-turnover investors. The high liquidity of high-fee ETFs is self-perpetuating due to a "liquidity begets liquidity" effect and because of the high-turnover clientele. Because of these effects, ETFs with higher fees than their competitors not only survive but also flourish in equilibrium due to the value of liquidity.

The empirical analysis of U.S. equity ETFs that face competition from other ETFs following the same index confirms the key model predictions. More liquid ETFs have higher market shares, tend to charge higher fees, and attract shorter horizon investors compared to their same-index competitors. Competition among same-index ETFs tends to emerge when investor demand (combined assets under management) to trade the particular index is high and investors have highly heterogeneous holding horizons.

We quantify the value of ETF liquidity. More liquid ETFs charge a 10.95-bps management fee premium and offer 3.02-bps lower bid-ask spreads than their direct competitors, holding fixed the index and other ETF characteristics. They also maintain a 38% market share advantage and have 7.31 times higher trading volumes. The liquidity channel is economically important, as it accounts for two-thirds of the explained variation in market shares, and for 25% of the explained variation in fees.

The more liquid ETF's liquidity advantage increases in the level of investor trading urgency. Our model suggests that if investor average holding horizons decrease, they are more likely to choose a more liquid ETF, which leads to larger differences in liquidity and fees for competing same-index ETFs. We also find that an ETF's launch date is the most important predictor of future market dominance; that is, 80% of the large and liquid ETFs in our sample are the first movers in tracking a particular index.

The results imply that liquidity can be a double-edged sword for investors. On one hand, highly liquid markets are beneficial to allocating resources efficiently and allowing investors to trade at low cost. On the other hand, liquidity leads to rent extraction from liquidity-sensitive investors. Competition between ETFs tracking the same index gives rise to welfare losses due to negative liquidity externalities by splitting the investor pool into liquidity clienteles.

Our paper has implications for regulation of the ETF industry. The liquidity effects in ETF product markets introduce a network effect that regulators should consider when addressing antitrust enforcement and shaping the industry's structure. A key consideration is that network effects can create barriers to entry, which directly affect antitrust considerations and enforcement strategies.

**Code Availability:** The replication code and data are available in the Harvard Dataverse at https://doi.org/10.7910/DVN/FVTGIU.

## Appendix A. Notation Summary

**Table A1.**
**Model notation**

| **Exogenous parameters and their interpretation** | |
|---|---|
| Parameter | Definition |
| $v_t$ | Common value of the ETF index at time $t$ |
| $\eta$ | Arrival rate of news |
| $\sigma$ | Size of common value innovation upon news arrival |
| $2Q\mathrm{d}i$ | Trade quantity for an individual investor |
| $\lambda_\ell, \lambda_h$ | Arrival rate for low- and high-turnover investors |
| $\alpha$ | Share of high-turnover investors |
| **Endogenous quantities and their interpretation** | |
| Variable | Definition |
| $f_k, k \in \{L, F\}$ | Management fee for ETF $k$ |
| $s_k$ | Half-spread posted by a market maker for ETF $k$ |
| $\Lambda_k, k \in \{L, F\}$ | Aggregate investor trading rate in ETF $k$ |
| **Agent labels** | |
| Label | Agent |
| **I** | Risk-neutral investors who trade for liquidity reasons |
| **M** | Risk-neutral market-makers |
| **A** | Arbitrageurs who trade on news arrival events |

## Appendix B. Proofs

**Lemma 1**

**Proof of Lemma** We show that the following strategies form an equilibrium in the trading subgame starting at $t = 0$ with an empty order book:

1. If the order book is empty at time $t > 0$, for each ETF $k$, all **M**s submit a buy limit order for $2Q\mathrm{d}i$ units at $v_t - s_k^\star$ and a sell limit order for $2Q\mathrm{d}i$ units at a price $v_t - s_k^\star$, where $s_k^\star = \frac{\eta}{\eta + \lambda_k}\sigma$. The first arriving market-maker fills the order book for both ETFs; the rest immediately cancel their quotes.

2. An trigger event happens at some random time $\tau$. If the event is a news arrival, the arbitrageur submits two market orders aimed at the stale quotes in each ETF $k$ – the ask (bid) side upon good (bad) news.
3. Following the trigger event, all **M**s cancel their outstanding quotes and the order book returns to an empty state.

From Equation (2), it is never optimal for **M**s to submits quotes larger than the $2Q\text{d}i$, that is the order size of an uninformed investor. The rationale is that the losses upon trading with an arbitrageurs scale linearly with the quote size, whereas the potential profits from trading with **I** are capped by their fixed order size. For any non-negative expected profit, a market-maker is at least weakly better off by submitting a limit order for $2Q\text{d}i$ units of the ETF.

Further, as in Menkveld and Zoican (2017), only the first market-maker to arrive at the exchange lets her quotes rest in the order book. Subsequent market-makers immediately cancel their orders as they have negative expected profit. The rationale is that uninformed investors trade only against the quote of the first market-maker; however, all quotes in order book are exposed to latency arbitrage losses upon the realization of news.

Next, we show that the equilibrium spread is indeed $s_k^\star$. Market makers are worse off by deviating and submitting quotes at a lower half-spread, that is, $s_k^\star - \epsilon$ for $\epsilon > 0$. If one **M** deviates and posts quotes at $s_k^\star - \epsilon$, then she undercuts her competitors and earns negative expected profit:

$$\mathbb{E}\pi_{\mathbf{M}}\left(s_k^\star - \epsilon\right) = \left[\frac{\lambda_k}{\lambda_k + \eta}\left(s_k^\star - \epsilon\right) + \frac{\eta}{\eta + \lambda_k}\left(s_k^\star - \epsilon - \sigma\right)\right] 2Q\,\text{d}i < \mathbb{E}\pi_{\mathbf{M}}\left(s_k^\star\right) \equiv 0 \tag{A1}$$

Second, a market maker is indifferent between submitting a quote with half spread $s_k^\star + \epsilon$ and a quote at half-spread $s_k^\star$. The reason is that such a quote is immediately undercut by competitor market makers and therefore never rests in the order book.

We note that $s_k^\star < \sigma$ for any $k$, and therefore the arbitrageur earns positive expected profit from "sniping" stale quotes upon the realization of news in both ETFs. Finally, upon the realization of a trade, market makers refill the order book around the current asset value. ■

## Appendix C. Data-Cleaning Procedures

We apply the following minor filters in the sample selection (in addition to those described in the data section of the main paper) :

- Two ETFs (SPLG, SPSM) are removed as they changed the index they track. Prior to November 16, 2017, SPLG tracked the Russell 1000 Index and traded under the ticker ONEK. Between November 16, 2017, and January 24, 2020, the SPLG fund tracked SSgA Large Cap Index. Prior to November 16, 2017, the SPSM tracked the Russell 2000 Index and traded under the ticker TWOK. Between November 16, 2017, and January 24, 2020, the SPSM fund tracked the SSGA Small Cap Index.
- We manually check that several ETFs are recorded twice in ETF Global database. We remove the instances of those ETFs that have incomplete or inaccurate data, and leave the ones with data entries (e.g., fee) that align with alternative data sources such as ETF.com. The ETFs for which duplicates are removed are IVV, IJH, and IJR.
- We remove several erroneous ETF comps matches that refer to ETFs that track fundamentally opposite versions (e.g., long vs. short) of the same index. We do so by searching the ETF description field for words such as "short" or "bear" and removing the resultant tickers.
- We remove two ETFs that are erroneously matched with their group although they track a different index than the rest of ETFs in their group (S&P Mid Cap 400 Pure Growth vs. S&P Mid Cap 400 Growth). These ETFs are RFG and RFV.

- We remove two ETF pairs that are erroneously matched with the same benchmark while they have rather different holdings. ASHR, targeting the CSI 300 Index, emphasizes sectors that reflect the broader Chinese economy. However, PEK provides a more balanced exposure to China A-shares.[18] Further, Global X Scientific Beta Asia ex-Japan ETF (SCIX) and Global X Scientific Beta Japan ETF (SCIJ) have nonoverlapping geographical areas, while they have the same primary benchmark on ETF global.
- We remove ETFs by the same issuer that have nonoverlapping portfolios (e.g., Gold Miners large cap vs. Junior Gold Miners small cap portfolios). These ETFs are GDX and GDXJ.

We apply the following data cleaning procedures in computing the variables used in the main regression analysis:

- For each ETF, the turnover and tracking difference variables are winsorized at the 99% level;
- For 37 ETFs for which FactSet does not report underlying index values, the tracking error and performance drag in the 2020 cross-section are replaced with the average value of tracking error and performance drag in the 100 remaining ETFs. In panel data, the same procedure is performed in each year-month.

## Appendix D. Matching ETF Characteristics in a Given Index

In our analysis, competing ETFs are assumed to be the same or similar with respect to structure. In Table D2, we outline the strategies to ensure ETFs in a given index are effectively identical: (a) controlling for the observed ETF characteristics in regressions, and (b) matching the same-index ETFs by observed characteristics.

**Table D2.**
**Strategies to ensure ETFs within a given index are very similar**

| ETF characteristic | Control variables | Set to identical in sample selection |
|---|---|---|
| Currency of returns | No | Yes |
| ETN structure | No | Yes |
| Index exposure | Yes | Yes |
| Investment style | No | Yes |
| Legal structure | Yes | No |
| Performance drag | Yes | No |
| Tax rate on capital gains | No | Yes |
| Tax on distributions | No | Yes |
| Tracking error | Yes | No |
| Marketing expenses | Yes | No |
| Security lending revenue | Yes | No |

This table lists observed ETF characteristics and how we ensure those characteristics are either controlled for in regressions or matched to be the same for same-index ETFs through the sample selection. There are two types of ETFs in our sample: open-end investment funds and unit investment trusts (UITs). We introduce the UIT dummy variable to control for the UIT structure. There are only two ETFs in the sample (MDY and SPY) that are UITs. Being a UIT, these ETFs reinvest their dividends not daily, but quarterly. They also cannot lend out securities.

## Appendix E. Additional Robustness Checks

In this section we present several robustness checks of our empirical analysis, as discussed in the main text. Table E7 lists all the same-index ETFs in our sample.

[18] See, for example, the ETF Insider article PEK versus ASHR: A Comprehensive Comparison of ETFs.

**Table E3.**
**Management fee and liquidity for competing ETFs: Identical benchmarks**

| | Fee (1) | Quoted spread (2) | log dollar volume (3) | Turnover (4) | Market share (5) | log profit (6) |
|---|---|---|---|---|---|---|
| High fee | 8.40*** | −4.06*** | 2.03*** | 1.61* | 0.57*** | 2.20*** |
| | (7.26) | (−3.02) | (4.60) | (1.79) | (5.85) | (6.89) |
| Name recognition | −0.14 | −0.26** | 0.01 | 1.47*** | 0.01 | −0.03 |
| | (−0.95) | (−2.62) | (0.16) | (3.74) | (0.50) | (−1.43) |
| log index AUM | 1.14*** | −18.71** | 1.57*** | −0.26 | 0.03 | 1.62*** |
| | (10.74) | (−2.18) | (3.65) | (−0.33) | (0.44) | (6.82) |
| Lending income (bps of AUM) | 0.80 | −2.49** | 0.44 | 0.57 | 0.23* | 0.51 |
| | (1.13) | (−2.35) | (1.03) | (0.99) | (1.74) | (1.10) |
| Marketing expense (bps) | 4.93** | −2.94** | −1.25** | −1.06 | −0.23* | −1.06** |
| | (2.19) | (−2.23) | (−2.18) | (−1.12) | (−2.08) | (−2.71) |
| Other net expenses | 6.14** | −3.47*** | −1.44*** | −1.94* | −0.21** | −0.93*** |
| | (2.77) | (−3.60) | (−3.07) | (−1.75) | (−2.47) | (−3.09) |
| Tracking error (bps) | −0.74*** | 2.96* | −0.22 | −0.07 | −0.07 | −0.24 |
| | (−6.70) | (1.73) | (−1.19) | (−0.15) | (−1.58) | (−1.53) |
| Performance drag (bps) | −0.12 | 0.99 | −0.10 | −0.45* | −0.02 | −0.07 |
| | (−1.28) | (1.10) | (−1.22) | (−1.86) | (−1.04) | (−1.10) |
| Unit investment trust | −3.16 | 7.45*** | 1.33 | 9.28*** | −0.39* | −0.13 |
| | (−0.81) | (3.59) | (1.64) | (3.86) | (−2.04) | (−0.23) |
| Tax-insensitive investors (TII) | 0.00 | 0.01 | −0.01 | 0.01 | −0.00 | −0.01 |
| | (0.17) | (0.57) | (−0.71) | (1.13) | (−0.64) | (−1.02) |
| Lagged ETF return | −3.09** | 16.47*** | −3.74* | −10.55 | −0.34** | −1.32** |
| | (−2.61) | (2.90) | (−2.04) | (−1.57) | (−2.33) | (−2.18) |
| TII $\times$ Lagged ETF return | −0.00 | −0.08* | 0.05** | 0.13* | 0.00 | 0.01 |
| | (−0.25) | (−2.04) | (2.29) | (2.05) | (1.51) | (1.10) |
| Constant | 17.78*** | 9.40*** | 16.09*** | 1.05 | 0.28 | 23.97*** |
| | (18.41) | (4.62) | (21.70) | (1.10) | (1.67) | (40.89) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 790 | 790 | 790 | 789 | 790 | 789 |
| Adjusted $R$-squared | .93 | .68 | .83 | .69 | .69 | .85 |

We replicate the analysis in Table 4 using only ETFs that track identical benchmarks. The unit of observation is an ETF-quarter. The dependent variables in the regressions are (1) ETF fee (management expense ratio) in bps, (2) Quoted bid-ask spread (in bps), (3) log dollar volume, (4) Turnover as a fraction (annualized secondary market traded dollar volume divided by the ETF's AUM), (5) Market share as a fraction (AUM of the ETF divided by the total AUM of all ETFs tracking the given index), (6) log profitability (AUM times fee). *Quoted spread* (in bps) is computed as the time-weighted bid-ask spread in the ETF secondary market divided by the ETF's midpoint price. *High fee* is a dummy variable that is one if the ETF has the highest fee among competing same-index ETFs. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *log index AUM* is the logarithm of the aggregate AUM for all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average difference in daily returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts.*Tax-insensitive investors* is the share of AUM held by tax-insensitive investors for a given index and quarter. *Lagged ETF return* is the percentage change in ETF price from the previous quarter. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

**Table E4.**
Management fee and liquidity for competing ETFs: Cross-sectional regressions

| | Fee (1) | Quoted spread (2) | log dollar volume (3) | Turnover (4) | Market share (5) | log profit (6) | Transient investor share (7) |
|---|---|---|---|---|---|---|---|
| High fee | 9.49*** | −2.21* | 1.52*** | 4.19*** | 0.34*** | 1.44*** | 0.05** |
| | (7.04) | (−1.78) | (3.18) | (3.25) | (3.12) | (3.87) | (2.29) |
| Name recognition | −1.05*** | −0.40 | 0.34*** | 2.32*** | 0.08*** | 0.14 | 0.00 |
| | (−2.83) | (−1.59) | (3.20) | (8.41) | (3.10) | (1.64) | (0.85) |
| log index AUM | 14.06 | −25.59* | 13.77*** | 4.12 | 2.26** | 12.08*** | −0.74** |
| | (0.82) | (−1.96) | (3.44) | (0.37) | (2.58) | (3.64) | (−2.14) |
| Lending income (bps of AUM) | 2.36** | −2.36*** | 0.61** | 0.54 | 0.12* | 0.63*** | −0.02 |
| | (2.09) | (−2.88) | (2.55) | (1.39) | (1.85) | (3.28) | (−1.06) |
| Marketing expense (bps) | −2.27 | −0.50 | 0.07 | 4.64** | −0.08 | −0.76 | 0.08*** |
| | (−0.96) | (−0.33) | (0.10) | (2.35) | (−0.59) | (−1.53) | (3.09) |
| Other net expenses | 0.65 | −3.95** | 0.90 | 4.24* | 0.14 | 0.29 | 0.03 |
| | (0.23) | (−2.28) | (1.07) | (1.82) | (0.75) | (0.45) | (0.96) |
| Tracking error (bps) | 10.63*** | −3.06* | 0.70* | 4.37*** | −0.05 | 0.51 | −0.04 |
| | (5.07) | (−1.81) | (1.80) | (2.69) | (−0.61) | (1.65) | (−1.25) |
| Performance drag (bps) | 3.91** | 0.71 | −0.22 | −0.18 | −0.09 | −0.12 | 0.01 |
| | (2.17) | (0.45) | (−0.82) | (−0.17) | (−1.27) | (−0.48) | (0.34) |
| Unit investment trust | 6.27 | 7.16** | −2.28* | −5.67 | −0.93*** | −1.53* | −0.03 |
| | (1.31) | (2.65) | (−1.91) | (−1.60) | (−3.58) | (−1.80) | (−0.43) |
| Tax-insensitive investors (TII) | 0.16 | 0.73 | −0.24 | 1.05 | −0.07* | −0.33** | 0.07*** |
| | (0.17) | (1.33) | (−1.22) | (1.50) | (−1.82) | (−2.29) | (2.83) |
| Constant | 21.30*** | 8.78*** | 16.98*** | 4.07*** | 0.40*** | 24.32*** | 0.20*** |
| | (25.26) | (9.96) | (53.59) | (5.12) | (5.18) | (90.97) | (10.65) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 103 | 103 | 103 | 103 | 103 | 103 | 103 |
| Adjusted $R$-squared | .85 | .69 | .55 | .44 | .05 | .67 | .59 |

We replicate the analysis in Table 4 using cross-sectional averages for each ETF, across all measures within our 2016-2020 sample. The unit of observation is the ETF. The dependent variables in the regressions are (1) ETF fee (management expense ratio) in basis points, (2) Quoted spread (in bps), computed as the time-weighted bid-ask spread in the ETF secondary market divided by the ETF's midpoint price, (3) log dollar volume, (4) Turnover as a fraction (the annualized ratio of daily dollar volume in the ETF's secondary market and assets under management, AUM), (5) market share as a fraction (AUM of the ETF divided by the total AUM of all ETFs tracking the given index), (6) log profitability (AUM times fee), (7) The AUM share of transient investors in a given ETF. *High fee* is a dummy taking value one if the ETF has the highest fee among competitors tracking similar indices. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *Log index AUM* is the logarithm of the aggregate AUM across all ETFs tracking the same index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in basis points) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in basis points) is the average daily difference in returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. *Tax-insensitive investors* is the AUM share held by tax-insensitive investors for a given index and quarter. *Lagged ETF return* stands for the percentage change in ETF price from the previous quarter. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are clustered by index. $t$-statistics are reported in parentheses.
$^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

**Table E5.**
**Management fee and liquidity for competing ETFs: Alternative liquidity measures**

| | Quoted spread | | Effective spread | | Realized spread | |
|---|---|---|---|---|---|---|
| | (bps) (1) | (US¢) (2) | (bps) (3) | (US¢) (4) | (bps) (5) | (US¢) (6) |
| High fee | −3.02*** | −1.82*** | −1.67*** | −1.06** | −1.75*** | −1.06*** |
| | (−4.15) | (−3.39) | (−3.32) | (−2.77) | (−3.80) | (−3.01) |
| Name recognition | −0.17** | −0.37** | −0.03 | −0.16** | 0.02 | −0.02 |
| | (−2.22) | (−2.57) | (−0.53) | (−2.82) | (0.35) | (−0.61) |
| log index AUM | −6.09*** | −1.29 | −4.36*** | −1.14 | −3.34** | −1.04 |
| | (−3.40) | (−1.13) | (−3.19) | (−1.51) | (−2.53) | (−1.64) |
| Lending income | −2.23*** | −0.47*** | −1.89*** | −0.50*** | −1.73*** | −0.51*** |
| (bps of AUM) | (−6.21) | (−3.50) | (−5.61) | (−3.12) | (−5.44) | (−3.36) |
| Marketing expense (bps) | 0.20 | −2.52*** | 0.89 | −1.21** | 0.80 | −0.78* |
| | (0.25) | (−4.14) | (1.41) | (−2.72) | (1.33) | (−1.96) |
| Other net expenses | −1.88** | −1.90*** | −0.64 | −0.80* | −0.66 | −0.55 |
| | (−2.38) | (−3.26) | (−1.22) | (−1.92) | (−1.42) | (−1.44) |
| Tracking error (bps) | −1.57** | −0.64* | −1.12** | −0.52* | −0.86* | −0.33 |
| | (−2.42) | (−1.84) | (−2.16) | (−1.98) | (−2.00) | (−1.46) |
| Performance drag (bps) | −0.21 | −0.01 | −0.07 | −0.03 | −0.08 | −0.02 |
| | (−0.66) | (−0.07) | (−0.27) | (−0.24) | (−0.41) | (−0.24) |
| Unit investment trust | 4.63*** | 5.51*** | 1.71** | 2.53*** | 1.53*** | 0.85** |
| | (4.11) | (3.69) | (2.69) | (4.15) | (2.94) | (2.52) |
| Tax-insensitive investors | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | (1.32) | (0.52) | (0.93) | (0.93) | (1.25) | (1.45) |
| Lagged return | 2.50** | 2.43 | −5.54 | −1.57 | −2.64 | −0.83 |
| | (2.59) | (1.03) | (−0.75) | (−0.65) | (−0.80) | (−0.69) |
| TII × Lagged return | −0.03 | 0.00 | 0.05 | 0.02 | 0.04 | 0.02* |
| | (−1.16) | (0.21) | (0.50) | (1.11) | (0.80) | (1.91) |
| Constant | 7.08*** | 5.66*** | 4.43*** | 3.14*** | 2.90*** | 1.96*** |
| | (5.47) | (6.50) | (4.31) | (5.23) | (3.48) | (3.87) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,703 | 1,703 | 1,703 | 1,703 | 1,703 | 1,703 |
| Adjusted $R$-squared | .65 | .57 | .61 | .50 | .54 | .47 |

We replicate the analysis in Table 4 using alternative liquidity measures. The unit of observation is ETF-quarter. The dependent variables in the regressions are (1) *Quoted spread* (in bps and US cents), computed as the time-weighted bid-ask spread in the ETF secondary market divided by the ETF's midpoint price; (2) *Effective spread* (in bps and US cents), computed as the time-weighted average of the signed difference between trade price and midpoint, and (3) *Realized spread* (in bps and US cents), computed as the time-weighted average of the signed difference between trade price and the midpoint five minutes after the trade. *High fee* is a dummy variable that is one if the ETF has the highest fee among competing same-index ETFs. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *log index AUM* is the logarithm of the aggregate AUM for all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other net expenses* (in bps) is taken from ETF Global and capture ETF costs such as licensing fees and distribution costs, net of fee waivers and marketing expenses. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average difference in daily returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. *Tax-insensitive investors* is the share of AUM held by tax-insensitive investors for a given index and quarter. *Lagged ETF return* is the percentage change in ETF price from the previous quarter. All dependent variables that are not dummies are standardized to have zero mean and unit variance. Standard errors are double clustered by ETF and quarter. $t$-statistics are reported in parentheses. *$p$ <.1; **$p$ <.05; ***$p$ <.01.

**Table E6.**
**Liquidity channel versus alternatives (dollar spreads)**

| | Market share | | | Fee | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Liquidity $R^2$ contribution** | **48.74%** | **50.33%** | **55.83%** | **5.52%** | **5.51%** | **7.54%** |
| Quoted spread (cents) | −0.04*** | | | −0.32* | | |
| | (−5.30) | | | (−1.93) | | |
| Effective spread (cents) | | −0.06*** | | | −0.45* | |
| | | (−4.82) | | | (−1.74) | |
| Realized spread (cents) | | | −0.07*** | | | −0.59** |
| | | | (−5.55) | | | (−2.04) |
| **Reputation $R^2$ contribution** | **34.77%** | **34.32%** | **30.18%** | **40.71%** | **40.96%** | **39.90%** |
| Name recognition | 0.01 | 0.01 | 0.02 | −0.31 | −0.27 | −0.22 |
| | (0.70) | (0.94) | (1.28) | (−1.26) | (−1.18) | (−1.03) |
| Marketing expense (bps) | −0.02 | −0.01 | −0.01 | −0.29 | −0.26 | −0.25 |
| | (−1.52) | (−1.19) | (−1.08) | (−0.98) | (−0.88) | (−0.86) |
| Other net expenses | 0.42 | 0.77 | 0.80 | 25.44 | 28.78 | 28.56 |
| | (0.40) | (0.73) | (0.80) | (0.91) | (1.04) | (1.02) |
| **TII $R^2$ contribution** | **0.91%** | **1.00%** | **0.99%** | **2.25%** | **2.32%** | **2.33%** |
| Tax-insensitive investors (TII) | −0.00 | 0.00 | 0.00 | 0.02 | 0.02 | 0.02 |
| | (−0.00) | (0.20) | (0.37) | (0.63) | (0.70) | (0.78) |
| **Other $R^2$ contribution** | **15.56%** | **14.33%** | **12.99%** | **51.51%** | **51.19%** | **50.21%** |
| Lending income (bps of AUM) | 0.02*** | 0.02*** | 0.01** | 0.23 | 0.22 | 0.21 |
| | (3.35) | (2.75) | (2.50) | (0.79) | (0.73) | (0.68) |
| Tracking error (bps) | −0.00 | −0.00 | −0.00 | 0.16*** | 0.16*** | 0.16*** |
| | (−0.42) | (−0.56) | (−0.43) | (3.18) | (3.12) | (3.17) |
| Performance drag (bps) | −0.00 | −0.00 | −0.00 | 0.31** | 0.31** | 0.31** |
| | (−0.77) | (−0.82) | (−0.91) | (2.32) | (2.31) | (2.32) |
| Unit investment trust | −0.11 | −0.17 | −0.27 | 10.32** | 9.75** | 9.03** |
| | (−0.49) | (−0.68) | (−0.96) | (2.11) | (2.11) | (2.03) |
| Constant | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.03 |
| | (1.56) | (1.02) | (0.75) | (1.42) | (1.20) | (1.03) |
| Index FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,703 | 1,703 | 1,703 | 1,703 | 1,703 | 1,703 |
| Adjusted $R$-squared | .34 | .34 | .37 | .34 | .34 | .35 |

This table reports the Shapley-Owen (Owen 1977) relative contributions to ETF market shares and fees, computed as in Equation (23). The unit of observation is an ETF-quarter. The dependent variables in the regressions are (1) ETF fee (management expense ratio) in bps and (2) market share as a fraction. *Quoted spread* (in bps) is the bid-ask spread in the ETF secondary market divided by the ETF's midpoint price. *Turnover* is the annualized daily dollar volume in the ETF's secondary market divided by AUM. *Name recognition* is the daily average count of StockTwits messages mentioning the ETF. *log index AUM* is the logarithm of the aggregate AUM for all ETFs tracking the index. *Lending income* (in bps) is $ETF lending income from 2020, divided by AUM. *Marketing expenses* (in bps) are 12b-1 fees. *Other expenses* and *Fee waivers* are taken from ETF Global and capture ETF costs such as licensing fees and distribution costs. *Tracking error* (in bps) is the standard deviation of daily differences in returns of the ETF and its underlying index. *Performance drag* (in bps) is the average daily difference in returns of the ETF and its underlying index. *Unit investment trust* is a dummy variable for ETFs that are structured as unit investment trusts. *Tax-insensitive investors* is the AUM share held by tax-insensitive investors for a given index and quarter. Standard errors are clustered by quarter. $t$-statistics are reported in parentheses. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

**Table E7.**
List of same-index ETFs

| Index | Ticker | Issuer | Inception date | Primary benchmark | AUM (US$bn) |
|---|---|---|---|---|---|
| All World | ACWI | Blackrock | 3/26/2008 | MSCI ACWI NR USD | 8.87 |
| All World | VT | Vanguard | 6/24/2008 | FTSE Global All Cap TR USD | 10.6 |
| All World Ex US | VXUS | Vanguard | 1/26/2011 | FTSE Global All Cap ex US TR USD | 13 |
| All World Ex US | VEU | Vanguard | 3/2/2007 | FTSE AW Ex US TR USD | 20.83 |
| Asia ex Japan | AAXJ | Blackrock | 8/13/2008 | MSCI AC Asia Ex Japan NR USD | 3.8 |
| Asia ex Japan | FLAX | Franklin Templeton | 2/6/2018 | FTSE Asia ex-Japan Capped Index | 0.02 |
| Australia | EWA | Blackrock | 3/12/1996 | MSCI Australia NR USD | 1.5 |
| Australia | FLAU | Franklin Templeton | 11/2/2017 | FTSE Australia Capped Index | 0.01 |
| Brazil | EWZ | Blackrock | 7/10/2000 | MSCI Brazil 25/50 NR USD | 6.33 |
| Brazil | FLBR | Franklin Templeton | 11/3/2017 | FTSE Brazil Capped Index | 0.06 |
| Canada | BBCA | JPMorgan | 8/7/2018 | Morningstar Canada Target Market Exposure Index | 3.19 |
| Canada | EWC | Blackrock | 3/12/1996 | MSCI Canada NR USD | 2.74 |
| China Technology Sector | CQQQ | Guggenheim | 12/8/2009 | AlphaShares China Technology USD | 0.43 |
| China Technology Sector | KWEB | KraneShares | 7/31/2013 | CSI Overseas China Internet GR USD | 1.4 |
| Consumer Discretionary Sector | XLY | SSgA | 12/16/1998 | S&P Consumer Dis Select Sector TR USD | 13 |
| Consumer Discretionary Sector | VCR | Vanguard | 1/26/2004 | MSCI US IMI/Consumer Discr 25-50 GR USD | 2.67 |
| Consumer Staples Sector | XLP | SSgA | 12/16/1998 | S&P Cons Staples Select Sector TR USD | 10.72 |
| Consumer Staples Sector | VDC | Vanguard | 1/26/2004 | MSCI US IMI/Cons Staples 25-50 GR USD | 4.27 |
| EAFE | VEA | Vanguard | 7/20/2007 | FTSE Developed ex North America NR USD | 61.49 |
| EAFE | IEFA | Blackrock | 10/18/2012 | MSCI EAFE IMI NR USD | 46.84 |
| EAFE Small Cap | SCZ | Blackrock | 12/10/2007 | MSCI EAFE Small Cap NR USD | 8.72 |
| EAFE Small Cap | SCHC | Charles Schwab | 1/14/2010 | FTSE Dvlp Small Cap Ex US Liquid NR USD | 1.79 |
| Emerging Markets | EEM | Blackrock | 4/7/2003 | MSCI EM NR USD | 29.46 |
| Emerging Markets | VWO | Vanguard | 3/4/2005 | FTSE Emerging NR USD | 56.43 |
| Energy Sector | VDE | Vanguard | 9/23/2004 | MSCI US IMI/Energy 25-50 GR USD | 3.59 |
| Energy Sector | XLE | SSgA | 12/16/1998 | S&P Energy Select Sector TR USD | 14.01 |
| Europe | IEV | Blackrock | 7/25/2000 | S&P Europe 350 TR USD | 2.21 |
| Europe | FEZ | SSgA | 10/15/2002 | EURO STOXX 50 NR USD | 2.8 |
| Eurozone | VGK | Vanguard | 3/4/2005 | FTSE Developed Europe NR USD | 14.09 |
| Eurozone | EZU | Blackrock | 7/25/2000 | MSCI EMU NR USD | 8.01 |

(continued)

**Table E7.**
Continued

| Index | Ticker | Issuer | Inception date | Primary benchmark | AUM (US$bn) |
|---|---|---|---|---|---|
| France | EWQ | Blackrock | 3/12/1996 | MSCI France NR USD | 0.68 |
| France | FLFR | Franklin Templeton | 11/2/2017 | FTSE France RIC Capped Index | <0.01 |
| Germany | ZDEU | SSgA | 6/11/2014 | Solactive GBS Germany Large & Mid Cap Index | 0.01 |
| Germany | EWG | Blackrock | 3/12/1996 | MSCI Germany NR USD | 3.35 |
| Hong Kong | EWH | Blackrock | 3/12/1996 | MSCI Hong Kong NR USD | 1.92 |
| Hong Kong | FLHK | Franklin Templeton | 11/2/2017 | FTSE Hong Kong Capped Index | 0.02 |
| JPX-Nikkei 400 Net Total Return Index | JPN | Deutsche Bank | 6/24/2015 | JPX-Nikkei 400 Total Return Index | 0.02 |
| JPX-Nikkei 400 Net Total Return Index | JPXN | iShares | 10/23/2001 | JPX-Nikkei 400 Net Total Return Index | 0.1 |
| Japan | EWJ | Blackrock | 3/12/1996 | MSCI Japan NR USD | 15.35 |
| Japan | BBJP | JPMorgan | 6/15/2018 | Morningstar Japan Target Market Exposure Index | 3.87 |
| Korea | FLKR | Franklin Templeton | 11/2/2017 | FTSE South Korea Capped Index | 0.02 |
| Korea | EWY | Blackrock | 5/9/2000 | MSCI Korea 25/50 NR USD | 4.1 |
| MSCI ACWI Low Carbon Target Index | LOWC | SSgA | 11/25/2014 | MSCI ACWI Low Carbon Target Index | 0.11 |
| MSCI ACWI Low Carbon Target Index | CRBN | Blackrock | 12/8/2014 | MSCI ACWI Low Carbon Target Index | 0.43 |
| MSCI ACWI ex-US Index | CWI | SSgA | 1/10/2007 | MSCI ACWI Ex USA NR USD | 1.36 |
| MSCI ACWI ex-US Index | ACWX | Blackrock | 3/26/2008 | MSCI ACWI Ex USA NR USD | 2.98 |
| MSCI All Argentina 25/50 Index | ARGT | Global X | 3/2/2011 | FTSE Argentina 20 USD | 0.1 |
| MSCI All Argentina 25/50 Index | AGT | iShares | 4/25/2017 | MSCI All Argentina 25/50 Index | 0.02 |
| MSCI China All Shares Index | KALL | KraneShares | 2/12/2015 | MSCI China All Shares Index | 0.01 |
| MSCI China All Shares Index | CN | Deutsche Bank | 4/30/2014 | MSCI China All Shares Index | 0.04 |
| MSCI Emerging Markets ESG Leaders Index | EASG | DWS | 9/6/2018 | MSCI EAFE ESG Leaders Index | 0.01 |
| MSCI Emerging Markets ESG Leaders Index | EMSG | DWS | 12/6/2018 | MSCI Emerging Markets ESG Leaders Index | 0.01 |
| MSCI Norway IMI 25/50 Index | ENOR | Blackrock | 1/23/2012 | MSCI Norway IMI 20/50 NR USD | 0.03 |
| MSCI Norway IMI 25/50 Index | NORW | Global X | 11/9/2010 | FTSE Norway 30 USD | 0.13 |
| Mexico | EWW | Blackrock | 3/12/1996 | MSCI Mexico IMI 25/50 TR USD | 1.08 |
| Mexico | FLMX | Franklin Templeton | 11/3/2017 | FTSE Mexico RIC Capped Index | <0.01 |
| NASDAQ-100 Equal Weighted Index | QQEW | First Trust | 4/19/2006 | NASDAQ 100 Equal Weighted TR USD | 0.62 |
| NASDAQ-100 Equal Weighted Index | QQQE | Direxion | 3/21/2012 | NASDAQ 100 Equal Weighted TR USD | 0.18 |
| Preferred Shares | PGX | Invesco PowerShares | 1/31/2008 | BofAML Preferred Stock Fixed Rate TR USD | 5.11 |
| Preferred Shares | PFF | Blackrock | 3/26/2007 | S&P Preferred Stock TR USD | 16.39 |

(continued)

**Table E7.**
Continued

| Index | Ticker | Issuer | Inception date | Primary benchmark | AUM (US$bn) |
|---|---|---|---|---|---|
| Russell 1000 Growth Index | IWF | Blackrock | 5/22/2000 | Russell 1000 Growth TR USD | 41.66 |
| Russell 1000 Growth Index | VONG | Vanguard | 9/20/2010 | Russell 1000 Growth TR USD | 2.19 |
| Russell 1000 Index | IWB | Blackrock | 5/15/2000 | Russell 1000 TR USD | 19.37 |
| Russell 1000 Index | VONE | Vanguard | 9/20/2010 | Russell 1000 TR USD | 1 |
| Russell 1000 Value Index | IWD | Blackrock | 5/22/2000 | Russell 1000 Value TR USD | 36.08 |
| Russell 1000 Value Index | VONV | Vanguard | 9/20/2010 | Russell 1000 Value TR USD | 1.55 |
| Russell 2000 Growth Index | IWO | Blackrock | 7/24/2000 | Russell 2000 Growth TR USD | 8.58 |
| Russell 2000 Growth Index | VTWG | Vanguard | 9/20/2010 | Russell 2000 Growth TR USD | 0.26 |
| Russell 2000 Index | IWM | Blackrock | 5/22/2000 | Russell 2000 TR USD | 39.82 |
| Russell 2000 Index | VTWO | Vanguard | 9/20/2010 | Russell 2000 TR USD | 1.33 |
| Russell 2000 Value Index | IWN | Blackrock | 7/24/2000 | Russell 2000 Value TR USD | 8.58 |
| Russell 2000 Value Index | VTWV | Vanguard | 9/20/2010 | Russell 2000 Value TR USD | 0.22 |
| Russell 3000 Index | VTHR | Vanguard | 9/20/2010 | Russell 3000 TR USD | 0.41 |
| Russell 3000 Index | IWV | Blackrock | 5/22/2000 | Russell 3000 TR USD | 8.37 |
| Russia | ERUS | Blackrock | 11/9/2010 | MSCI Russia 25/50 NR USD | 0.5 |
| Russia | RSX | Van Eck | 4/24/2007 | MV Russia NR USD | 1.61 |
| S&P 500 Dividend Aristocrats Index | NOBL | ProShares | 10/9/2013 | S&P 500 Dividend Aristocrats USD | 4.03 |
| S&P 500 Dividend Aristocrats Index | KNG | Cboe Vest Financial | 3/26/2018 | S&P 500 Dividend Aristocrats USD | 0.04 |
| S&P 500 Growth Index | SPYG | SSgA | 9/25/2000 | S&P 500 Growth Index | 5.3 |
| S&P 500 Growth Index | IVW | Blackrock | 5/22/2000 | S&P 500 Growth TR | 20.96 |
| S&P 500 Growth Index | VOOG | Vanguard | 9/7/2010 | S&P 500 Growth TR | 1.29 |
| S&P 500 Index | SPY | SSgA | 1/22/1993 | S&P 500 TR USD | 252.25 |
| S&P 500 Index | IVV | Blackrock | 5/15/2000 | S&P 500 TR USD | 147.71 |
| S&P 500 Value Index | IVE | Blackrock | 5/22/2000 | S&P 500 Value TR USD | 14.32 |
| S&P 500 Value Index | VOOV | Vanguard | 9/7/2010 | S&P 500 Value TR USD | 0.6 |
| S&P 500 Value Index | SPYV | SSgA | 9/25/2000 | S&P 500 Value Index | 3.46 |
| S&P Global Infrastructure Index | GII | SSgA | 1/25/2007 | S&P Global Infrastructure TR USD | 0.24 |
| S&P Global Infrastructure Index | IGF | Blackrock | 12/10/2007 | S&P Global Infrastructure TR USD | 2.31 |
| S&P Midcap 400 Growth Index | MDYG | SSgA | 11/8/2005 | S&P Midcap 400 Pure Growth Index | 1.51 |

continued

**Table E7.**
Continued

| Index | Ticker | Issuer | Inception date | Primary benchmark | AUM (US$bn) |
|---|---|---|---|---|---|
| S&P Midcap 400 Growth Index | IJK | Blackrock | 7/24/2000 | S&P MidCap 400 Growth TR | 6.82 |
| S&P Midcap 400 Growth Index | IVOG | Vanguard | 9/7/2010 | S&P MidCap 400 Growth TR | 0.61 |
| S&P Midcap 400 Index | MDY | SSgA | 5/4/1995 | S&P MidCap 400 TR | 18.05 |
| S&P Midcap 400 Index | IJH | Blackrock | 5/22/2000 | S&P MidCap 400 TR | 42.12 |
| S&P Midcap 400 Value Index | IVOV | Vanguard | 9/7/2010 | S&P MidCap 400 Value TR USD | 0.52 |
| S&P Midcap 400 Value Index | IJJ | Blackrock | 7/24/2000 | S&P MidCap 400 Value TR USD | 5.51 |
| S&P Midcap 400 Value Index | MDYV | SSgA | 11/8/2005 | S&P Midcap 400 Pure Value Index | 1.17 |
| S&P SmallCap 600 Index | SLY | SSgA | 11/8/2005 | S&P SmallCap 600 Index | 1.13 |
| S&P SmallCap 600 Index | IJR | Blackrock | 5/22/2000 | S&P SmallCap 600 TR USD | 36.18 |
| S&P SmallCap 600 Index | VIOO | Vanguard | 9/7/2010 | S&P SmallCap 600 TR USD | 0.49 |
| S&P Smallcap 600 Growth Index | SLYG | SSgA | 9/25/2000 | S&P Smallcap 600 Growth Index | 1.81 |
| S&P Smallcap 600 Growth Index | IJT | Blackrock | 7/24/2000 | S&P SmallCap 600 Growth TR | 4.74 |
| S&P Smallcap 600 Growth Index | VIOG | Vanguard | 9/7/2010 | S&P SmallCap 600 Growth TR | 0.2 |
| S&P Smallcap 600 Value Index | IJS | Blackrock | 7/24/2000 | S&P SmallCap 600 Value TR | 5.21 |
| S&P Smallcap 600 Value Index | SLYV | SSgA | 9/25/2000 | S&P Smallcap 600 Value Index | 1.97 |
| S&P Smallcap 600 Value Index | VIOV | Vanguard | 9/7/2010 | S&P SmallCap 600 Value TR | 0.18 |
| South Africa | FLZA | Franklin Templeton | 10/10/2018 | FTSE/JSE South Africa RIC Capped Index | <0.01 |
| South Africa | EZA | Blackrock | 2/3/2003 | MSCI South Africa NR USD | 0.39 |
| Switzerland | FLSW | Franklin Templeton | 2/6/2018 | FTSE Switzerland RIC Capped Index | 0.01 |
| Switzerland | EWL | Blackrock | 3/12/1996 | MSCI Switzerland 25/50 TR USD | 1.17 |
| Taiwan | FLTW | Franklin Templeton | 11/2/2017 | FTSE Taiwan RIC Capped Index | 0.01 |
| Taiwan | EWT | Blackrock | 6/20/2000 | MSCI Taiwan NR USD | 3.45 |
| UK | EWU | Blackrock | 3/12/1996 | MSCI UK NR USD | 2.23 |
| UK | FLGB | Franklin Templeton | 11/2/2017 | FTSE UK Capped Index | 0.04 |
| US Communications Sector | XLC | SSgA | 6/18/2018 | MSCI US IMI/Telecom Svc 25-50 GR USD | 5.93 |
| US Communications Sector | VOX | Vanguard | 9/23/2004 | MSCI US IMI/Telecom Svc 25-50 GR USD | 1.66 |
| US Defense Sector | PPA | Invesco PowerShares | 10/26/2005 | AMEX SPADE Defense TR USD | 0.75 |
| US Defense Sector | ITA | Blackrock | 5/1/2006 | DJ US Select Aerospace&Defense TR USD | 3.7 |
| US Financials Sector | XLF | SSgA | 12/16/1998 | S&P Financial Select Sector TR USD | 23.27 |
| US Financials Sector | VFH | Vanguard | 1/26/2004 | MSCI US IMI/Financials 25-50 GR USD | 6.29 |

continued

**Table E7.**
Continued

| Index | Ticker | Issuer | Inception date | Primary benchmark | AUM (US$bn) |
|---|---|---|---|---|---|
| US Healthcare Sector | XLV | SSgA | 12/16/1998 | S&P Health Care Select Sector TR USD | 17.66 |
| US Healthcare Sector | VHT | Vanguard | 1/26/2004 | MSCI US IMI/Health Care 25-50 GR USD | 8.06 |
| US Industrials Sector | XLI | SSgA | 12/16/1998 | S&P Industrial Select Sector TR USD | 10.53 |
| US Industrials Sector | VIS | Vanguard | 9/23/2004 | MSCI US IMI/Industrials 25-50 GR USD | 3.13 |
| US Internet Sector | PNQI | Invesco PowerShares | 6/12/2008 | NASDAQ Internet TR USD | 0.52 |
| US Internet Sector | FDN | First Trust | 6/19/2006 | DJ Internet Composite TR USD | 6.73 |
| US Materials Sector | VAW | Vanguard | 1/26/2004 | MSCI US IMI/Materials 25-50 GR USD | 1.93 |
| US Materials Sector | XLB | SSgA | 12/16/1998 | S&P Materials Select Sector TR USD | 3.89 |
| US Small Caps | SMMD | iShares | 7/6/2017 | Russell 2500 Index | 0.04 |
| US Small Caps | VXF | Vanguard | 12/27/2001 | S&P 500 TR USD | 6.5 |
| US Technology Sector | XLK | SSgA | 12/16/1998 | S&P Technology Select Sector TR USD | 21.04 |
| US Technology Sector | VGT | Vanguard | 1/26/2004 | MSCI US IMI/Info Tech 25-50 GR USD | 19.47 |
| US Total Market | ITOT | iShares | 1/20/2004 | S&P Total Market Index | 20.88 |
| US Total Market | VTI | Vanguard | 5/24/2001 | MSCI US Broad Market NR USD | 103.85 |
| US Utilities Sector | XLU | SSgA | 12/16/1998 | S&P Utilities Select Sector TR USD | 8.94 |
| US Utilities Sector | VPU | Vanguard | 1/26/2004 | MSCI US IMI/Utilities 25-50 GR USD | 3.13 |

This table lists the ETFs competing across same or similar indices in our sample. For each ETF, we report the ticker, issuer, inception day, primary benchmark, and the average assets under management (AUM) between 2016 and 2020.

**References**


Agapova, A. 2011. Conventional mutual index funds versus exchange-traded funds. *Journal of Financial Markets* 14:323–43.

Agarwal, V., P. Hanouna, R. Moussawi, and C. W. Stahel. 2018. Do ETFs increase the commonality in liquidity of underlying stocks? *Working paper, Georgia State University*.

Amihud, Y., and H. Mendelson. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 17:223–49.

———. 1991. Liquidity, maturity, and the yields on us treasury securities. *Journal of Finance* 46:1411–25.

Balchunas, E. 2016. *The institutional etf toolbox: How institutions can understand and utilize the fast-growing world of etfs*. Hoboken, NJ: John Wiley & Sons.

Barber, B. M., T. Odean, and L. Zheng. 2005. Out of sight, out of mind: The effects of expenses on mutual fund flows. *Journal of Business* 78:2095–120.

Ben-David, I., F. Franzoni, and R. Moussawi. 2018. Do ETFs increase volatility? *Journal of Finance* 73:2471–535.

Bhattacharya, A., and M. O'Hara. 2018. Can ETFs increase market fragility? Effect of information linkages in ETF markets. Working paper, Baruch College.

Biglaiser, G., and J. Crémer. 2020. The value of incumbency when platforms face heterogeneous customers. *American Economic Journal: Microeconomics* 12:229–69.

Blouin, J. L., B. J. Bushee, and S. A. Sikes. 2017. Measuring tax-sensitive institutional investor ownership. *Accounting Review* 92:49–76.

Box, T., R. Davis, and K. Fuller. 2020. The dynamics of ETF fees. *Financial Analysts Journal* 76:11–8.

Brolley, M., and M. Zoican. 2022. Liquid speed: A micro-burst fee for low-latency exchanges. *Journal of Financial Markets* 64:100785–.

Broman, M. S., and P. Shum. 2018. Relative liquidity, fund flows and short-term demand: Evidence from exchange-traded funds. *Financial Review* 53:87–115.

Budish, E., P. Cramton, and J. Shim. 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics* 130:1547–621.

Bushee, B. J. 1998. The influence of institutional investors on myopic r&d investment behavior. *Accounting Review* 73:305–33.

Cabral, L. 2011. Dynamic price competition with network effects. *Review of Economic Studies* 78:83–111.

Chinco, A., and V. Fos. 2021. The sound of many funds rebalancing. *Review of Asset Pricing Studies* 11:502–51.

Chordia, T. 1996. The structure of mutual fund charges. *Journal of Financial Economics* 41:3–39.

Comerton-Forde, C., and T. Marta. 2021. Etf effects: the role of primary versus secondary market activities. *Working paper, University of Melbourne*.

Cookson, J. A., and M. Niessner. 2020. Why don't we agree? evidence from a social network of investors. *Journal of Finance* 75:173–228.

Cremers, M., and A. Pareek. 2016. Patient capital outperformance: The investment skill of high active share managers who trade infrequently. *Journal of Financial Economics* 122:288–306.

Da, Z., and S. Shive. 2018. Exchange traded funds and asset return correlations. *European Financial Management* 24:136–68.

Dannhauser, C. D. 2017. The impact of innovation: Evidence from corporate bond exchange-traded funds (ETFs). *Journal of Financial Economics* 125:537–60.

Dannhauser, C. D., and S. Hoseinzade. 2022. The Unintended Consequences of Corporate Bond ETFs: Evidence from the Taper Tantrum. *Review of Financial Studies* 35:51–90.

Dow, J. 2004. Is liquidity self-fulfilling? *Journal of Business* 77:895–908.

Easley, D., D. Michayluk, M. O'Hara, and T. J. Putniņš. 2021. The Active World of Passive Investing. *Review of Finance* 25:1433–71.

Edelen, R. M. 1999. Investor flows and the assessed performance of open-end mutual funds. *Journal of Financial Economics* 53:439–66.

Ernst, T. 2022. Stock-Specific Price Discovery From ETFs. *Working paper, University of Maryland*.

Evans, R., R. Moussawi, M. S. Pagano, and J. Sedunov. 2021. ETF short interest and failures-to-deliver: Naked short-selling or operational shorting? *Working paper, University of Virginia*.

Foucault, T., O. Kadan, and E. Kandel. 2005. Limit order book as a market for liquidity. *Review of Financial Studies* 18:1171–217.

Foucault, T., and C. A. Parlour. 2004. Competition for listings. *RAND Journal of Economics* 35:329–55.

Friewald, N., R. Jankowitsch, and M. G. Subrahmanyam. 2013. Illiquidity or credit deterioration: A study of liquidity in the US corporate bond market during financial crises. In *Managing and Measuring Risk: Emerging Global Standards and Regulations After the Financial Crisis*, 159–200. Singapore: World Scientific.

Glosten, L., S. Nallareddy, and Y. Zou. 2021. ETF activity and informational efficiency of underlying securities. *Management Science* 67:22–47.

Green, J. R. 1973. Information, efficiency and equilibrium. Discussion Paper, Harvard Institute of Economic Research.

Halaburda, H., B. Jullien, and Y. Yehezkel. 2020. Dynamic competition with network externalities: how history matters. *RAND Journal of Economics* 51:3–31.

Hamm, S. 2014. The effect of ETFs on stock liquidity. Working paper, Tulane University.

Holden, C., S. Jacobsen, and A. Subrahmanyam. 2014. The empirical analysis of liquidity. *Foundations and Trends in Finance* 8:263–365.

Hortaçsu, A., and C. 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* 119:403–56.

Huang, S., M. O'Hara, and Z. Zhong. 2020. Innovation and Informed Trading: Evidence from Industry ETFs. *Review of Financial Studies* 34:1280–316.

Israeli, D., C. Lee, and S. Sridharan. 2017. Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies* 22:1048–83.

Johnson, W. T. 2004. Predictable investment horizons and wealth transfers among mutual fund shareholders. *Journal of Finance* 59:1979–2012.

Katz, M. L., and C. Shapiro. 1985. Network externalities, competition, and compatibility. *American Economic Review* 75:424–40.

Krause, T., S. Ehsani, and D. Lien. 2014. Exchange-traded funds, liquidity and volatility. *Applied Financial Economics* 24:1617–30.

Krishnamurthy, A. 2002. The bond/old-bond spread. *Journal of Financial Economics* 66:463–506.

Lettau, M., and A. Madhavan. 2018. Exchange-traded funds 101 for economists. *Journal of Economic Perspectives* 32:135–54.

Li, F. W., and Q. Zhu. 2016. Synthetic shorting with ETFs. Working paper, Hong Kong University of Science and Technology.

Longstaff, F., S. Mithal, and E. Neis. 2005. Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *Journal of Finance* 60:2213–53.

Lopez Avila, E., C. Martineau, and J. Mondria. 2023. When Crowds Aren't Wise: Biased Social Networks and its Price Impact. *Working paper, University of Toronto*.

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

Madhavan, A., and A. Sobczyk. 2016. Price dynamics and liquidity of exchange-traded funds. *Journal of Investment Management* 14:1–17.

Malamud, S. 2016. A dynamic equilibrium model of ETFs. Working paper, Ecole Polytechnique Federale de Lausanne.

Markovich, S., and Y. Yehezkel. 2022. Group hug: Platform competition with user groups. *American Economic Journal: Microeconomics* 14:139–75.

Marshall, B., N. Nguyen, and N. Visaltanachoti. 2013. ETF arbitrage: Intraday evidence. *Journal of Banking and Finance* 37:3486–98.

Menkveld, A. J., and M. A. Zoican. 2017. Need for Speed? Exchange Latency and Liquidity. *Review of Financial Studies* 30:1188–228.

Owen, G. 1977. Values of games with a priori unions. In *Mathematical Economics and Game Theory, eds. R. Henn and O. Moeschlin*, 76–88. Berlin: Springer.

Pagano, M. 1989. Trading volume and asset liquidity. *Quarterly Journal of Economics* 104:255–74.

Pagnotta, E., and T. Philippon. 2018. Competing on speed. *Econometrica* 86:1067–115.

Rakowski, D. 2010. Fund flow volatility and performance. *Journal of Financial and Quantitative Analysis* 45:223–37.

Wermers, R., and J. Xue. 2015. Intraday ETF trading and the volatility of the underlying. Working paper, University of Maryland.

Yao, C., and M. Ye. 2018. Why trading speed matters: A tale of queue rationing under price controls. *Review of Financial Studies* 31:2157–83.

Zhu, H. 2014. Do dark pools harm price discovery? *Review of Financial Studies* 27:747–89.