The Active World of Passive Investing☆

David Easley[1], David Michayluk[2], Maureen O'Hara[1], and Talis J. Putnins[2,3]

[1]*Cornell University,* [2]*University of Technology Sydney,* [3]*Stockholm School of Economics in Riga*

Forthcoming, *Review of Finance*

July 2021

Abstract

We investigate the new reality of exchange-traded funds (ETFs). We show that most ETFs are active investments in *form* (designed to generate alpha) or *function* (serve as building blocks of active portfolios). We define a new activeness index to capture these dimensions, finding that the cross-section of ETFs is now increasingly characterized by highly active investment vehicles. Active-in-form ETFs have positive flow-performance sensitivity, charge the highest fees among ETFs, and have high within-portfolio turnover. Active-in-function ETFs have more concentrated holdings, less within-portfolio turnover, but higher turnover in the secondary market. We show how more active ETFs are gaining market share over less active ETFs, leading to competitive fee pressure both within the ETF space and across the investment management industry. We suggest that the growing activeness of ETFs may assuage concerns about ETFs harming price discovery.

JEL classification: G11, G14, G23

Keywords: exchange-traded fund, ETF, Activeness Index, passive investing, index investing.

☆ We are grateful for the comments of the editor (Marcin Kacperczyk), two anonymous referees, Pedro Matos, Luboš Pástor and conference/seminar participants at UBS, Chi-X Australia, EPFL Lausanne, the University of South Australia, the Stockholm School of Economics in Riga, FRIC 18 at the Copenhagen Business School, and the WFA 2019 meetings. We thank Jennie Bai, Alexi Savov and Thomas Philippon for sharing data. Putnins gratefully acknowledges funding from the Australian Research Council under grant ARC DE150101889.
Email addresses: dae3@cornell.edu, david.michayluk@uts.edu.au, mo19@cornell.edu, and talis.putnins@uts.edu.au.

1. Introduction

Worldwide, there are approximately 43,192 publicly listed stocks but there are estimated to be 3.3 million indexes, many of which are related to some sort of ETF or index fund investing.[1] This staggering disparity reflects the new reality of investing that for many investors the building blocks of their portfolios are index vehicles, not individual stocks. While large, passive index products continue to flourish, they are now joined by more specialized industry and characteristic-based ETFs, providing exposure to everything from artificial intelligence (tickers AIEQ, BIKR, and ROBO), to obesity-related investments (SLIM), to products favored by millennials (GENY and MILN), to companies with alignment to biblical values (BIBL), to meme stocks and social sentiment (BUZZ).[2] The rise of factor investing and smart beta strategies fostered a new generation of ETF products characterized by changing index constituent weights and active (even daily) rebalancing. And now, an even newer generation of "active ETFs" feature portfolio managers who make security selections that may not be constrained by any benchmark or even visible to market participants.[3]

In this paper, we investigate the new reality of "passive" investing. When passive investing was viewed as synonymous with index mutual funds, the dichotomy between passive and active investing was clear: select a one-size-fits-all, low-cost market index fund or invest in a higher-alpha but higher-cost selectively chosen portfolio of assets. The equilibrium size of passive and active investing would be determined by the relative costs of providing investment products and the cost and availability of information needed to generate alpha. But the rise of ETFs undermines this simple depiction. ETFs provide a variety of benefits that influence both the cost of providing investment products as well as the ability to craft alpha-seeking portfolios. Now there is a middle ground, one occupied by investment products with both passive and active features. As we demonstrate, this ETF incursion into the "active" management sphere has important implications for both the evolution of investment products and the nature of active management in the economy.

But what exactly does activeness mean in the ETF context? We propose a paradigm based on *form* and *function*. While some ETFs are active in *form* meaning that the portfolio of assets tracked by the ETF is chosen with the objective of generating alpha, other ETFs are active in

[1] Data from "Indices: The Tug of War between the Popular and the Personalized", Curatia news brief, June 22, 2018.
[2] Sadly, the spirits and whiskey ETF (WSKY) ceased trading on June 15, 2018.
[3] See Mackintosh (2017) for discussion of active ETFs and Hill (2016) for insights into index strategies using ETFs.

function meaning that they are used by investors as building blocks of active portfolios. Their ability to be easily and cheaply traded is what sets ETFs apart from index mutual funds and enables activeness in function. We extend the Berk and Green (2004) model of mutual fund management to incorporate these new investment products, demonstrating a role for what we call “aggressive-passive” investing with ETFs. Our model suggests a variety of hypotheses regarding fees, sizes, and fund flows across these investment products and their competitive impact on investment management. We also consider the larger question of why activeness matters for the market more generally. We argue that the growth in ETFs, by filling the void in the spectrum between active and passive products, need not change the equilibrium level of activeness in the market.

To investigate these hypotheses, we propose a new empirical metric to capture ETF activeness, and to distinguish active ETFs into those that are active in *form* and those that are active in *function*. This metric, which we call the Activeness Index, provides a lens with which to view the evolution of ETFs from their origin as simple index products to their modern incarnation as complex investment vehicles. This metric also gives us means to test the empirical implications of our model and to examine how the rise of ETFs has affected the competitive landscape of investment management.

Our research yields a variety of results. We show that most ETFs are fairly active, this activity level is increasing over time, and more active ETFs are gaining market share at the expense of less active ETFs. Consistent with our model, we find that broad, passive ETFs tend to be larger, less numerous, and charge lower fees than the more active ETFs. More intriguing is that these passive ETFs trade less actively on secondary markets than ETFs that are active in form or in function. Moreover, the “spread” between active ETF fees and passive ETF fees on a size-weighted basis is decreasing over time, consistent with increasing competition among ETFs.

We find that ETFs, like mutual funds, have positive flow-performance sensitivity: investor funds flow disproportionately to ETFs with better relative performance. However, as predicted by our analysis, this positive flow-performance sensitivity is driven entirely by the active ETFs, and it is strongest among the group of ETFs that are active in form, further highlighting the similarities between such ETFs and active mutual funds.

We use cluster analysis to partition the active ETFs into those that are active in form and those that are active in function. We show that active-in-form ETFs tend to charge higher fees and have higher levels of portfolio turnover *within* the fund, consistent with active management on the part of the ETF issuer. In contrast, active-in-function ETFs are larger, lower cost and have higher turnover in secondary markets, consistent with active investing on the part of investors that dynamically adjust exposures in their active portfolios by using these ETFs as building blocks. Therefore, even ETFs in which the holdings are passively linked to an underlying index can contribute to informational efficiency through the active trading by investors.

We show how the cross-section of ETFs is now increasingly characterized by highly active investment vehicles. Active-in-form and active-in-function ETFs account for approximately 58% (15% + 43%) of the ETF market by assets, 93% (42% + 51%) by number, and 78% (10% + 68%) by dollar volume traded in the secondary market. This movement into more active ETFs contrasts sharply with what is happening in mutual funds, where modest fund flows into more passive mutual funds being swamped by outflows from more active mutual funds. These divergent dynamics underscore the competitive pressures in the investment management industry arising from ETFs, which we show is affecting fee differentials between mutual funds and ETFs.

This rise of ETF activeness is relevant for the current debate on market activeness: have active ETFs simply replaced actively managed mutual funds, resulting in little or no change to overall market activeness or, as some fear, does ETF "free riding" on the price discovery produced by active mutual funds undermine the efficiency of the market?[4] The complexity of this issue is beyond the focus of our analysis, but our findings on the increasingly active role of ETFs may assuage concerns about ETFs harming price discovery.

Our paper joins a growing literature investigating the impact of ETFs and index investing. A recent paper by Cheng et al. (2019) looks at how ETFs deviate from their benchmarks by using, for example, synthetic replications with affiliated bank sponsors, which can increase returns to the ETF sponsor but also raises concerns about conflicts of interest. The activeness we examine is different, but both papers make the point that ETFs can function in ways far removed from passive investments. Bhattacharya et al. (2017) find that retail traders tend to make similar mistakes trading ETFs as they do trading more traditional investment products, consistent with the notion that ETFs are treated as active investments by some traders. Levy and Lieberman (2016) argue that from an asset allocation perspective some traditionally passive investments can be used in active investment strategies, an insight akin to what we develop in this research.

[4] As an example of this negative view, see Fraser-Jenkins (2016).

The issue of investment activeness is also addressed in a recent paper by Pastor et al. (2020). These authors propose a broader activeness index including liquidity and diversification but restrict their analysis only to active mutual funds. They note that “we exclude index funds because the model is designed for active funds trying to outperform benchmarks.” Our analysis differs in that we do consider the various ways in which many ETFs are, in fact, active investments.

Theoretical research, discussed in more detail in the next section, looks at how indexing affects the informational efficiency of the market. Empirical research on this topic includes Kacperczyk et al. (2021), Ben-David et al. (2018), and Madhavan and Sobczyk (2016). At the individual stock level, Glosten et al. (2021) find that inclusion in an ETF improves the incorporation of earnings information. Huang et al. (2021) show similar positive information effects arising from industry ETFs, and Bhojraj et al. (2020) find positive information effects from sector funds, but negative effects from non-sector funds.[5]

This paper is organized as follows. The next section sets out why activeness matters, provides a framework for viewing investment products along a continuum of passive to active, and defines the concepts of active in form and active in function. Section 3 defines empirical measures of ETF activeness and characterizes the ETF activeness landscape. Section 4 tests our hypotheses, examining how active and passive ETFs differ in size, fees, and flow-performance sensitivity. Section 5 separates ETFs into active in form and active in function and investigates their differences. We provide evidence on the evolution of ETF activeness over time and its implications for both intra-ETF and industry competition in Section 6. Section 7 discusses implications for overall market activeness and efficiency. Section 8 is a conclusion.

2. Active Investing and Passive Investing

What characterizes an investment product as being active or passive and why does it matter? In this section we develop theoretical arguments to guide our empirical investigation. We first consider why the level of activeness matters, drawing on the literature that considers the equilibrium level of active and passive investing. We then adapt the canonical Berk and Green

[5] Index investing more broadly has been associated with larger bid-ask spreads (Israeli et al., 2017), increases in daily return reversals (Baltussen et al., 2019), short-term volatility (Coles et al., 2020), and a reduction in the usefulness of prices as signals (Brogaard et al., 2019).

(2004) model of investment management to include ETFs. Finally, we propose a paradigm of activeness in form and in function to capture the role played by active ETFs.

2.1 The Equilibrium Levels of Active and Passive Investment

Starting with the seminal work of Grossman and Stiglitz (1980), a large literature considers the optimal amount of information gathering in the economy. Put succinctly, if no one gathers information, then market prices are uninformative, and there would be a large return to an agent seeking out information to guide investment decisions. Conversely, if everyone gathers information, then market prices are highly informative and an agent could forego the costly acquisition step and free-ride off market prices. So, there must be an equilibrium amount of information gathering that balances these two effects and this, in turn, dictates the level of market efficiency.[6]

The more current context for this problem is the balance between active and passive investing. Active investing strives to generate alpha via selective stock picking using information. Purely passive investing involves holding the market (an index) and so eschews information gathering. A number of authors consider what changes the mix of active and passive investing and how that influences efficiency and other parameters of interest. Bond and Garcia (2018), for example, show that indexing (their metric for passive investing) induces more participation in the market by uninformed investors, which increases risk-sharing, and raises welfare. Malikov (2019), in a Grossman-Stiglitz framework, argues that more traders participating in passive investments enhances the returns to active traders.

More germane for our focus is research by Garleanu and Pedersen (2019) and Cremers et al. (2016) who examine facets of the decisions to invest passively or actively. Garleanu and Pedersen, using a rational expectations equilibrium model, show that if the cost of information falls, then the number of active managers increases, fees fall, and market efficiency increases. Whereas if the costs of passive investing falls, the number of active managers falls, active fees decrease by less than passive fees, and market efficiency decreases. Cremers et al. (2016) also consider substitutability between active and passive investing. They note that active managers

[6] The issue is a bit more complex because the original Grossman-Stiglitz paper hypothesized that if prices are fully revealing then no one would pay for information but then how did the prices get the information in the first place? The key to this paradox is that there must be impediments that result in a non-revealing equilibrium, which can produce the trade-off discussed here.

may adapt in part by becoming “closet indexers” which they argue would tend to reduce overall market efficiency. However, an offsetting effect is that increased competition from indexers might lead active funds to lower fees and provide stronger incentives to collect information and generate alpha, leading to higher efficiency. Their empirical work supports the latter effect.

What emerges from these theoretical papers is that the level of investing dedicated to alpha generation, i.e., “active” investing, matters for determining the level of market efficiency. Yet, there is a conspicuous blind spot in virtually all of this research: the “passive” investing considered does not allow for the possibility that it, too, may involve alpha generating activity. Thus, while ETFs that simply hold the market portfolio fit this narrow definition of passive funds, the new generation of ETF products does not. These new ETFs reduce the costs to investors through greater tax efficiency and liquidity, while also giving investors a greater ability to craft alpha-seeking portfolios.

Stambaugh (2014) in his AFA Presidential Address made the important point that active investing was becoming less active as measured by tracking error and stock holdings. It should not be surprising that ETFs are evolving in the other direction—becoming more active and less passive. This growing middle ground in investment management raises a variety of intriguing questions, not the least of which is how to characterize these new active-passive products. We turn to this issue in the next sub-section.

2.2 A Framework for Active, Passive, and Aggressive-Passive Investing

To understand the economic forces affecting the evolving forms of active and passive fund management, we need a model capable of including a spectrum of funds from active to passive. We modify the model of mutual fund management developed by Berk and Green (2004) and expanded by Berk and van Binsbergen (2015) to include passive investments and we use this framework to show how ETFs can be characterized in equilibrium.

The model focusses on the decisions of a fund manager with respect to fund size and fees. Let the gross investment return that a manager earns from a dollar amount q invested in the fund be given by

$$qR + q(a - bq) \tag{1}$$

where R is the benchmark return, a is the net of transaction cost amount the manager earns from skill, and b reflects the decline in returns due to size.[7] The investment return a fund earns per dollar under management is thus

$$R + (a - bq) \ . \tag{2}$$

The parameters in these returns, (a_A, b_A) for an active manager and (a_P, b_P) for a passive manager, have different magnitudes and interpretations. For an active manager, the skill term a_A is positive. The b_A term is also positive as it reflects declining value creation with size, which arises either from managers using their best ideas first so that growth necessitates choosing lower NPV investments, or as noted in Berk and Green (2004, p. 1273) it "can be due to larger trades entailing a larger price impact and therefore larger execution costs."[8] For a purely passive manager tracking a market index, a_P can take a small positive value (allowing for the skill in designing an index or a rebalancing algorithm), and b_P is positive, but very small, reflecting increasing trading cost as the fund becomes a larger share of the market.[9] Thus, we assume $a_A > a_P$ and $b_A > b_P$.

With these parameter assumptions, a small active fund would earn a larger excess return (its return minus the benchmark return R) per dollar invested than would a small passive fund. For larger fund sizes, this return differential falls as $b_A > b_P$, and we assume that for a large enough fund size a passive fund would earn more per dollar invested than would an active fund. That is, active funds do not always earn more than passive funds.[10] If this were not the case, we would not expect to see any passive funds and as we point out in the previous section, returns earned by funds adjust to provide a balance between active and passive investing. Formally these parameter restrictions are

$$a_A > a_P, \quad b_A > b_P, \quad \frac{a_P}{b_P} > \frac{a_A}{b_A}. \tag{3}$$

[7] We focus on a static model as it captures the particular trade-offs we wish to explore between active and passive fund size and fee-setting. Berk and Green (2004) also consider a more dynamic setting, and we discuss how their dynamic results relate to our results later in this sub-section.

[8] Notice that these parameters can also link back to the factors affecting active and passive investing discussed in the previous sub-section. There we pointed out that the cost of gathering information and the cost of running the fund affected the amount of active vs passive management. This information cost is related to a as a higher a means the manager is better able to extract information to generate alpha.

[9] See Pedersen (2018) who notes that index investors can lose to active investors from price pressures due to index rebalancing. We thank the referee for raising this point.

[10] This is consistent with Sharpe's (1991) Arithmetic of Active Management in that after deducting fees the active manager can underperform the passive manager. Note, however, that we argue here that this occurs at very large size where the market impact costs of the active fund increasingly exceed those of the passive fund. As we discuss in the next section, we agree with Pedersen (2018) that passive trading is mischaracterized and so active management can add value.

Increasingly in the asset management industry ancillary revenue may accrue to the fund, which we denote as a fixed return of $l \geq 0$ per dollar invested in the fund. This ancillary revenue derives primarily from two sources. First, the fund may lend out shares, receiving a small fee per share. Second, the fund family may benefit from attracting money to this fund as this may lead to spillovers to other funds run by the fund family. We use l to represent all potential non-investment per dollar revenues, and discuss later how these revenues may accrue differentially in active and passive fund settings. We assume that some fraction $\gamma < 1$ of this ancillary revenue is returned directly to investors, with the remaining fraction retained by the fund.[11]

The fund manager charges a fee of f per dollar under management. We represent the amount investors can earn per dollar of investment without using this manger's fund by $R - \delta$, where $\delta \geq 0$ represents the gap between the market return and the amount that investors can earn using the best alternative to the manager's fund. This best alternative could be investing on their own or in another fund.[12] Investors will be willing to put money into the manager's fund if it earns more than $R - \delta$ and they may choose to withdraw funds if it earns less than $R - \delta$. So in equilibrium for any fund with return parameters a and b we have[13]

$$R + (a - bq) - f + \gamma l = R - \delta \tag{4}$$

where the left hand side is the investor's return per dollar invested in the fund. This investor participation constraint implies that the fee the manager can extract if the fund is of size q is

$$f = (a - bq) + \gamma l + \delta \tag{5}$$

The manager can have any size fund (up to the size at which the fund only earns $R - \delta - l$) he or she wants—by setting a low fee, the fund delivers high after-cost after-fee return, which attracts investors and increases size, while increasing the fee reduces the size. In addition, there is

[11] In practice, these stock lending fees are typically allocated in part to the investors and in part to the manager. For example, BlackRock notes that "BlackRock publishes security lending revenues in the fund's annual shareholder report and in [each fund's Statement of Additional Information] SAI and includes a separate line item that details BlackRock's portion of the revenues". For more discussion, see page 3 of https://www.blackrock.com/us/individual/literature/brochure/us-retail-securities-lending-brochure.pdf. See also Blocher and Whaley (2016) and Madhavan (2016).

[12] If investors could earn the benchmark return after transaction costs without using any fund then a passive fund could produce value for them only through its small return to skill and their share of ancillary revenue. This small revenue does not seem adequate to explain the rise of passive funds. So we expect a positive discount from the benchmark return, but our analysis does not depend on this. Alternatively, $R - \delta$ could be the best return that investors can earn using another ETF or mutual fund. The manager's fund clearly needs to pay this market determined best alternative to investors. Industry conditions and competition between various types of funds will determine this best alternative fund and thus influence how a manger structures a fund.

[13] We suppress fund type subscripts for clarity in these calculations.

a fixed cost, $F \geq 0$, of running the fund which the manager pays out of fees. This fixed cost is likely to be smaller for passive funds than for active funds.

The manager's decision problem can be formulated as either selecting f which, in turn, determines q, or as selecting q and thereby determining f. We focus on the latter formulation so the manager will choose a fund size q to maximize the value of running the fund

$$MAX(qf + q(1-\gamma)l) - F \tag{6}$$
$$\text{s.t. } q \geq 0, f = (a - bq) + \gamma l + \delta.$$

After substituting in the constraint, the unconstrained objective function is[14]

$$-bq^2 + q(a + \delta + l) - F. \tag{7}$$

If it is optimal for the fund to operate, the solution to this problem (which requires $b > 0$) is

$$q^* = \frac{a + \delta + l}{2b} \tag{8}$$

and the fee the manager charges is

$$f^* = \frac{a + \delta - l}{2} + \gamma l. \tag{9}$$

It is apparent that γ, the sharing rule on how ancillary revenues are split between the fund and investors, does not affect the optimal size of the fund q^*, or the net fee, $f - \gamma l$, paid by the investors.

The manager's maximized value of running the fund is

$$\frac{(a + \delta + l)^2}{4b} - F. \tag{10}$$

Provided $(a + \delta + l)^2 > 4bF$, the fund operates and the manager earns a positive profit. Ancillary revenue plays an interesting role here as, for a large enough l, it can be optimal for the manager to charge a zero fee, or even to pay investors to invest in the fund.[15] Calculation shows that as l increases, the optimal size of the fund increases and the fee charged to investors decreases.

With this fund size and fee, the net return per dollar invested earned by the investor is $R - \delta$. If $\delta = F = l = 0$, then the fund size and fee is the same as in Berk and Green. We focus on

---

[14] The two problems are identical and yield the same solutions for f and q. But this formulation reveals the equivalence between the manager's problem and the textbook monopolist problem of choosing a quantity to produce given a linear demand with intercept $(a + \delta + \gamma l)$ and slope $-b$.

[15] This is currently the case for 16 of the 737 ETFs on Forbes "Best ETFs" list, where the fund revenues from stock lending exceeds the fees the fund charges the investors, see "The Stock Lending Bonanza in your ETF", www.forbes.com/sites/baldwin/2018/06/26/the-stock-lending-bonanza-in-your-etf/#3b9a7b4e4841

the static problem confronting the manager, not on the dynamic multi-period model with shocks to fund flows. In a multi-period setting, the optimal fee changes each period and Berk and Green demonstrate that in equilibrium the active manager would be indifferent between this outcome and an alternative strategy of charging a non-time varying fee and then investing any funds that flow in above the fund's optimum size in a passive investment. Thus, dynamic issues muddy even more the distinction between active and passive funds.

Using this model for guidance, we now evaluate how active and passive funds differ in equilibrium. Two hypotheses immediately arise:

Hypothesis 1: In equilibrium, passive funds will be larger than active funds.

Hypothesis 2: In equilibrium, passive funds will have lower fees than active funds.

Our assumption (3) on parameters imply $q_P^* > q_A^*$, so in equilibrium passive funds will be larger than active funds. It is easy to show that the larger the fund, the lower the fee, so the second hypothesis follows from simple arithmetic. This latter result may be reinforced due to ancillary income from stock lending being a much larger activity for passive funds than it is for active funds.[16] In this case, $l_P > l_A$, which further reinforces the equilibrium size gap between $q_P^* > q_A^*$ and results in the passive fund offering even lower fees than the active fund.

This relation between the sizes of funds is illustrated in Figure 1. The blue solid curve illustrates the total return to the manager of an active fund as a function of the size of the fund, while the green dotted curve illustrates this relation for a passive fund. The manager chooses the size of the fund to maximize his total return so the passive fund is larger than the active fund.

< Figure 1 here >

In our model, as in the standard Berk and Green model, more talented active managers pick better investments, yielding higher returns for any given size of the fund. If investors do not initially know managers' skill levels then money will flow to active managers who generate higher returns as these returns are signals of skill. This causes subsequent performance to fall due to the

[16] An excellent discussion of stock lending by index funds and ETFs is "Examining the Risks and Rewards of Securities Lending" available at http://www.morningstar.com/904334/examining-the-risks-and-rewards-of-securities-lending.html

decreasing returns to scale. In equilibrium, the return to investors net of fees will be the same, but better managers will have larger funds. Thus, the active fund management industry will feature many funds of different sizes, money will chase past returns, managers will not outperform their benchmarks, and any after-fee excess return (the "alpha") will not reflect manager ability.

2.3 The Many Shades of ETF Activeness – Active in Form and Active in Function

What might this equilibrium look like in a world of ETFs? Consider a passive ETF based on a broad market index. ETF manager skill will play a relatively small role because the universe of stocks in the ETF and their weights are given (but the need to handle rebalancing suggests some skill is still involved). And, since the ETF is essentially holding the entire market, the gross return to the ETF should be the same regardless of the scale of the fund, although the ability to lend the underlying securities could add substantial ancillary revenues to the overall return. Scale may affect the net return to the ETF, however, due to (i) the greater operational efficiencies that come with greater size; and (ii) the (potentially small) transaction fees connected with trading larger and larger amounts of the market. In equilibrium, our model predicts the following:

Hypothesis 3: For passive ETFs based on a broad market index:

A. funds do not chase returns but rather flow to the largest (and therefore cheapest) ETFs;
B. there will be a few large passive ETFs;
C. competition (lowering δ) would force fees to fall.

This characterization of passive ETF investment is typified by Vanguard's Total Stock Market Shares Index ETF (VTI) which charges a remarkably low 3 basis point fee on its $216.4 billion asset size, the even larger iShares S&P 500 Index ETF (IVV) with $253.44 billion in assets and fees of only 3 basis points, or the very first ETF, the SPDR (SPY) with assets of $337.2 billion. This fund charges a slightly higher fee of 9 basis points, reflecting its near monopolistic role in S&P index arbitrage against the e-mini futures contract. The massive scale of these funds is consistent with our hypothesis of truly passive large-scale index funds being dominated by a few fund providers.

Focusing on large equity indexers, however, misses an important point: with $7 trillion in global ETF holdings spread across more than 5,000 ETFs, the ETF industry is now much more diverse (and considerably less passive).[17] Many ETFs begin with designer indices intended to

[17] See "ETF assets reach $7tn milestone", Financial Times, September 9 2020.

capture better performance. The rise of "smart beta" ETFs in which the holdings are adjusted to capture factor exposures, the development of leveraged ETFs whose holdings are adjusted to keep leverage constant, and even the rise of industry ETFs all suggest the characterization of ETFs as passive investments is far from accurate—or as Sharpe (1991) noted, "the passive manager in question may not be truly passive."

Consider a smart beta ETF focusing on a momentum factor. Unlike the passive ETFs depicted above, this ETF features skill on the part of the manager (or sponsoring firm) in designing an algorithm to select stocks and optimally adjust the portfolio over time. The ETF may also exhibit decreasing returns to scale, because greater scale can result in larger price impacts and execution costs when rebalancing the portfolio following the fund's algorithm. These features suggest that value creation by such an ETF manager leads to an equilibrium closer to that of active management than passive management—the "aggressive-passive" middle ground. Figure 1 illustrates how these aggressive-passive funds fit into our model. The red dashed curve illustrates the total return to the manager of an aggressive-passive ETF as a function of the size of the fund.

Our model yields five hypotheses predicting how passive ETFs, aggressive-passive ETFs, and active mutual funds should differ. First, consider an aggressive-passive ETF run by a manager who attempts to exploit investing skill versus a passive ETF that acts as if the manger has little or no investing skill and instead strictly follows an index. In our model, this difference is reflected in a higher level of the skill parameter, a, and a higher level of the cost, b, of implementing the aggressive-passive ETF strategy. Comparative statics shows that a one-unit increase in a increases the optimal fee by ½ and that this optimal fee is unaffected by b. So, relative to a passive fund run by an otherwise similar manager, we expect the aggressive-passive fund to charge a higher fee. If over time the information the active manger exploits becomes cheaper and easier to incorporate into an investment strategy then the manager's effective skill parameter, a, should increase and we expect these more active ETFs to grow relative to passive ones. Their relative sizes, however, are determined jointly by the parameters of the model and if these parameters satisfy our assumptions then the passive ETF will be larger than the aggressive-passive ETF.

Second, adding "aggressive-passive" ETFs to the universe of funds reveals a very different ecosystem. These funds will be smaller than purely passive vehicles and, depending on the skill and scale parameters, can approach the same profiles as active funds. Our model suggests that fund flows for these products will chase performance, in contrast to the lack of return chasing for the broad market index ETFs.

Third, if we compare an active ETF and an active mutual fund, we again expect both the effective skill and cost parameters to differ. A mutual fund does not have to disclose its positions with the same frequency as an ETF and this makes it easier for an otherwise identical manger to exploit investing skill resulting in a smaller effective skill parameter, a, for the ETF. However, the cost of implementing a strategy may also differ across the two types of funds. If the ETF has a smaller cost, b, this would tend to increase its optimal size, but it does not affect the fund's optimal fee. Overall we expect to see active ETFs have lower fees than similar mutual funds, but their relative sizes are indeterminate.

Fourth, competition between various types of ETFs and mutual funds affects an individual manager's choice of fees through its effect on the discount investors receive relative to the market return (δ in our model). Investors now have a wide array of alternative investment opportunities ranging from low cost passive ETFs to higher cost active ETFs and mutual funds. This increased opportunity set should result in a smaller δ and thus lower fees for all funds (a one unit decrease in δ reduces the optimal fee by ½); smaller funds (a one unit decrease in δ reduces the optimal fund size by ½ b); and a lower value of the funds to managers. If passive funds are already charging nearly a zero fee, and are unwilling or unable to charge a negative fee, then our model predicts the increased competition should decrease fee differences between more active ETFs and purely passive ETFs. So, over time we expect the difference in ETF fees to decline.

Fifth, it is plausible that over time the fixed cost of running a fund, whether it be a mutual fund or an ETF, declines as a result of technological advances. This has no direct impact on the optimal fee or size for a fund, but it will have an indirect impact through its effect on the market for funds. Lower fixed cost will result in entry of funds that would otherwise not find it profitable to operate. These funds have managers with lower skill or have higher trading costs than existing funds. As these funds enter the market, competition for dollars to invest increases and we expect the discount (δ) relative to the benchmark to decline. This will result in lower fees for all funds and a smaller optimal fund size; there will be more but smaller funds than before.

These hypotheses are summarized below:

Hypothesis 4: Active ETFs will have higher fees, but smaller size than passive ETFs.[18]

Hypothesis 5: Fund flows for active ETFs will chase performance.

Hypothesis 6: Active ETFs will have lower fees than mutual funds, but their relative size is indeterminant.

Hypothesis 7: Increased competition between active ETFs and passive ETFs will lower the fee differentials between these ETF products.

Hypothesis 8: Entry into asset management will result in lower fees for all funds and a smaller optimal fund size.

The aggressive-passive (henceforth simply called active) ETFs discussed above can be quite diverse. Consider, for example, the AdvisorShares (VEGA) STAR Global Buy-Write ETF. This ETF uses a proprietary investment strategy, which it terms the Volatility Enhanced Global Appreciation ("VEGA") strategy. It uses both ETFs and individual equities to implement its tactical allocation strategy and it does not track any specific index. It has an annual management fee of 135 basis points, annual expense ratio of 203 basis points, and $15 million in assets under management.[19] Clearly, this ETF is much closer to the active end of the spectrum than to the passive end.

To capture the diversity of this middle ground, we propose a paradigm for these active ETF products based on form and function. In particular, while some ETFs are active in *form* meaning that the portfolio of assets tracked by the ETF is chosen with the objective of generating alpha, other ETFs are active in *function* meaning that they designed to be used by investors as building blocks of active portfolios. The smart beta and VEGA STAR examples above are active in *form*—they involve strategies intended to directly generate alpha for investors.

Alternatively, we could look at how ETFs are used by investors—the ETF *function*. Increasingly, both professional and individual investors are using ETFs, rather than individual securities, to form portfolios. As Cong and Xu (2016) show, such a portfolio approach explains the increasing use of ETFs for factor investing. Market commentary also highlights this active-in-function property of ETFs, pointing out that ETFs are "used as Lego-like building blocks for

[18] Note that this hypothesis is just Hypotheses 1 and 2 applied to ETFs.
[19] As at September 2019, data from Bloomberg.

strategic portfolios or tactical punts, by anyone from sophisticated hedge funds to ordinary retail investors."[20]

This functional approach to investment management recognizes that the greater liquidity, low transaction fees, and tax efficiency of ETFs provide new building blocks for implementing investment strategies. What matters in this context is the exposure that can be obtained with a particular ETF and its contribution to the investor's total portfolio, not the particular properties it has as a stand-alone investment. Linking this to our model above, the skill element here is designing ETFs as vehicles that can be used by investors for selecting exposure to particular industries or factors while also optimizing the diversification and implementation costs. That skill can be from the index designer in carefully tailoring an index to efficiently achieve a desired exposure, or the ETF manager in strategically selecting which of the many specialized indexes to follow, or both.[21] Upon having designed these ETFs with tailed exposures, active trading of those ETFs by investors can facilitate price discovery and contribute to informational efficiency even though the ETFs themselves passively track the underlying indexes. In this respect, many ETFs function as active investments even if their behavior is passively linked to that of a specific index.

To give an example, consider the hedge fund Twin Tree Management LP that has approximately $2.5 billion in assets under management and is regarded as a high-turnover, large-cap focused, highly active fund. The fund's 13F filings with the SEC indicate that of its top 20 reportable positions, half are ETFs and half are stocks, with the ETFs making up 61% of the value and the individual stocks making up the remaining 39%.[22] Clearly, ETFs play an important role as the building blocks of this hedge fund's portfolio. The ETFs in their top-20 holdings include five sector ETFs (covering semiconductors, oil & gas, metals and mining), a real estate ETF, two commodity ETFs (gold and silver), and two country/region ETFs (Brazil and Europe/Australasia). That combination of ETFs provides very specific active exposures to various segments of the market, reflecting the fund's strategic, active asset allocation decisions.

---

[20] See Financial Times, December 20, 2018, "Passive attack: the story of a Wall Street revolution." See also Hill (2016), for a discussion of strategic and tactical uses of ETFs.

[21] In the interests of simplicity, the model does not have a separate role for an index designer. One can consider the ETF manager skill in the model as including skill in index design, i.e., the index designer and ETF manager are treated as one agent. Index designers will typically charge a fee to ETF providers to use the index, which is how they monetise their skill, which the ETF ultimately passes on as a fee to investors much like a skilled fund manager monetises their skill through the fees they charge investors.

[22] As at June 30, 2019, data from Factset.

ETFs can also be inputs in more complex trading strategies, particularly by institutional investors. Huang et al. (2021) investigate industry ETFs, focusing on their role as hedging vehicles. They show that hedge funds use industry ETFs in a "long the stock, short the ETF" strategy, essentially using the ETF to hedge out industry risk from their underlying long positions in individual stocks. An interesting finding of that paper is that hedge funds trade more aggressively when they can use an industry ETF to hedge, which in turn allows information to be incorporated more quickly into asset prices. Far from playing a passive role, ETFs now function directly and indirectly as active investment products.

This evolution, from purely passive index products to active investment tools, has important implications for investors, the investment management industry, and the market. For example, exactly how active are individual ETFs and where do they place on the active-passive spectrum? Much like closet indexing in mutual funds, individual ETFs may be far more active than investors and policymakers realize. Can this activeness, combined with the decreasing costs arising from technology, introduce competitive pressures within ETFs and across the broad investment management industry? Moreover, to the extent that the growth of ETFs is increasingly driven by the "aggressive-passive" ETF middle ground, are fears of market inefficiency resulting from the transition from active to passive investing overstated? To address any of these issues, we need to measure empirically the "activeness" of ETFs, an issue we turn to in the next section.

3. How "Active" are ETFs?

3.1 Data

We analyze U.S. ETFs that hold US equities during the period from 2000 to 2017. We combine data from several sources. First, we identify U.S.-listed ETFs using Center for Research in Security Prices (CRSP) daily security files (share code 73 identified ETFs). We screen these ETFs, keeping only those that focus on U.S. equities and are domiciled in the U.S.[23] We obtain daily data on each ETF's volume, market capitalization, price, and shares outstanding from CRSP.

We obtain the holdings of each ETF from the Thomson Reuters Fund Holdings database, which is based on the quarterly filings each fund makes to the US SEC.[24] To merge the CRSP

[23] We eliminate funds with Investment Objective Codes 1, 5, 6, and 8. In this step, we discard bond funds, commodity funds, and funds that predominantly hold non-US stocks.
[24] We use a copy of this dataset downloaded after WRDS and Thomson Reuters, in 2018, corrected issues with missing data.

data with the ETFs holdings data, we use the MFLINKS tables developed by Russ Wermers and Wharton Research Data Service (WRDS). We use the latest available version of these tables, which include links through to the end of 2016 (see Cao et al., 2017). We adjust the holdings data to account for funds that have multiple share classes similar to Ben-David et al. (2018).[25] Where holdings data are missing or incomplete, we supplement them with holdings data from the ETF Global database, where we also obtain fund fees. Finally, we add information about mutual fund fees from the CRSP mutual fund database.

In total, this gives us 452 distinct U.S. equity ETFs for which we have both holdings data and pricing/volume data. These ETFs have a combined market capitalization of $1.48 trillion at the end of December 2017. This sample is similar to that used by Ben-David et al. (2018).

3.2 Measuring ETF "Activeness"

A key issue in this paper is how to measure activeness. There are a variety of approaches in the literature. Kacperczyk et al. (2005) capture activeness in mutual funds using an Industry Concentration Index, which measures the concentration of mutual fund holdings in particular industries relative to the weights of those industries in the stock market. Ivkovic et al. (2008) measure the concentration of investor holdings of individual stocks using a Herfindahl Index. Cremers and Petajisto (2009) propose a measure of activeness for mutual funds based on how far a fund departs from its benchmark. Pastor et al. (2020) suggest an alternative measure for mutual fund activeness which incorporates the portfolio's liquidity and diversification. Each of these measures is designed to capture specific features of activeness, particularly as it relates to standard active management behavior.

The focus we have here is on the *form* and *function* of ETFs. To capture activeness in this sense, we start from the concept of what defines a passive investment. Restricting our attention to U.S. equities and considering a U.S.-based investor that invests only domestically, a completely passive investment as defined by Sharpe (1991) involves holding a portfolio of all U.S. stocks with weights proportional to their market capitalizations. Our notion of ETF activeness is the extent to which the ETF deviates from this completely passive strategy. ETFs can do so in two ways: (i)

[25] Specifically, when the aggregate value of a fund's reported holdings exceeds the market capitalization of the ETF by more than 20% (this happens when the fund has share classes other than the ETF) we scale back the holdings to match the market capitalization of the ETF, thereby considering only the holdings attributable to the ETF and not other (unlisted) share classes.

choosing a benchmark that embraces or departs from the market (active in function); and/or (ii) choosing holdings that depart from the chosen benchmark (active in form).

Figure 2 gives a schematic of how these components relate to overall activeness. Broad-based ETFs such as Vanguard's Total Stock Market Index Fund (VTI), which tracks the total US stock market by holding over 3,000 US stocks in market capitalization weights, come close to this completely passive portfolio. At the other end of the spectrum, highly specialized ETFs such as narrow industry ETFs, or smart beta ETFs that deliberately load on particular factors depart from the completely passive portfolio and instead constitute active bets on some segment of the market or factor exposure outperforming.

< Figure 2 here >

More formally, we define the Activeness Index of a given ETF, i, at a point in time, t, as:

$$ActivenessIndex_{i,t} = \sum_{s=1}^{N} \left| w_{i,s,t} - w_{market,s,t} \right| \tag{11}$$

where $w_{i,s,t}$ is the weight of stock s in fund i's portfolio at time t and $w_{market,s,t}$ is the weight of stock s in the value-weighted portfolio of all US listed stocks (the "market portfolio"). For an all-equity fund that has no leveraged or short positions, the Activeness Index lies between 0 and 1 (0% and 100%) and indicates the fraction or percentage of the fund's portfolio that differs from the passive market benchmark.

This Activeness Index is related to the measure proposed by Kacperczyk et al. (2005) in that fund holdings are compared to the market portfolio (not the fund benchmark) to gauge activeness, but it differs in that our measure considers holdings of individual stocks not holdings aggregated to industry exposures. The Activeness Index is also related to the approach used by Cremers and Petajisto (2009) to identify mutual funds that are "closet indexers"; the key difference is that we compare ETFs to the market portfolio rather than to the index they track. This design allows the Activeness Index to capture both activeness in *form* (when the ETF issuer seeks to create a portfolio that will generate alpha) and activeness in *function* (ETFs that track narrow segments of the overall market and are used as building blocks of active portfolios). In the former case, the active-in-form ETF manager is responsible for making active investment decisions in how to depart from the given benchmark, while in the latter case, investors make active investment decisions in selecting and then dynamically adjusting which exposures to include in their portfolios through combinations of active-in-function ETFs.

While the Activeness Index is our main measure, we supplement it with an alternative metric, which we term the "Active Return Deviation", based on how much the ETF return differs from the returns of the passive market portfolio:

$$ActiveReturnDeviation_{i,t} = Stddev(R_{i,d} - R_{market,d}) \quad (12)$$

where $R_{i,d}$ and $R_{market,d}$ are daily returns on ETF i's holdings and the market portfolio respectively.[26] We calculate (annualized) Active Return Deviation for each ETF at each quarter t using the daily returns during that quarter (holding fixed the ETF's portfolio weights from the end of the quarter).[27]

The difference between Activeness Index and Active Return Deviation is that Active Return Deviation picks up the covariance of active bets and is high when that covariance is high (when the bets are highly correlated). A very passive broad market ETF will have low Activeness Index and low Active Return Deviation because it closely replicates the market. An active ETF that departs from the broad market portfolio either in function or form will have a high Activeness Index. Such ETFs can have high or low Active Return Deviation depending on the correlation of the active positions. For example, an ETF that is very selective within industries (e.g., holds only the largest stock in each industry), but holds industries in their market weights, will have a high Activeness Index but a fairly low Active Return Deviation because most of the risk in the active positions is diversified away. In contrast, an ETF that is very selective across industries (e.g., holds only one industry) will have a high Activeness Index and a high Active Return Deviation because the risk in the active positions will not be diversified away.

3.3 The Extent of ETF Activeness

Figure 3 provides a static snapshot of the activeness of the US equity ETFs. The horizontal axis measures the Active Return Deviation of the ETFs, while the vertical axis measures their Activeness Index. Each circle on the plot is an ETF. For each ETF in each quarter, we measure the values of Activeness Index and Active Return Deviation and then take time-series averages for each ETF. In Panel A, larger circles are used for larger ETFs according to their assets under

---

[26] A variety of metrics in the literature for predicting mutual fund performance consider similar inputs. For example, Amihud and Goyenko's (2013) R-squared metric is obtained from a regression of fund returns on a multi-factor model. They interpret a lower R-squared as greater selectivity. Our measure captures differences in the standard deviation of returns, which we interpret as greater activeness of the ETF.

[27] We estimate Active Return Deviation only if there is a minimum of 30 valid daily returns for the given ETF in the given quarter and winsorize it at a maximum of 25% p.a..

management (AUM). In Panel B, larger circles are used for more actively traded ETFs measured by the ETF's secondary market turnover (volume scaled by shares outstanding).

< Figure 3 here >

The figure shows that ETFs are quite active. Most ETFs lie in the top two quadrants indicating that the holdings of these ETFs are far from merely replicating the broad market portfolio. Most ETFs have an Activeness Index above 50%, indicating that if their holdings were decomposed into a holding of the market portfolio (the passive part of their holdings) and a holding in a zero-investment long-short portfolio (the active part of their holdings), the value of the active bets would make up more than 50% of the ETF's value.

Furthermore, most ETFs have returns that deviate substantially from those of the passive market portfolio. Many have annual Active Return Deviation greater than 7% p.a. The absence of any ETFs in the bottom right-hand quadrant is simply a result of how the activeness metrics are constructed: it is not possible to generate a large Active Return Deviation without having holdings that substantially depart from the market portfolio. In essence, Figure 3 shows that most ETFs constitute an active bet, or a series of bets, on factors, individual stocks, or segments of the market.

When we consider the size of ETFs (Panel A), we see that the most passive ETFs (bottom left-hand quadrant) tend to be the big ones, consistent with Hypothesis 3. Such funds likely accrue operational efficiencies from becoming large and because they are fairly passive, they do not face the same concerns about alpha erosion as do more active funds. In this quadrant, we see funds such as Vanguard's Total Stock Market Index Fund (VTI). In contrast, the very active ETFs (top two quadrants, and in particular the top right-hand side quadrant), are far more numerous and tend to be smaller, results in line with Hypothesis 4. Such diseconomies of scale for active management are also predicted by Berk and Green (2004) and Pastor et al. (2020). There are, however, a few fairly large active ETFs with average AUM greater than $10 billion. These large active ETFs tend to be the active in *function* ETFs, such as sector ETFs, while the smaller active ETFs tend to be active in *form* (later we empirically separate these two types of activeness).

Figure 3 Panel B shows that in terms of secondary market turnover, more active ETFs tend to be more actively traded. In fact, many of the ETFs with a high Activeness Index have an annual turnover in excess of seven times, that is, their secondary market annual traded dollar volume is more than seven times their market capitalization. Such levels of turnover are considerably higher than the typical US stock, which has a turnover closer to one.

The high turnover of active ETFs is consistent with being used by investors to make relatively short horizon bets on segments of the market or on factors—what we refer to as activeness in function. The high turnover of active ETFs suggests that active ETFs might contribute to price discovery in individual stocks, as active bets made by buying and selling these ETFs leads to trading of the ETF constituent securities via the ETF creation/redemption mechanism and ETF arbitrage.[28]

< Figure 4 here >

Figure 4 further illustrates the composition of ETFs and their trading activity by levels of activeness. We assign each ETF to one of four baskets based on its average Activeness Index: (i) "Very Passive" (Activeness Index < 25%), (ii) "Moderately Passive" (25% < Activeness Index < 50%), (ii) "Moderately Active" (50% < Activeness Index < 75%), and (ii) "Very Active" (Activeness Index > 75%). We then measure the proportion of aggregate ETF AUM held by ETFs in each of the four activeness categories (Panel A), the proportion of ETFs in each category (Panel B), and the proportion of ETF secondary market traded dollar volume (Panel C). One ETF in particular, the SPY, is anomalous in terms of its turnover due to its role in S&P 500 index arbitrage against the e-mini futures contract. It has a turnover ratio of around 25, not because of active investment, but because of cross-security arbitrage and short horizon hedging. We therefore exclude SPY from the turnover measures in Panel C.

The figure shows that active ETFs dominate the ETF landscape. This tendency is particularly apparent in terms of the number of ETFs. Very Active ETFs alone account for 84% of the ETF population. Adding the Moderately Active ETFs brings the share of these "aggressive-passive" ETFs to 92% of all US equity ETFs. Because active ETFs tend to be smaller than their passive peers, their share of aggregate ETFs assets is smaller but nevertheless substantial. Active ETFs hold over half (58%) of aggregate ETF assets, with Very Active ETFs holding the majority of these (43% of aggregate ETF assets). The relatively large size of passive ETFs is apparent in the proportions—they account for a mere 8% of ETFs, but 42% of aggregate ETF asset value.

[28] For a discussion of this process see Lettau and Madhavan (2018).

The dominance of active ETFs is also apparent in secondary market trading. Active ETFs account for a much larger share of ETF traded dollar volume than their share of aggregate ETF asset value, suggesting they are more actively traded by investors than their passive counterparts. For example, Very Active ETFs account for as much as 86% of the dollar volume of ETFs traded in the secondary market, which is twice their share of aggregate ETF assets. Thus, aggressive-passive ETFs are active not only in their holdings, due to active decisions made by the fund manager or index provider (active in *form*), but also in the way they are used by investors who frequently trade in and out of active ETFs (active in *function*).

< Table I here >

Table I reports the distribution of the ETF Activeness Index in the cross-section of ETFs. The pooled sample descriptive statistics in Panel A confirm what was illustrated in the previous figures—most ETFs are highly active with a median Activeness Index of 93.1% and median Active Return Deviation of 8.8% p.a. The partition by size quartiles in Panel B shows that the largest quartile of ETFs is more passive than the smaller ETFs, consistent with our Hypothesis 4. For example, ETFs in the largest quartile have a mean Activeness Index of 78.3%, compared to 86.7% in the next quartile, and Active Return Deviation of 7.8% p.a. compared to 8.8% p.a. in the next quartile. This inverse relation between size and activeness arises from an absence of alpha for passive funds and thus economies of scale as the primary determinant of the optimal size for highly passive funds (see the equilibrium model).

Table I Panel C partitions ETFs by turnover quartiles and confirms that highly active ETFs tend to have the highest turnover. While the Activeness Index generally increases with turnover, the relation between Active Return Deviation and turnover is even stronger—the highest turnover quartile of ETFs have median Active Return Deviation almost twice as large as that of the lowest turnover quartile.

Finally, Table I Panel D shows that more recently launched ETFs are more active than earlier ETFs. In fact, ETFs launched in the last five years of our sample are extremely active with a median Activeness Index of 99.1% and a mean of 97.1%. This shift illustrates that while the first ETFs were largely passive investment vehicles, ETFs evolved into a more active form of investing. Thus, a spectrum of investment products is emerging, with new investment vehicles available to support active investing.

4. How do "Active" and "Passive" ETFs Differ?

Our analysis of active, passive, and aggressive-passive funds predicts that active ETFs should be smaller (Hypothesis 4) and charge higher fees in equilibrium (Hypothesis 4) compared to their more passive ETF counterparts. To test for these effects, we estimate cross-sectional regressions of ETF activeness (logit transformation of the Activeness Index) on various fund characteristics including size and fees. Our model also predicts that the fund flow-performance relation should differ between these active and passive investment products, with passive products having little or no fund flow sensitivity unlike active funds that will (Hypotheses 3 and 5).

The results in Table II Model 1 show that more active ETFs tend to be smaller, consistent with the theoretical model. In fact, fund size alone explains around one-third (32%) of the variation in ETF activeness. Model 2 in Table II shows that more active funds tend to charge higher fees, consistent with Hypothesis 4, and both size and fees remain significant when included together (Model 3). Figure 5 graphically illustrates the relations between ETF activeness, size, and fees. Panel A shows the tendency for more active ETFs to be smaller than the passive ETFs, while Panel B shows that generally, more active ETFs tend to charge higher fees.

< Table II here >

< Figure 5 here >

The results (Model 4 of Table II) also show that more active ETFs tend to have higher secondary market turnover (dollar trading volume scaled by market capitalization). Both activeness in form and activeness in function are likely to contribute to higher secondary market turnover but for different reasons. For active in function ETFs (e.g., narrow industry funds), high secondary market trading is a natural consequence of investors using such ETFs as building blocks of active portfolios. For active-in-form ETFs, if investors learn about the skill of the fund manager by observing realized performance, funds should flow from worse performing funds to better performing funds, generating secondary market trading of the active-in-form ETFs. The evidence in Table II is consistent with both effects.

Our model suggests that active investments would exhibit fund flows linked to performance (Hypothesis 5), but that this would not be the case for passive funds. Levy and Lieberman (2016) presented intriguing empirical evidence consistent with this notion. They examine data from mutual funds and ETFs and show that performance-related flows are stronger for more active groupings. The authors do not use either individual fund data or an activeness metric, and instead aggregate their data by segments. This aggregation may understate the strength of this flow-performance relation among ETFs.[29] Moreover, much has changed in the ETF universe since the period considered by these authors (their sample ends in 2012), with the rise of active ETFs particularly pronounced in recent years. A natural question is whether this performance-flows relation has changed as well.

We examine these flow-performance relations at a more granular level, focusing on individual funds and then contrasting passive ETFs with "active" ETFs. For active ETFs, we expect investor flows to be sensitive to past performance: when investors learn about managerial skill in generating alpha, flows should "chase" good past performance. In contrast, there is very little alpha in a truly passive, broad market ETF and so we expect little or no flow-performance sensitivity. For each ETF in each month of our sample we measure the ETF's relative performance ($Perf_{i,t}$) as the percentile of its return in the cross-sectional distribution of all ETF returns that month. We then regress monthly percentage flows into individual ETFs ($Flow_{i.t}$) on their past month's relative performance ($Perf_{i,t-1}$).[30]

< Table III here >

Table III reports the results of the flow-performance sensitivity regressions, with standard errors double clustered by ETF and date. Model 1 shows that in the pooled sample of all ETFs there tends to be a positive flow-performance relation. That is, investor flows tend to chase high

[29] For example, at a sector level some sectors may have great performance, others poor returns, so aggregating may obscure these differences. Even within size and style categories, aggregation may miss the activeness resulting from factor timing/ factor rotation strategies used by active in function ETFs.

[30] We measure flows using the changes in the ETF shares outstanding: $\%Flow_{i,t} = Shares_{i,t}/Shares_{i,t-1} - 1$ where $Shares_{i,t}$ is the number of outstanding shares in ETF i at the end of month t. Our measures of net flows as a percentage of the fund's assets and performance as the ranking (in our case, percentile) of the fund relative to its peers, mirrors the mutual fund literature (e.g., Sirri and Tufano, 1998; Huang et al., 2007). We winsorize the monthly percentage flows at −50% and +100%.

past ETF returns, similar to the performance chasing that is documented among investors in active mutual funds (e.g., Sirri and Tufano, 1998; Huang et al., 2007). The positive flow-performance relation is robust to fund and time fixed effects (Models 2 – 3).

However, we only expect a positive flow-performance relation in *active* ETFs. Therefore, in Model 4 of Table III, we include a dummy variable for active ETFs (those with an Activeness Index > 50%) and interact that dummy variable with the past performance measure to allow for different flow-performance sensitivities for active and passive ETFs. The results confirm our model's hypotheses: there is a significant positive flow-performance sensitivity for active funds (the positive coefficient on the interaction term $Perf_{i,t-1} \times D_i^{Active}$), yet *no* relation between flow and performance for passive funds (the near-zero coefficient on the variable $Perf_{i,t-1}$). Thus, the analysis indicates a strong heterogeneity among ETFs, with some behaving like active mutual funds in terms of how investors respond to performance (those ETFs with a high Activeness Index), while the passive ETFs show a complete absence of performance-chasing behavior among investors.

Among the active ETFs, we expect to see a stronger tendency for investors to learn from past performance and chase returns in ETFs that are active in *form*, that is, where the ETF seeks to generate alpha through active management within the ETF. To test this conjecture, we use a simple proxy for active-in-form ETFs based on the amount of turnover within the ETF portfolio. The Pastor et al. (2020) activeness measure also uses within-portfolio turnover, which they interpret as managers chasing opportunities to earn alpha.[31] Active-in-form ETFs are expected to have higher turnover within their portfolios due to active management within the ETF, whereas active-in-function ETFs can have little turnover within the ETF portfolio, and instead have highly active trading in secondary markets due to their function as active portfolio building blocks. We create a dummy variable for the subset of active ETFs that are within the top quartile of active ETFs by portfolio turnover, $D_i^{ActiveInForm}$, as a simple proxy for ETFs that are active in form. We then rerun the flow-performance sensitivity regressions for just the subsample of active ETFs, including an interaction term with this active-in-form proxy.

[31] Our measure of portfolio turnover follows the mutual fund literature, where portfolio turnover is the percentage of the fund's average net asset value that is replaced per annum (e.g., Chen et al., 2000). This measure considers changes to the ETF portfolio composition excluding those that arise from inflows and outflows.

The results in Model 5 show that the flow-performance sensitivity of active in form ETFs is about twice as strong as for active-in-function ETFs.[32] Model 6 confirms that the results are robust to using the full sample of ETFs in the regressions and two sets of interaction terms: there is zero flow-performance sensitivity for passive ETFs, a moderate (0.04) flow-performance sensitivity for active-in-function ETFs, and a strong (0.09) flow-performance sensitivity for active-in-form ETFs.[33] The flow-performance sensitivity for active-in-function ETFs (0.04) implies that if a fund's performance increases such that it moves up 10 percentiles (e.g., from the 50th percentile to the 60th percentile of ETFs), then the expected fund inflow in the following month is 0.4% (10×0.04) of its AUM. In contrast, for active-in-form ETFs, that same increase in performance is expected to result in fund inflow equivalent to 0.9% (10×0.09) of its AUM.

5. Active-in-Form vs Active-in-Function ETFs

We have discussed that ETFs can play a role in active funds management in two ways: (i) either the ETF's holdings are themselves actively managed by the ETF issuer with the objective of delivering an excess risk-adjusted return, what we term "active in form", or (ii) the ETF is designed to track a narrow segment of the market (e.g., a sector) and so can be used by active investors as an efficient way to place a bet on a particular exposure, what we term "active in function". The previous section suggested two empirical characteristics that can distinguish between these two forms of activeness: (i) portfolio turnover within the ETF, and (ii) the extent to which investor flows are sensitive to the fund's performance.

We now take these two characteristics and seek to separate empirically the ETFs that are active in form from those that are active in function. To do so, we employ k-means cluster analysis of the set of active ETFs (those with an Activeness Index > 50%). This form of cluster analysis is useful for identifying distinct groups in the data based on observable characteristics. In this case, we provide the algorithm with two observable characteristics (the portfolio turnover and flow-

[32] In Model 5, which includes only active ETFs, for active-in-form ETFs the flow-performance sensitivity is the sum of the coefficients of $Perf_{i,t-1}$ and $(Perf_{i,t-1} \times D_i^{ActiveInForm})$, i.e., 0.04+0.05=0.09, while for active-in-function ETFs, the sensitivity is the coefficient of $Perf_{i,t-1}$, i.e., 0.04.

[33] In Model 6, which includes all ETFs, for passive ETFs the flow-performance sensitivity is the coefficients of $Perf_{i,t-1}$, for active-in-function ETFs it is the sum of the coefficients of $Perf_{i,t-1}$ and $(Perf_{i,t-1} \times D_i^{Active})$, i.e., 0.00+0.04=0.04, while for active-in-form ETFs it is the sum of the coefficients of $Perf_{i,t-1}$, $(Perf_{i,t-1} \times D_i^{Active})$, and $(Perf_{i,t-1} \times D_i^{ActiveInForm})$, i.e., 0.00+0.04+0.05=0.09.

performance sensitivity of each fund) and get it to find the two groups of ETFs such that the within group variation in minimized and between group variation is maximized.

The output of the cluster analysis is striking. Figure 6 shows a clear separation of the ETFs into two groups. One group (blue circles) has high, almost always positive, flow-performance sensitivity, and a range of portfolio turnover levels including all of the highest portfolio turnover ETFs. That group is the active-in-form ETFs. The other group, active-in-function ETFs (red squares), has lower portfolio turnover and has flow-performance sensitivities that cluster around zero. Interestingly, while the cluster analysis algorithm was not told anything about the relation between portfolio turnover and flow-performance sensitivity, it identified the tendency for high portfolio turnover to be accompanied by high flow-performance sensitivity in one group of ETFs and low portfolio turnover to be accompanied by low flow-performance sensitivity in another group of ETFs.

< Figure 6 here >

To compare the characteristics of active-in-*form* and active-in-*function* ETFs, we take the set of active ETFs and regress ETF-level measures (size, fees, turnover, and so on) on an indicator variable for the active-in-form ETFs that are identified through cluster analysis. The results of these regressions are in Table IV Panel A, with Panel B further controlling for the differences in the activeness levels. The results show that there is a tendency for active-in-form ETFs to be smaller than active-in-function ETFs (around 32% smaller), although this result is only marginally statistically significant (at the 5% level in Panel A, Model 1).[34] There is also a significant difference in fees (Model 2), with active-in-form ETFs charging around 7 bps higher fees or 22% more than the average fee of active-in-function ETFs (32 bps). The higher fee for active-in-form ETFs reflects the additional skill and costs of active management within the ETF portfolio, as compared to constructing an ETF that is active in function. Pastor et al. (2020) find that more active mutual funds are smaller and charge higher fees, a finding mirrored here by the active in function ETFs.

There is a difference in the levels of secondary market trading activity of the two types of active ETF, although the difference is only statistically significant when controlling for the overall

[34] The coefficient −0.38 is the difference in log dollar holding value, which equates to $e^{-0.38} - 1 = -32\%$ lower dollar value of holdings in active-in-form ETFs.

level of activeness (Panel B, Model 3). The secondary market turnover ratio of active-in-form ETFs is around 19% lower than for active-in-function ETFs, consistent with the notion that active-in-function ETFs are more actively traded as the building blocks of active investor portfolios.[35]

In contrast, the turnover *within* the ETF portfolio, that is, the tendency for the ETF portfolio holdings to change (excluding the effects of flows) is significantly higher for ETFs that are active in form (Model 4). Pastor et al. (2020) find a similar relation for mutual funds, finding that more active mutual funds have higher within-portfolio turnover. The higher portfolio turnover of the active-in-form ETFs reflects the active management that occurs within the ETF. Finally, consistent with our earlier results, investor fund flows are significantly more sensitive to the performance of ETFs that are active in form (Model 5), where performance provides a signal about managerial skill, compared to ETFs that are active in function.

< Table IV here >

In Figure 7 we extend our earlier comparison of the prevalence of active vs passive ETFs to include a partition of the active ETFs into active-in-form and active-in-function types based on the cluster analysis described above. The figure shows that both active-in-function and active-in-form ETFs account for a substantial share of all ETFs, 51% and 42% respectively. However, because active-in-function ETFs tend to be larger, they account for a larger share of ETF assets (43%) compared to active-in-form ETFs (15%). Finally, because active-in-function ETFs are very actively traded in the secondary markets, they account for the dominant share of ETF traded dollar volume (68%).

< Figure 7 here >

6. The Evolution of ETF Activeness Through Time: Competitive Impacts

To capture the evolution of ETF activeness, Figure 8 repeats our two-dimensional mapping of ETFs on an Activeness Index / Active Return Deviation plane, but for four different points in time, each five years apart. In 2001, there are a few passive ETFs, which are relatively small, and a few active ones. By 2006, some passive ETFs have grown large, and the active ETFs have

[35] The coefficient –0.21 is the difference in log turnover ratios, which equates to $e^{-0.21} - 1 = -19\%$ lower turnover ratios for active-in-form ETFs.

become far more numerous. This trend continues into 2011. By 2016, we have several very large passive ETFs and a very large number of active ETFs. Many of the ETFs at this point are very active and some of these active ETFs have also attracted considerable investment (assets).

< Figure 8 here >

The tendency for newer ETFs to be more active than older ETFs is also reflected in Figure 9, which shows how the cross-sectional distribution of Activeness Index has evolved through time. The figure reiterates just how active ETFs are overall: even the 25th percentile of Activeness Index has remained above 50% throughout the past 17 years. The tendency for newer ETFs to be more active than older ETFs shifts the cross-sectional distribution of Activeness Index towards higher values through time. However, given that ETFs are already quite active overall, the changes in the cross-sectional distribution of activeness are only really apparent in the 10th, 25th, and 50th percentiles of the distribution, which tend to shift higher through time.

< Figure 9 here >

The growth of active ETFs speaks to the growing competition both within the ETF sphere and across the investment management universe. Technology lowered the cost of operating all types of funds, so even active ETFs tend to have fairly low management fees. For some investors, this can induce shifting from passive ETFs to these newer alpha-generating ETFs. Competition between active ETFs and mutual funds should be even more intense. As discussed earlier, mutual funds have some advantages and some disadvantages relative to active ETFs when it comes to turning information into alpha. As ETFs evolve into more complex products, and as technology lowers the cost of information gathering, we expect to see both disparate fund flows and fee pressure arising from greater competition between active and passive ETFs and between active ETFs and mutual funds.

To investigate these effects, we partition fund flows to ETFs into activeness categories: Very Passive, Moderately Passive, Moderately Active, and Very Active. We compute dollar flows for each fund i in each quarter t as follows (see Sirri and Tufano, 1998):

$$\$Flow_{i,t} = HoldingValue_{i,t} - HoldingValue_{i,t-1}(1 + R_{i,t}) \qquad (13)$$

where $HoldingValue_{i,t}$ is the dollar value of the fund's holdings of stocks at the end of quarter t and $R_{i,t}$ is the return on the fund's holdings during that quarter.

< Figure 10 here >

Figure 10 shows the cumulative dollar flows to ETFs of varying activeness during the past 17 years. While all activeness categories attracted net inflows of funds, the biggest inflows of funds have gone to the Very Active ETFs. In contrast, Very Passive ETFs received the least inflows. Consequently, not only have active ETFs become more numerous, they also collectively attracted more investment through time—in effect, winning the intra-ETF competitive battle and turning ETFs into increasingly active investment products.

The competitive battle between active ETFs and mutual funds reveals a different story. In a similar manner as above, we construct the flows to traditional US equity mutual funds segmented across activeness categories. These flows are illustrated in Figure 10 Panel B.[36] In contrast to ETFs, mutual funds in aggregate experienced net outflows during the sample period. Interestingly, the outflows are from the Very Active and Moderately Active mutual funds. In fact, the only mutual funds to receive positive inflows are the Moderately Passive and Very Passive ones, consistent with the technology-induced fall in operating costs attracting more investor funds. These passive mutual fund inflows, however, are smaller than the inflows to the ETF passive products, perhaps reflecting a lower overall cost structure of passive ETFs, or perhaps greater ancillary revenue accruing to ETFs. The greater competitiveness of the active ETF products suggests that ETFs are becoming more adept at alpha generation. Lower information costs due to enhanced technology may be the culprit here, but whatever the cause the flows speak to an eroding advantage for mutual funds in active management.

A second dimension of enhanced competition is fee pressure. For the ETFs, we calculate fees for active-in-form, active-in-function, and passive categories. Our model, and the results above, predict that increasing competition should result in fees decreasing through time. Figure 11 illustrates these fee dynamics. Panel A shows that average ETF fees (weighted by assets under

[36] The mutual fund holdings data are also sourced in a similar manner as the ETF holdings from the Thomson Reuters Fund holdings database. We exclude non-equity funds, and non-US funds, and funds with fewer than 10 holdings or $10m in assets.

management) exhibit a broad decline across all ETF categories, consistent with increased competition. Hypothesis 7 is that the fee spread between the active and passive products would decline, which is clearly supported by the data—the decline in fees is much larger for the two types of active ETFs compared to broad, passive ETFs.

Panel B of Figure 11 adds mutual funds for comparison, separating the index funds from the rest (active mutual funds). Hypothesis 8 argued that entry should lower fees for both mutual funds and ETFs, and these declining fees are evident in the Figure. To the extent that ETFs also put competitive pressure on mutual funds, we would expect that as the number of ETFs increases, mutual fund fees fall to narrow the spread between ETF fees and mutual fund fees. Indeed, this is what the data show—mutual fund fees for both index and non-index funds fall through time, consistent with the effects of competition. We also see, consistent with Hypothesis 6, that ETF fees are lower than those of mutual funds (active ETF fees are lower than active mutual fund fees and passive ETF fees are lower than index mutual fund fees). As well, we find support for Hypothesis 4 that fees for passive funds are lower than for active funds (passive ETF fees are lower than active ETF fees and index mutual fund fees are lower than active mutual fund fees).

< Figure 11 here >

7. Market Implications of ETF Activeness

Our results show that most ETFs are active investment vehicles, they get used in an active manner by investors who frequently turn over positions in ETFs, they have become more active through time, and they have increased competition across the asset management industry. In this section we turn to some broader implications of ETF activeness, but caution at the outset that our goals here are modest. The evolution of ETF products coincides with a range of other market changes including high-frequency and algorithmic trading, regulatory changes, new quantitative trading strategies, and the like. Ascribing causality for changes in market fundamentals to the effects of ETFs is not our intention. We offer, instead, some observations of what the changing nature of ETFs may imply for some important policy questions.

7.1 Is the Overall Market Becoming More Passive?

The large flows of funds out of active US mutual funds and into index funds and ETFs raise the question of whether investors overall are becoming more passive in how they invest. While the flows within mutual funds, from the more active ones to the more passive ones (Figure 10) contribute to making investment more passive, the flows into the more active categories of ETFs may act as an opposing force. Given the increasing activeness of ETFs over time, is it possible that the activeness level overall is not greatly affected?

Measuring the overall activeness of investors is a daunting task, and it is compounded by the different scales of investment in mutual funds and ETFs. Previous literature, however, provides the valuable insight that stock turnover provides information about the overall activeness or passiveness of investors.[37] More precisely, the cross-sectional standard deviation of individual stock turnover (traded dollar volume divided by market capitalization) provides a measure of investor overall activeness (e.g., Tkac, 1999; Lo and Wang, 2000; Bhattacharya and Galpin, 2011).

The intuition is as follows. If all market participants were completely passive in their investment, we should only see buying and selling of the market portfolio. Trading the market portfolio involves buying or selling all stocks in proportion to their market capitalization, so the dollar trading volume of each individual stock would be proportional to its market capitalization. Therefore, the ratio of dollar trading volume to market capitalization for individual stocks (their turnover ratios) would be equal for all stocks, and equal to the market turnover ratio. If this were the case, the cross-sectional standard deviation of turnover ratios would be zero. In contrast, active investors do not buy and sell all stocks together in market capitalization weights and therefore cause the cross-sectional standard deviation of turnover to be greater than zero. More active investing (or correspondingly, less passive investing) causes a larger cross-sectional standard deviation of turnover.

< Figure 12 here >

To implement this measure of overall market-wide activeness, we compute the turnover of each stock in each quarter and then in each quarter we take the cross-sectional standard deviation

[37] Aggregate activeness cannot simply be measured by aggregating up the positions of investors and measuring the activeness of the aggregate portfolio. This is because collectively all investors hold the market portfolio (the basis of Sharpe's "arithmetic of investing") even though individual investors may be making active investment decisions and departing markedly from the market portfolio.

of the log turnover values.[38] Figure 12 shows the time-series of this measure of overall activeness (solid line, left-hand side axis). The figure shows that, at least as measured by this metric, the overall activeness of investors throughout this 17-year period has not changed markedly. Our results suggest that while investors have changed the *vehicles* through which they invest, they may not have substantially changed their aggregate appetite for seeking outperformance through bets on certain segments of the market, factors, or individual stocks. Investors used to do it via mutual funds; now many do it via active ETFs.

7.2 ETFs and Stock Price Efficiency

Many concerns have been raised about the effects of ETFs on market efficiency. Kacperczyk et al. (2020), for example, argue that the rise of passive investors has a negative effect on price informativeness. They reason that as the size of the passive sector increases, informed investors focus their information gathering on a smaller subset of stocks. Some stocks increase in information content, while others lag behind, but the overall effect is a reduction in price informativeness. A more positive case for price informativeness is given by Buss and Sundaresan (2020) who argue that the inelastic demand characterizing passive investing lowers firms' cost of capital, inducing firms to take on riskier projects. The higher variability of the resulting cash flows encourages more information gathering and so improves price informativeness. Implicit in both analyses is the premise that ETFs largely reflect passive investment that does not seek to exploit mispricing, but our results suggest this characterization may not be accurate. Thus, the overall effect of ETFs on market informativeness remains unclear.

A more targeted area of research considers the simpler issue of whether measures of stock-specific information in prices have changed over time. We turn our attention to three such measures of stock-specific information in prices. The first is a measure based on a return variance decomposition that separates noise from information, and then partitions information into various sources. We obtain these information and noise components from Brogaard et al. (2021), who estimate them for each stock in each year and then aggregate across stocks in each year by taking the average of each variance component. Of primary interest is the component of return variance due to firm-specific information.

[38] We log the turnover values similar to Bhattacharya and Galpin (2011) to account for the tendency for the standard deviation to increase with the level of market turnover.

The second measure is the Bai et al. (2016) stock price informativeness measure, which is the extent to which current individual stock valuations predict individual stock earnings five years in the future.[39] The measure is estimated from annual cross-sectional regressions of future earnings on current market valuations. Because the measure requires earnings five years in the future, we cannot calculate the measure through the end of our sample period.

The third measure is idiosyncratic variation, which is widely used in the literature as a measure of firm-specific information in prices (e.g., Morck et al., 2000). We measure the proportion of idiosyncratic variation in stock returns by estimating a market model regression for each stock in each year (regressing daily stock returns on daily market returns), saving the R^2 from each regression, calculating the average R^2 each year, and finally inverting the measure $(1 - R^2)$ so that it measures firm-specific variation rather than market-wide variation.

Using the time frame of our analysis, Figure 13 plots the time-series of all three of these measures of firm-specific information (solid lines, left-hand side axes) with the first two measures in Panel A and the third in Panel B. There is little evidence of a deterioration of firm-specific information in prices. While firm-specific information as a proportion of stock return variance tends to decrease during the crisis years when stocks tend to fall and recover together, firm-specific information is slightly higher at the end of the sample than at the beginning for the first measure and about the same at the beginning and end for the third measure. Although the second measure, the Bai et al. (2016) stock price informativeness measure, cannot be calculated through to the last years in the sample, it shows a very similar trend to the Brogaard et al. (2021) stock-specific information share and like the other measures shows no signs of deterioration during a period of strong growth in ETF assets.

< Figure 13 here >

These time-series patterns do not prove one way or the other if ETFs have harmed the amount of firm-specific information in prices, but they do show that price informativeness in general has remained relatively unchanged. The patterns are consistent with our earlier findings that collectively ETFs, in particular the active or "aggressive-passive" ones, might contribute to price discovery. By facilitating efficient and low-cost bets on segments of the market, industries, factors, and various other tilts, ETFs can be attractive vehicles for active investors. The strong

[39] We thank the authors for providing the data for their measure.

growth in the number of highly active ETFs through time and their success in attracting fund flows suggest that this is indeed the case.

8. Conclusion

This paper investigated the active world of passive investing. Our results demonstrate that many ETFs are actually active investment products in both form and function. We find that most ETFs are highly active with a median Activeness Index of 93.1% and median Active Return Deviation of 8.8% p.a. Moreover, fund flows over time are concentrating into the most active ETFs. The highly active ETFs also account for the vast majority of ETF trading activity, consistent with such ETFs being used by investors in an active manner. Consequently, overall activeness of investment in the market, cutting across the different investment vehicles and how they are used, has not changed much over the past two decades.

What emerges from our research is the clear finding that ETFs are best viewed as a continuum of products across the passive-active space. Much as mutual funds rely on manager ability to pick stocks, the new generation of ETFs relies on skill in designing indices and trading algorithms to generate returns.

This transition of ETFs from purely passive large-scale index products to more "aggressive-passive" investment vehicles has important implications. For many investors these new generation ETFs allow for implementation of more complex factor-based and industry-focused strategies. These products also provide new ways of hedging, with shorting of ETFs effectively sidestepping traditional constraints. Such innovations can potentially enhance the risk-return trade-off for investors. But, as our findings demonstrate, as ETFs become more narrowly focused, the same forces affecting mutual funds come into play for ETFs: alpha can degrade, scale becomes a negative, and ETFs may struggle to beat benchmarks.

These latter trends suggest a new complexity to the competitive landscape in investment management. That ETFs have successfully challenged the business model of the active mutual fund industry is clear. The ability of ETFs to provide both a lower cost of passive investing and new vehicles for alpha-generation introduced a new competitive environment, with lower fees and many more funds competing for investor's funds. Mutual funds have struggled to compete, but now that same fate may characterize intra-ETF competition as well. In 2020, more ETFs were liquidated than launched, suggesting that older ETF products may also be at risk of obsolescence.[40] This industry shake-up may become even more pronounced when the new "active" ETFs featuring opacity of holdings become live. Regardless of their success, a new era of investment management competition is at hand.

Finally, another important implication of our research is that ETFs are not simply "free-riders" in the market—investors are still betting on the outperformance of segments of the market, factors, or individual stocks, albeit now these bets are placed via active ETFs and index products. But precisely because many ETFs are active investment vehicles, their impact on the market more generally is an important area for future research. Issues such as whether ETFs enhance market stability, or how ETFs impact liquidity, or whether particular types of ETFs (such as leveraged products or exchange-traded notes) can be harmful seem fruitful areas for inquiry.

[40] See "The Shakeout in the ETF industry is Accelerating", Wall Street Journal, Aug. 25, 2020.

References

Amihud, Y. and Goyenko, R. (2013) Mutual fund's R^2 as predictor of performance, *Review of Financial Studies* 26, 667–694.

Bai, J., Philippon, T., and Savov, A. (2016) Have financial markets become more informative?, *Journal of Financial Economics* 122, 625–654.

Baltussen, G., van Bekkum, S., and Da, Z. (2019) Indexing and stock market serial dependence around the world, *Journal of Financial Economics* 132, 26–48.

Ben-David, I., Franzoni, F., and Moussawi, R. (2018) Do ETFs increase volatility?, *Journal of Finance* 73, 2471–2535.

Berk, J. and Green, R. (2004) Mutual fund flows and performance in rational markets, J*ournal of Political Economy* 112, 1269–1295.

Berk, J. and van Binsbergen, J. (2015) Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

Bhattacharya, U. and Galpin, B. (2011) The rise of value-weighted portfolios, *Journal of Financial and Quantitative Analysis* 46, 737–756.

Bhattacharya, U., Loos, B., Meyer, S., and Hackethal, A. (2017) Abusing ETFs, *Review of Finance* 21, 1217–1250.

Bhojraj, S., Mohanram, P.S., and Zhang, S. (2020) ETFs and information transfer across firms, *Journal of Accounting and Economics* 70, 1–20.

Blocher, J. and Whaley, R. (2016) Two-sided markets in asset management: Exchange-traded funds and securities lending, Working paper.

Bond, P. and Garcia, D. (2018) The equilibrium consequences of indexing, Working paper.

Brogaard, J., Nguyen, T.H., Putnins, T.J., and Wu, E. (2021) What moves stock prices? The role of news, noise, and information, Working paper.

Brogaard, J., Ringgenberg, M.C., and Sovich, D. (2019) The economic impact of index investing, *Review of Financial Studies* 32, 3461–3499.

Buss, A. and Sundaresan, S. (2020) More risk, more information: How passive ownership can improve informational efficiency, Working paper.

Cao, B., Xue J., and Wermers, R. (2017) Mutual Fund Links update procedure, Working paper.

Chen, H.L., Jegadeesh, N., and Wermers, R. (2000) The value of active mutual fund management: An examination of the stockholdings and trades of fund managers, *Journal of Financial and Quantitative Analysis* 35, 343–368.

Cheng, S., Massa, M., and Zhang, H. (2019) The unexpected activeness of passive investors: A worldwide analysis of ETFs, *Review of Asset Pricing Studies* 9, 296–355.

Coles, J., Heath, D., and Ringgenberg, M.C. (2020) On index investing, Working paper.

Cong, L.W. and Xu, D.X. (2016) Rise of factor investing: Asset prices, informational efficiency, and security design, Working paper.

Cremers, M.. Ferreira, J., Matos, P., and Starks, L. (2016) Indexing and active fund management: International evidence, *Journal of Financial Economics* 120, 539–560.

Cremers, M. and Petajisto, A. (2009) How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Fraser-Jenkins, I. (2016) The silent road to serfdom: Why passive investing is worse than Marxism, Sanford-Bernstein Research Report (August 23, 2016).

Garleanu, N. and Pedersen, L.H. (2019) Active and passive investing, Working Paper.

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

Grossman, S.J. and Stiglitz, J.E. (1980) On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hill, J.M. (2016) The evolution and success of index strategies in ETFs, *Financial Analysts Journal* 72, 8–13.

Huang, J., Wei, K.D., and Yan, H. (2007) Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance* 62, 1273–1311.

Huang, S., O'Hara, M., and Zhong, Z. (2021) Innovation and informed trading: Evidence from industry ETFs, *Review of Financial Studies* 34, 1280–1316.

Israeli, D., Lee, C., and Sridharan, S.A. (2017) Is there a dark side to exchange traded funds? An information perspective, *Review of Accounting Studies* 22, 1048–1083.

Ivkovic, Z., Sialm, C., and Weisbenner, S. (2008) Portfolio concentration and the performance of individual investors, *Journal of Financial and Quantitative Analysis* 43, 613–655.

Kacperczyk, M., Nosal, J., and Sundaresan, S. (2020) Market power and price informativeness, Working paper.

Kacperczyk, M., Sialm, C., and Zheng, L. (2005) On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.

Kacperczyk, M., Sundaresan, S., and Wang, T. (2021) Do foreign institutional investors improve price efficiency? *Review of Financial Studies* 34, 1317–1367.

Lettau, M. and Madhavan, A. (2018) Exchange-traded funds 101 for economists, *Journal of Economic Perspectives* 32, 135–154.

Levy, A. and Lieberman, O. (2016) Active flows and passive returns, *Review of Finance* 20, 373–401.

Lo, A.W. and Wang, J. (2000) Trading volume: Definitions, data analysis, and implications for portfolio theory, *Review of Financial Studies* 13, 257–300.

Mackintosh, P. (2017) It's all about active ETFs, *Journal of Index Investing* 7, 6–15.

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

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

Malikov, G. (2019) Information, participation, and passive investing, Working paper.

Morck, R., Yeung, B., and Yu, W. (2000) The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58, 215–260.

Pastor, L.R., Stambaugh, R.F., and Taylor, L.A. (2020) Fund tradeoffs, *Journal of Financial Economics* 138, 614–634.

Pedersen, L.H. (2018) Sharpening the arithmetic of active management, *Financial Analysts Journal* 74, 21–36.

Sharpe, W.F. (1991) The arithmetic of active management, *Financial Analysts Journal* 47, 7–9.

Sirri, E. and Tufano, P. (1998) Costly search and mutual funds flows, *Journal of Finance* 43, 1589–1622.

Stambaugh, R.F. (2014) Presidential address: Investment noise and trends, *Journal of Finance* 69, 1415–1453.

Tkac, P.A. (1999) A trading volume benchmark, *Journal of Financial and Quantitative Analysis* 89, 89–114.

Table I. Activeness in the cross-section of ETFs

This table shows the distribution of two measures of activeness (the Activeness Index and the Active Return Deviation) in the cross-section of ETFs. We compute the measures for each ETF in each quarter of our sample (2000-2017) and then take the average of each of the measures for each ETF to get a cross-sectional distribution of ETF activeness. The table reports the mean, median, and first and third quartiles (Q1 and Q3) of this distribution. Panel A is the pooled sample including all of the ETFs. Panel B partitions ETFs into four groups according to when the ETF commenced (we use the date the ETF first appears in the dataset). Panel C partitions the ETFs into quartiles by size (AUM). Panel D partitions the ETFs into quartiles by secondary market turnover (traded dollar volume divided by market capitalization).

	Activeness Index (%)				Active Return Deviation (% p.a.)			
	Mean	Q1	Median	Q3	Mean	Q1	Median	Q3
Panel A: Pooled sample (all ETFs)								
	85.9	86.1	93.1	97.5	9.2	5.4	8.8	11.7
Panel B: ETF size quartiles (AUM)								
Q1 = small	89.5	88.1	95.5	98.2	10.7	6.3	10.5	14.1
Q2	89.1	91.5	96.1	98.1	9.4	6.0	8.8	11.5
Q3	86.7	81.8	91.4	97.3	8.8	5.9	8.7	11.5
Q4 = big	78.3	63.2	89.6	94.4	7.8	3.9	6.6	10.2
Panel C: ETF secondary market turnover quartiles								
Q1 = low	82.6	75.3	89.6	96.6	7.0	4.3	6.1	9.4
Q2	84.6	83.0	93.9	97.5	8.1	5.2	8.0	10.3
Q3	86.1	87.3	92.9	97.9	9.9	6.6	9.5	12.4
Q4 = high	90.2	89.7	94.3	97.6	11.7	8.7	11.5	15.0
Panel D: Chronological groups of ETF start								
Before 2001	78.0	60.5	89.7	95.5	7.8	4.3	8.1	11.0
2001 – 2005	85.5	83.9	91.8	97.5	8.1	4.8	8.1	10.5
2006 – 2011	87.5	87.3	93.6	97.7	9.7	5.9	9.2	12.3
2012 – 2017	97.1	93.2	99.1	99.8	13.1	10.5	12.3	14.6

Table II. How ETF characteristics relate to activeness

This table reports cross-sectional OLS regression results of how ETF characteristics relate to ETF activeness. The dependent variable is the Activeness Index (a measure of how much the holdings of the ETF depart from the market portfolio) logit transformed. $Size_i$ is the log dollar value of the ETF's equity holdings. Fee_i is the annual management fee charged by the ETF (percent per annum). $Turnover_i$ is the log turnover ratio of the ETF in the secondary market (traded dollar volume divided by market capitalization). $FlowPerfSensitivity_i$ is an estimate of the change in percentage flows to the ETF per unit increase in performance (percentile in the cross-sectional return distribution). Standard errors are double clustered by ETF issuer (management company) and activeness category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	Model 1 $Activeness$ $Index\ (logit)$	Model 2 $Activeness$ $Index\ (logit)$	Model 3 $Activeness$ $Index\ (logit)$	Model 4 $Activeness$ $Index\ (logit)$	Model 5 $Activeness$ $Index\ (logit)$	Model 6 $Activeness$ $Index\ (logit)$
$Intercept$	6.09	1.89	5.81	3.03	2.67	6.36
	(9.58)***	(3.03)***	(9.91)***	(8.67)***	(5.70)***	(5.00)***
$Size_i$	–0.57		–0.62			–0.65
	(–4.91)***		(–6.84)***			(–14.36)***
Fee_i		4.22	2.31			–0.21
		(3.14)***	(3.42)***			(–0.14)
$Turnover_i$				0.71		0.69
				(2.71)***		(12.92)***
$FlowPerfSensitivity_i$					5.98	4.69
					(2.77)***	(2.02)**
R^2	32%	9%	42%	6%	4%	46%
N	452	393	393	402	412	338

Table III. Flow-performance sensitivity of ETFs

This table reports OLS regression results measuring the flow-performance sensitivity of ETFs. The dependent variable, $Flow_{i.t}$, is the monthly percentage flow of funds into ETF i, measured from the monthly percentage change in the ETF's number of shares outstanding. The ETF performance in the previous month, $Perf_{i,t-1}$, is measured as the percentile of the ETF's return (relative to other ETFs) in month $t-1$. D_i^{Active} is a dummy variable for active ETFs (those with an Activeness Index > 50%). $D_i^{ActiveInForm}$ is a dummy variable for the subset of active ETFs that are classified as active-in-form (those in the top quartile of active ETFs by portfolio turnover). Models 2 and 3 include fund and time (month) fixed effects, respectively. Model 5 restricts the sample to only active ETFs. Standard errors are double clustered by ETF and date. t-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	Model 1 $Flow_{i.t}$	Model 2 $Flow_{i.t}$	Model 3 $Flow_{i.t}$	Model 4 $Flow_{i.t}$	Model 5 $Flow_{i.t}$	Model 6 $Flow_{i.t}$
$Intercept$	0.15 (0.49)	0.77 (0.36)	–0.04 (–0.06)	2.71 (2.40)***	0.64 (1.85)*	2.71 (2.40)***
$Perf_{i,t-1}$	0.05 (10.40)***	0.05 (10.36)***	0.05 (10.45)***	0.00 (0.12)	0.04 (8.17)***	0.00 (0.12)
$Perf_{i,t-1} \times D_i^{Active}$				0.05 (2.23)**		0.04 (1.83)*
D_i^{Active}				–2.64 (–2.39)**		–2.07 (–1.89)*
$Perf_{i,t-1} \times D_i^{ActiveInForm}$					0.05 (4.74)***	0.05 (4.74)***
$D_i^{ActiveInForm}$					–3.13 (–5.58)***	–3.13 (–5.58)***
Fund fixed effects		Yes				
Time fixed effects			Yes			
Sample	Pooled	Pooled	Pooled	Pooled	Active	Pooled
N	52,678	52,678	52,678	52,678	48,970	52,678

Table IV. How active-in-form and active-in-function ETFs differ in characteristics.

This table reports cross-sectional OLS regression results for the subsample of active ETFs (those with an Activeness Index > 50%). The dependent variables are various characteristics. $Size_i$ is the log dollar value of the ETF's equity holdings. Fee_i is the annual management fee charged by the ETF (percent per annum). $Turnover_i$ is the log turnover ratio of the ETF in the secondary market (traded dollar volume divided by market capitalization). Portfolio turnover ($PortfTurn_i$) is the percentage of the fund's average net asset value that is replaced per annum, excluding portfolio changes that arise from inflows and outflows. $FlowPerfSensitivity_i$ is an estimate of the change in percentage flows to the ETF per unit increase in performance (percentile in the cross-sectional return distribution). The key independent variable ($D_i^{ActiveByForm}$) is an indicator variable for whether the ETF is in the active-in-form group of ETFs identified through cluster analysis. In Panel B we control for ETF activeness (logit transformation of the Activeness Index). Standard errors are double clustered by ETF issuer (management company) and activeness category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	Model1 $Size_i$	Model2 Fee_i	Model3 $Turnover_i$	Model4 $PortfTurn_i$	Model5 $FlowPerf$ $Sensitivity_i$
Panel A					
$Intercept$	5.09	0.32	0.59	0.16	0.01
	(13.98)***	(6.65)***	(4.15)***	(9.88)***	(1.66)*
$D_i^{ActiveByForm}$	–0.38	0.07	–0.21	0.55	0.10
	(–2.04)**	(2.05)**	(–1.61)	(12.62)***	(29.25)***
R^2	1%	3%	1%	27%	40%
N	406	350	358	406	383
Panel B					
$Intercept$	7.05	0.25	0.28	0.21	–0.01
	(9.62)***	(3.55)***	(1.75)*	(4.99)***	(–0.67)
$D_i^{ActiveByForm}$	–0.41	0.07	–0.20	0.55	0.10
	(–1.52)	(2.88)***	(–2.31)**	(16.29)***	(70.80)***
$Activeness_i$	–0.57	0.02	0.09	–0.02	0.00
	(–5.81)***	(1.80)*	(4.58)***	(–1.75)*	(2.26)**
R^2	26%	8%	6%	27%	41%
N	406	350	358	406	383

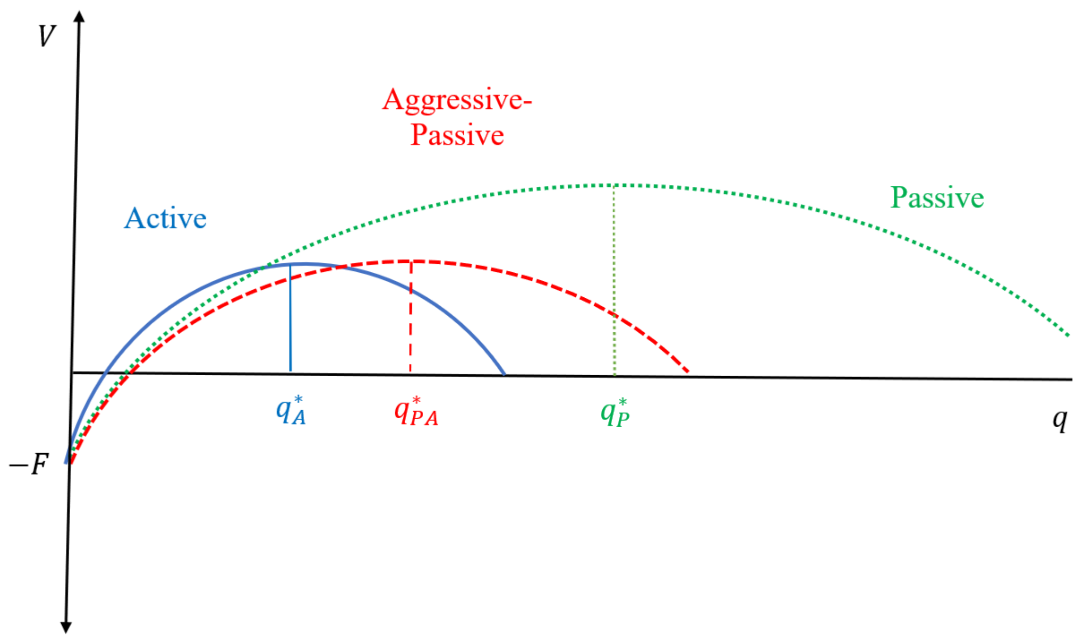


Figure 1. Active, passive, and aggressive-passive funds in equilibrium.
This figure provides the graph of the value of a fund ($V = qf$) to the manager for an active fund (A, blue solid line), a representative aggressive-passive fund (PA, red dashed line), and a passive fund (P, green dotted line). The starred quantities are the optimal sizes of the funds as chosen by the managers. The figure is drawn for active and passive parameters satisfying inequality (10) and aggressive-passive parameters satisfying parameters: $a_P < a_{PA} \leq a_A$ and $b_P < b_{PA} \leq b_A$.

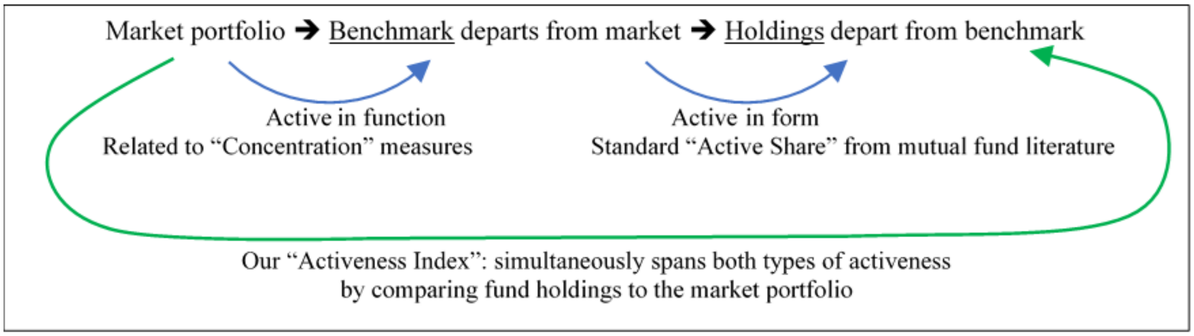
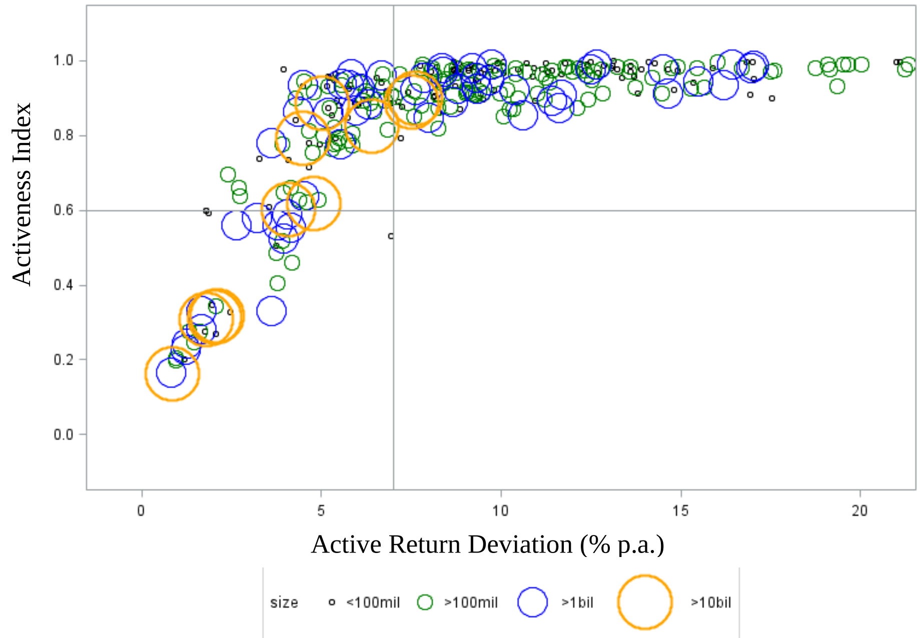


Figure 2. Schematic of active in function and active in form.
This figure illustrates how ETF activeness can arise from choosing a benchmark that embraces or departs from the market (active in function) or by choosing holdings that depart from the level in the chosen benchmark (active in form). Our notion of activeness is the extent to which the ETF deviates from the completely passive strategy of holding every component of the market in value-weighted measures.

Panel A: By AUM group



Panel B: By turnover group

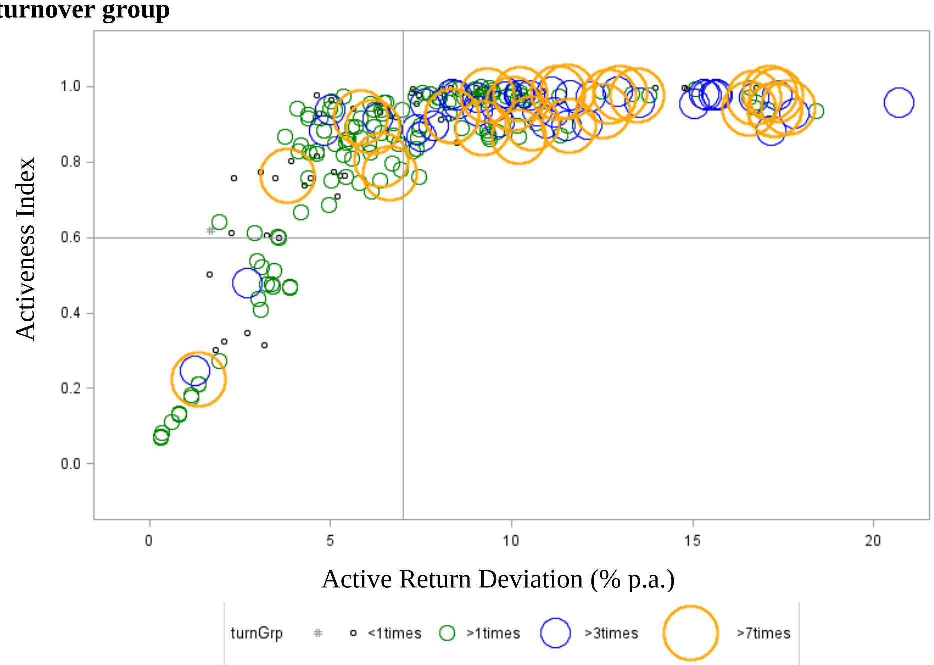


Figure 3. Activeness Index and Active Return Deviation of ETFs, by size and turnover.
The figure plots each of the ETFs in the two dimensions of activeness: the horizontal axis is the fund's average Active Return Deviation and the vertical axis is the fund's average Activeness Index. Each circle is an ETF. In Panel A, bigger circles are used for bigger ETFs in terms of assets under management (biggest circles are ETFs with more than $10bil of holdings on average, followed by $1bil to $10bil, followed by $100mil to $1bil, followed by under $100mil). In Panel B, bigger circles are used for ETFs that have higher turnover in the secondary market (biggest circles are ETFs with annual turnover more than seven times market capitalization, followed by turnover of three to seven times, followed by one to three times, followed by under one.

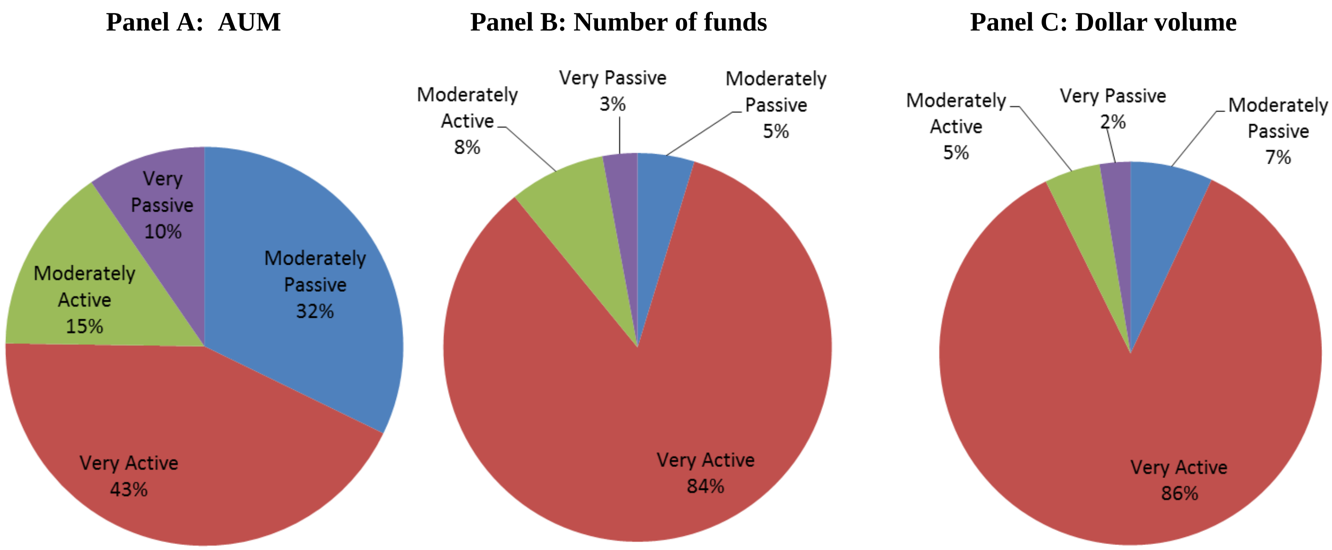
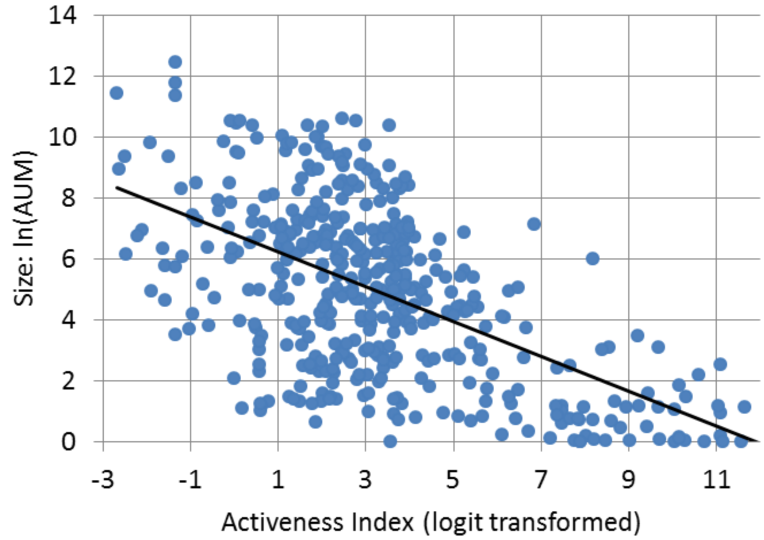


Figure 4. Breakdown of ETF AUM, turnover, and fund count by levels of activeness.
This set of pie charts shows the relative size of four ETF activeness categories, with the size of each category measured by the assets under management (AUM) in Panel A, number of ETFs in Panel B, and ETF secondary market traded dollar volume in Panel C. We assign each ETF to one of four baskets based on its average Activeness Index: (i) "Very Passive" (Activeness Index < 25%), (ii) "Moderately Passive" (25% < Activeness Index < 50%), (ii) "Moderately Active" (50% < Activeness Index < 75%), and (ii) "Very Active" (Activeness Index > 75%). These charts reflect the state of ETF activeness at the end of the sample period (last quarter of 2017).

Panel A: ETF Activeness and fund size



Panel B: ETF Activeness and management fees

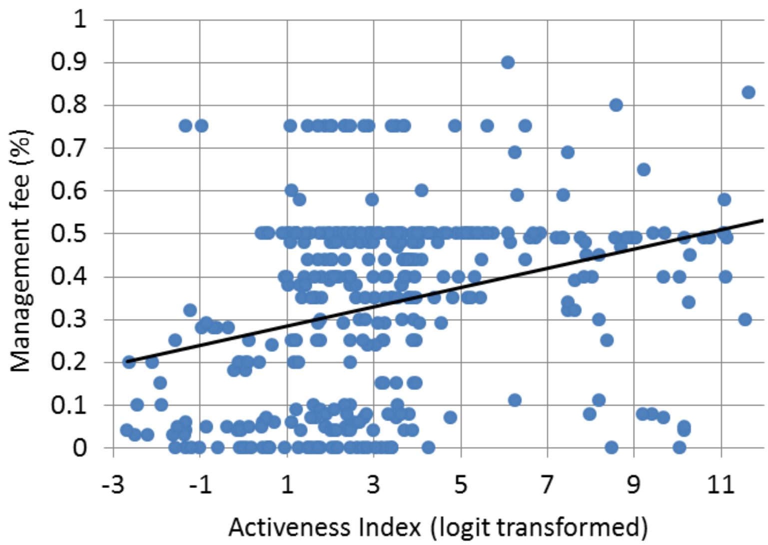


Figure 5. Relation between ETF activeness, size, and fees.
This figure plots each ETF as a dot. The horizontal axis measures the ETF's activeness (logit transformation of the Activeness Index). The vertical axis measures the ETF size (log of the dollar value of the ETF's equity holdings) in Panel A and the ETF's annual management fee (percent per annum) in Panel B.

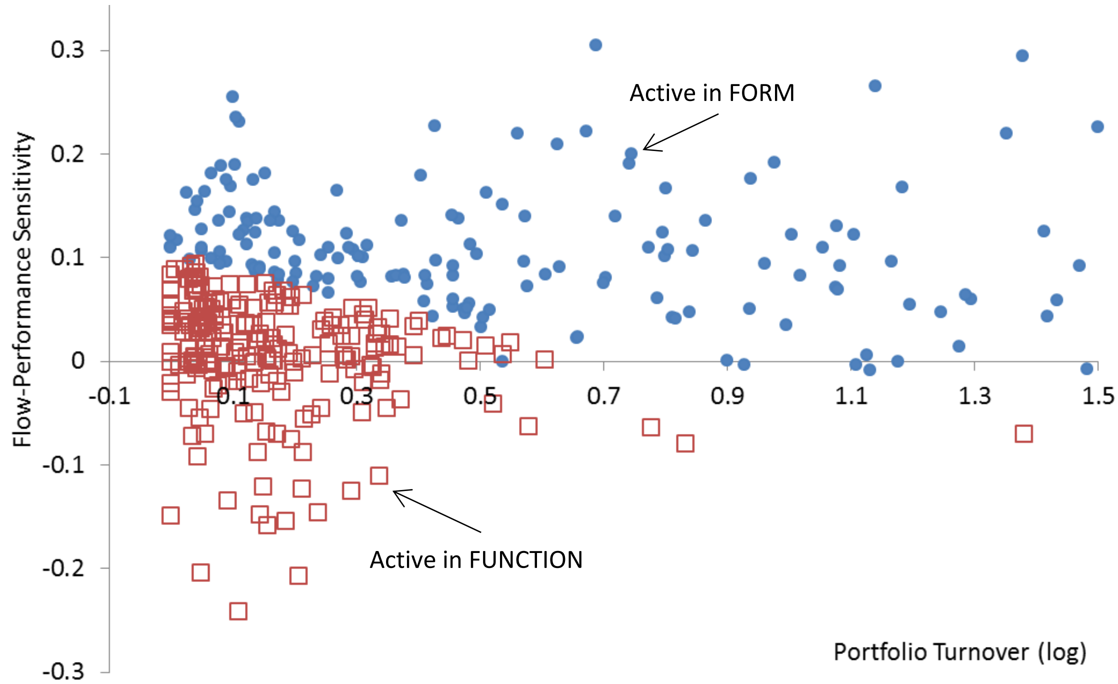


Figure 6. Active-in-form vs active-in-function ETF.
This figure plots each of the active ETFs: dots are used to represent ETFs that are active in form (active management within the ETF) and squares are used to represent ETFs that are active in function (used by active investors as building blocks). The horizontal axis measures the log of the turnover within the fund (the rate at which the fund's portfolio is turned over). The vertical axis measures the sensitivity of investor fund flows to past fund performance with positive values indicating that inflows to the particular fund tend to follow good past performance.

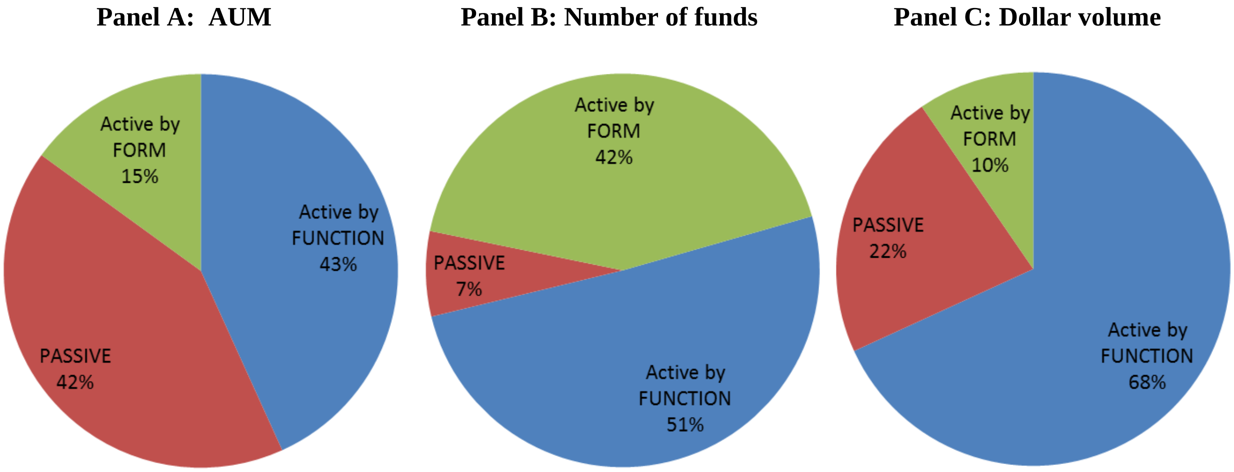


Figure 7. Breakdown of ETF AUM, turnover, and fund count by types of activeness and passiveness. This set of pie charts shows a comparison of ETFs that are active in form (active management within the ETF), ETFs that are active in function (used by active investors as building blocks), and passive ETFs. These three categories are compared with respect to their aggregate assets under management (AUM, Panel A), the number of ETFs in each category (Panel B), and ETF secondary market traded dollar volume (Panel C). Passive ETFs are those with an average Activeness Index < 50%. The active ETFs (Activeness Index > 50%) are partitioned into "active in form" and "active in function" using cluster analysis based on the ETF portfolio turnover and flow-performance sensitivity. These charts reflect the state of ETF activeness at the end of the sample period (last quarter of 2017).

Figure 8. Evolution of ETF activeness through time.

This figure provides four snapshots of ETF activeness at different points in time (2001, 2006, 2011, and 2016). The horizontal axis is the ETF's Active Return Deviation and the vertical axis is the ETF's Activeness Index. Each circle is an ETF. Bigger circles are used for bigger ETFs in terms of assets under management (biggest circles are ETFs with more than $10bil of holdings on average, followed by $1bil to $10bil, followed by $100mil to $1bil, followed by under $100mil).

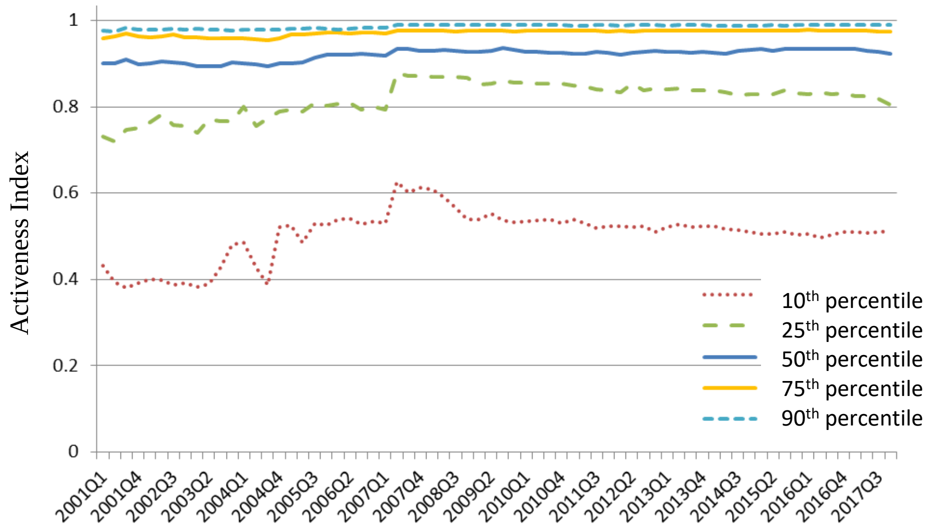
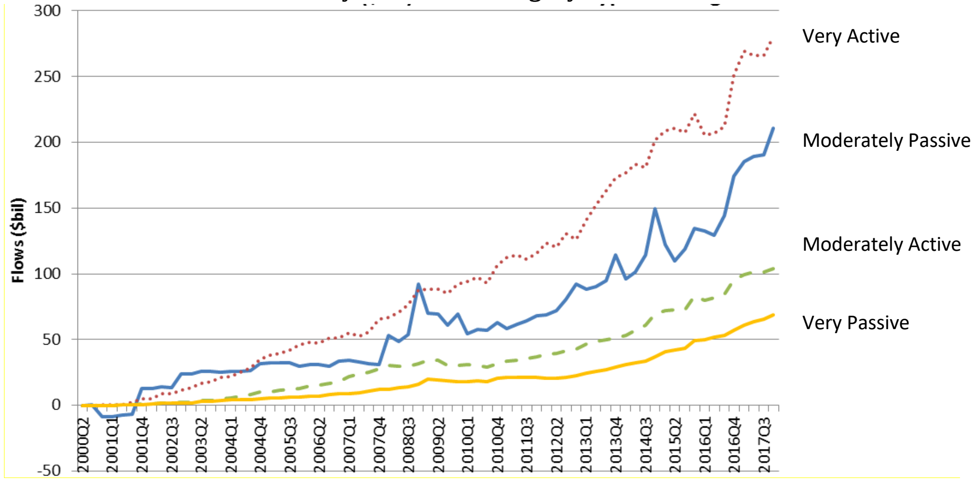


Figure 9. Cross-sectional distribution of ETF Activeness Index through time.
This figure shows the distribution (10th, 25th, 50th, 75th, 90th percentiles) of Activeness Index in the cross-section of ETFs through time.

Panel A: Cumulative fund flows to ETFs by activeness category



Panel B: Cumulative fund flows to mutual funds by activeness category

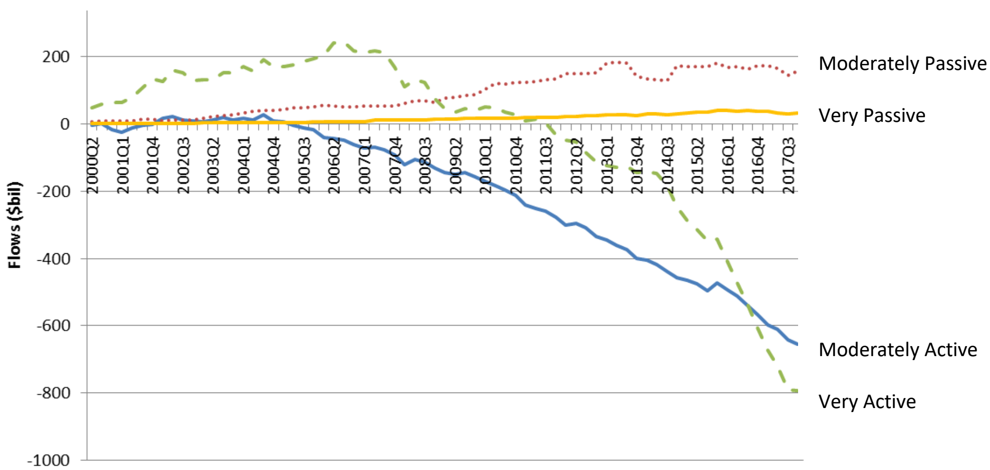
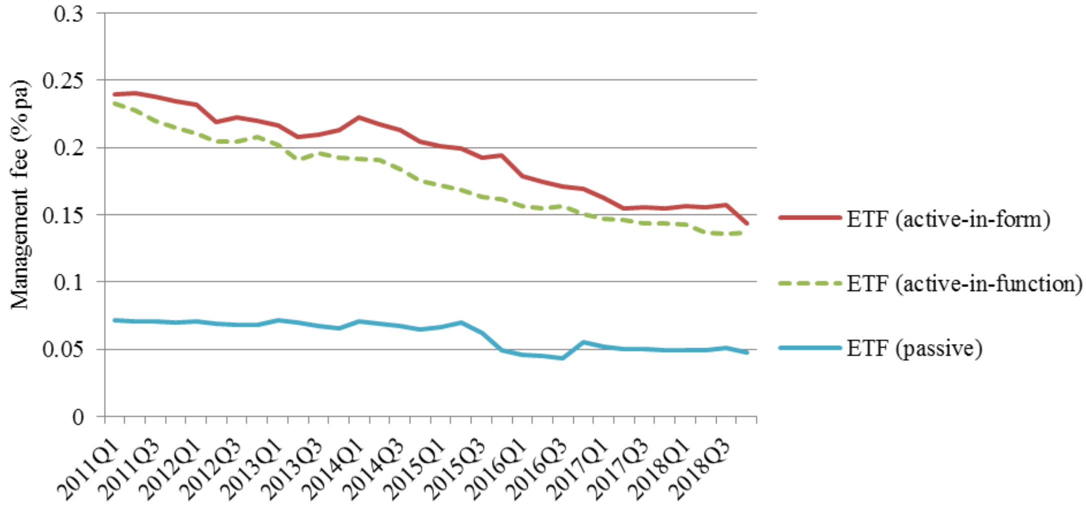


Figure 10. Fund flows to ETFs and mutual funds by activeness category.

This figure shows cumulative fund flows to ETFs (Panel A) and US mutual funds (Panel B) broken down into four activeness categories based on the fund's average Activeness Index: (i) "Very Passive" (Activeness Index < 25%), (ii) "Moderately Passive" (25% < Activeness Index < 50%), (ii) "Moderately Active" (50% < Activeness Index < 75%), and (ii) "Very Active" (Activeness Index > 75%).

Panel A: ETF fees through time



Panel B: Comparison of ETF and mutual fund fees through time

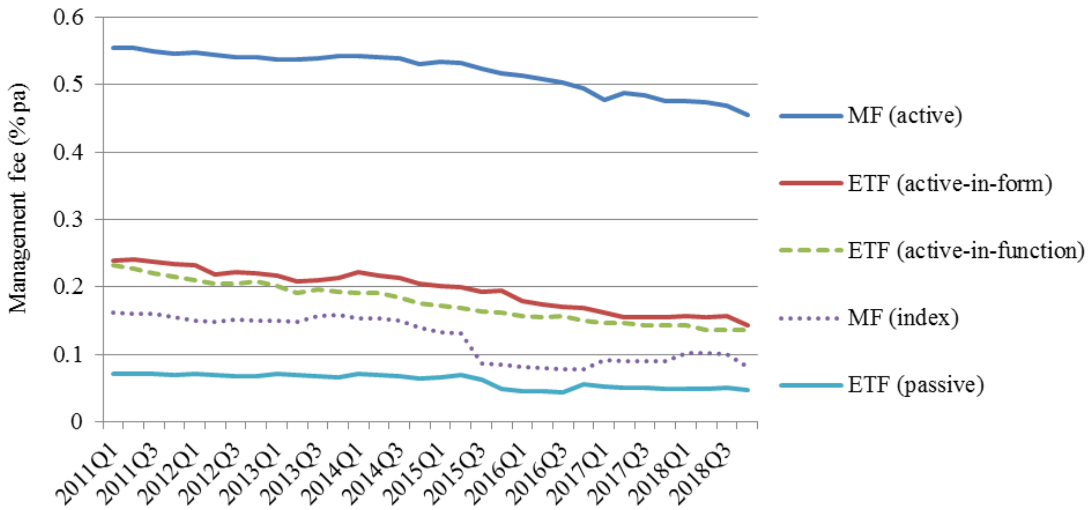


Figure 11: ETF and mutual fund fees by category through time.
This figure shows the average management fee (% per annum) for ETFs in Panel A segmented by active-in-form, active-in-function, and passive ETF categories. Panel B adds mutual funds for comparison, segmented into active mutual funds (non-index funds) and index mutual funds. The averages are weighted by the assets under management.

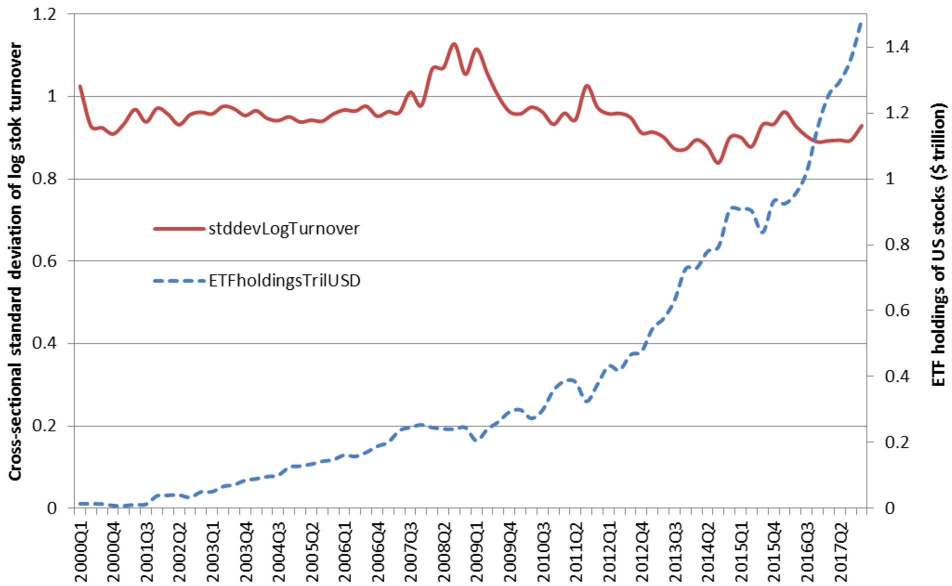
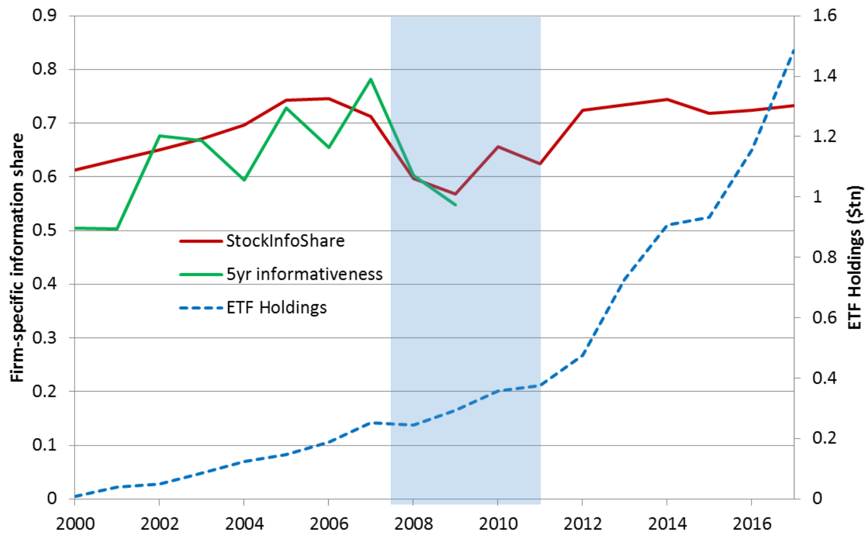


Figure 12. ETF holdings and overall activeness of stock market investments.
The solid line is a time-series measure of the overall activeness of investors in the stock market: the cross-sectional standard deviation of the turnover of individual stocks (left-hand axis). To compute this measure, each quarter, we measure each stock's log turnover (traded dollar volume divided by market capitalization) and then the standard deviation of the log turnover values. The dashed line (right-hand side axis) shows the time-series of the aggregate value of ETF holdings of US stocks. This measure is calculated each quarter by taking each ETF's holdings, valuing the holdings at market closing prices, and then summing the value of the holdings across all ETFs in our sample.

Panel A: Firm-specific information share and price informativeness



Panel B: Stock return idiosyncratic variance share

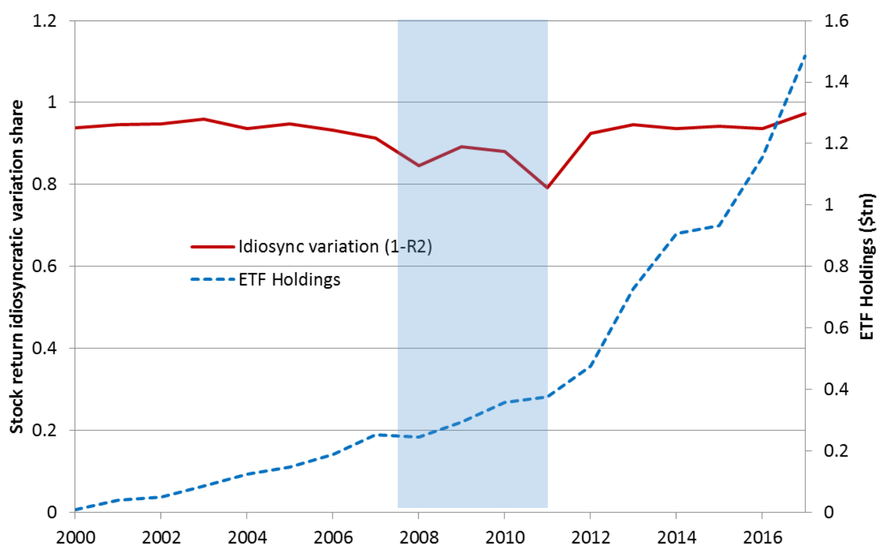


Figure 13. ETF holdings and firm-specific information in stock prices.
This figure shows time-series trends in the amount of firm-specific information in prices (solid lines, left-hand side axes), using three different measures. The first measure (Panel A, "StockInfoShare") is the share of stock return variance that is attributed to firm-specific information using a temporary-permanent component variance decomposition following Brogaard et al. (2021). The second measure (Panel A, "5yr informativeness") is the extent to which individual stock valuations predict individual stock earnings five years in the future, using annual cross-sectional regressions following Bai et al. (2016). The third measure (Panel B) is the idiosyncratic variation in stock returns as a share of total stock return variance. It is calculated as one minus the average R^2 from regressions of daily stock returns on market returns for each stock in each quarter, similar to Morck et al. (2000). The dashed lines (right-hand side axis) show the time-series of the aggregate value of ETF holdings of US stocks. This measure is calculated each quarter by taking each ETF's holdings, valuing the holdings at market closing prices, and then summing the value of the holdings across all ETFs in our sample.