Dominated ETFs*

David C. Brown† Scott Cederburg‡ Mitch Towner§

March 24, 2021

ABSTRACT

We study dominated products in the market for U.S. equity exchange-traded funds (ETFs). We identify a large number of dominated ETFs with returns that are highly correlated with those of cheaper, more liquid competitors. Counterintuitively, these dominated ETFs attract excess capital relative to expectations based on fund characteristics related to fees, liquidity, performance, strategy uniqueness, and investor awareness. We estimate the aggregate cost to investors from allocating capital to dominated ETFs to be \$1.0 billion to \$6.7 billion from 2000 to 2018. These costs are growing over time as newly listed ETFs claim unique strategies despite high correlations with cheaper ETFs.

Keywords: Exchange Traded Funds (ETFs), Dominated Products, Smart Beta
JEL Classification Numbers: D53, G11, G12, G23

*We thank Carole Comerton-Forde, Zhi Da, Caitlin Dannhauser, Shaun Davies, Travis Johnson, Michael O'Doherty, Bradley Paye, Matt Ringgenberg, Sophie Shive, Rick Sias, Yuri Tserlukevich, and seminar participants at the 2020 Arizona/ASU Junior Conference, Southern Methodist University, the University of Iowa, the University of New South Wales, the University of Virginia - Darden, and Virginia Tech for helpful comments and suggestions. Any errors are our own.

†Eller College of Management, University of Arizona, McClelland Hall, Room 315D, 1130 E. Helen Street, P.O. Box 210108, Tucson, AZ 85721-0108, Phone: (520)621-0746, Fax: (520)621-4261, Email: dcbrown@arizona.edu

‡Eller College of Management, University of Arizona, Email: cederburg@arizona.edu

§Eller College of Management, University of Arizona, Email: mitchtowner@arizona.edu

1 Introduction

Financial markets seem riddled with dominated products. These products attract substantial market share despite the existence of nearly identical, cheaper products. Recent studies identify this phenomenon in markets for mutual funds, bonds, and mortgage loans, among others.[1] Blame for the billions of dollars in losses to investors and customers is often cast on investor irrationality, search costs, or financial advisor incentives. Complexities in market structure, however, often make difficult the identification and measurement of dominated product costs. For example, most mutual funds offer multiple share classes with different fees, and access to share classes varies across investors. As such, an investment in an apparently dominated share class may reflect an inability to access better options rather than a poor decision.[2]

In this paper, we examine dominated products in the market for exchange-traded funds (ETFs). Relative to open-end mutual funds, the ETF market has two primary features that aid the identification of dominated funds. First, investors can access any listed ETF, such that we can directly observe their investment opportunity set. In contrast, access to mutual fund share classes varies, such that considering each share class as a separate option or treating a mutual fund as an asset-weighted combination of its share classes does not reflect the investment opportunity set for a given investor. Second, ETFs have simple fee structures. Expenses are charged as a stated percentage of the net asset value (NAV), and ETFs have no additional fees that provide brokers with incentives to promote inferior investments. Many mutual funds charge 12b-1 fees and front-end or back-end loads, which can distort financial advisor incentives and lead to differences in switching costs across funds.[3] Given these market features, investments in dominated ETFs are not attributable to a

[1]The prior literature documents dominated products in the markets for mutual funds (Elton, Gruber, & Busse, 2004; Hortaçsu & Syverson, 2004; Cooper, Halling, & Yang, 2020), money market funds (Christofferson & Musto, 2002), bonds (Green, Hollifield, & Schürhoff, 2007; Egan, 2019), insurance (Brown & Goolsbee, 2002; Bhargava, Loewenstein, & Sydnor, 2017), and mortgages (Allen, Clark, & Houde, 2014, 2019; Gurun, Matvos, & Seru, 2016).

[2]Access to the share class with the lowest expense ratio often requires a large minimum investment, such as the $5 billion minimum for the Vanguard Total Stock Market Index Fund Institutional Select Shares (VSTSX). Share class access within collective investment structures like 401(k) retirement plans can also depend on the total investment in a fund across all participants, such that an individual investor's access to share classes depends on the investment choices of other plan participants.

[3]State Street's Select Sector SPDR series of ETFs charges 12b-1 fees, making it the exception in the ETF market. State Street uses the fees for marketing expenses but does not provide incentives to financial advisors. According to Dan Dolan, director of management strategies at Select Sector SPDRs, "There are no broker dealers. And no one is getting paid," (http://https://www.bloomberg.com/news/articles/2017-02-08/where-do-spdr-fees-go-check-

lack of investor access to better alternatives or to incentive misalignment between investors and financial advisors.

Our analysis considers U.S. equity ETFs from January 2000 through June 2018. We identify dominated ETFs among the set of ETFs that deliver returns that are highly correlated with the returns of competing ETFs. Return correlations are calculated using daily returns over the trailing 12 months, and we use correlation thresholds of 95.0%, 97.5%, and 99.0%. Intra-day liquidity is an important feature of the ETF market structure, and investor heterogeneity produces a tradeoff between fees and liquidity (Khomyn, Putniņš, & Zoican, 2020). We classify an ETF as dominated if it both charges higher fees and offers lower liquidity compared with a highly correlated competitor.

We find that a large number of dominated ETFs collectively attract substantial assets. On average during the sample period, 38% of ETFs in a given quarter are classified as dominated by a competing fund at the 95.0% correlation threshold. Dominated ETFs collectively manage 36% of the assets under management (AUM) across all U.S. equity ETFs. In aggregate, our estimates of the additional costs of dominated ETFs from higher fees and additional trading costs range from $1.0 billion to $6.7 billion during the sample period depending on the correlation threshold. These costs have steadily increased over time, and we estimate the annual costs as of the end of our sample to be $146 million to $847 million.

Institutional investors could mitigate the costs of dominated ETFs through two channels. First, institutional investors could act as arbitrageurs in the markets of dominated ETFs. Elton et al. (2004) emphasize that arbitrage is not possible in the market for open-end mutual funds leading to violations of the law of one price, but the market for ETFs has no similar short-sale restrictions. Nonetheless, we find little evidence that arbitrageurs are active in addressing dominated ETFs. Short interest is lower for ETFs that are dominated (6% on average) versus the universe of U.S. equity ETFs (8%). Second, institutional investors could steer investments away from dominated ETFs in their roles as advisors and money managers. Institutional ownership percentages are similar for dominated ETFs (34% on average) versus all ETFs (36%), and institutional ownership is higher for dominated ETFs than for all ETFs in the last several quarters of our sample. These

the-ice-at-madison-square-garden).

findings suggest little difference in investment behavior across institutional and retail investors.

Our aggregate-level results show that investors are collectively making large investments in dominated ETFs. We proceed with an ETF-level analysis to further study these allocations. We hypothesize that investors will invest less in a dominated ETF than they otherwise would given the existence of a dominant alternative. To test this hypothesis, we use a panel regression approach to study ETF size. We first classify ETFs into five mutually exclusive categories: (i) Index ETFs that track well-known indexes, (ii) Quasi-Index ETFs that follow straightforward rule-based strategies (e.g., equal-weighted S&P 500), (iii) Active ETFs that have actively managed portfolios or use proprietary strategies, (iv) Sector ETFs that provide exposure to one of 11 broad industries, and (v) Smart Beta ETFs that pursue various strategies related to Value, Growth, Small Cap, Momentum, Profitability, Quality, and Low Volatility. Within each category, we relate the log of ETF market capitalization to fund characteristics that proxy for the fees, liquidity, performance, uniqueness, and investor awareness of the ETF. In each specification, we also include an indicator variable for ETFs that are dominated at the 95.0% correlation threshold.

Our initial tests relate ETF size to a set of fund characteristics that are relatively easy for investors to observe. Common intuition suggests that investor allocations to ETFs should decrease in fees and increase in liquidity. Regression results generally confirm these predictions across ETF categories with large economic magnitudes. Relative to fees and liquidity, the relations between the prior quarter return and size are much weaker, both statistically and economically. After controlling for these variables, we test the hypothesis that investors will avoid a dominated ETF because a dominant fund exists. This hypothesis predicts a negative association between the dominated ETF indicator variable and ETF size. The results show a positive coefficient estimate for each category, refuting the hypothesis. The association is weak and insignificant for Index ETFs. Estimates for the Quasi-Index, Active, Sector, and Smart Beta ETF categories are significantly positive, and the economic magnitudes are quite large. Quasi-Index ETFs, for example, are 531% larger on average (t-statistic of 15.58) than would otherwise be expected given their fund characteristics. Across categories, the effect ranges from 43% for Sector ETFs (t-statistic of 4.51) to 817% for Active ETFs (t-statistic of 18.33).

To better understand the unexpectedly large investments in dominated ETFs, we introduce additional fund characteristics. Investors may seek ETFs that deliver abnormal returns, strategy uniqueness, or desired risk exposures. Investor awareness may also drive ETF investments in the presence of search costs. We therefore include multiple proxies for performance, uniqueness, and awareness to study allocations to dominated ETFs.

To produce abnormal returns and measure strategy uniqueness for the Index, Quasi-Index, Active, and Sector ETF categories, we follow in the spirit of Berk and van Binsbergen (2015) and use 21 Vanguard ETFs to develop peer-based benchmarks. Abnormal performance bears little relation to ETF size, whereas uniqueness plays a more important role. We initially hypothesize that investors desire Index and Sector ETFs that closely track their benchmarks. In contrast, we expect that investors want Quasi-Index and Active ETFs to provide unique performance, as these ETFs charge higher fees for deviating from standard index benchmarks. Across all categories, we find that ETFs with lower measures of uniqueness are larger. The finding that investors are choosing Quasi-Index and Active ETFs that are more similar to their low-cost index benchmarks is consistent with our evidence of substantial allocations to dominated ETFs.

We generate two new measures of how well Smart Beta ETFs deliver on their promised strategies. These measures quantify the extent to which a Smart Beta ETF can simultaneously gain exposure to the factors associated with its stated strategy and minimize exposure to other factors and idiosyncratic risk. For both measures, we find that Smart Beta ETFs with greater risk in the dimensions of their stated factor strategies have more assets, whereas idiosyncratic risk is negatively associated with ETF size. We also find that investors prefer pure Smart Beta strategies that only track one factor as opposed to ETFs that target multiple factors.

We include three measures that proxy for investor awareness of a given ETF. First, the relations between ETF size and age are significantly positive across categories with large economic magnitudes. Second, we measure search volume from Google Keyword Planner, as investors must be aware of an ETF to search for its ticker. This measure is significantly positively associated with fund size in all categories except Active. Third, we develop a Sponsor Tilt measure that proxies for institutional investor awareness through the fund family channel. Some institutions tend to concentrate investments in ETFs from the same fund sponsor, and Sponsor Tilt measures whether a given ETF belongs to a family with loyal institutional investors. Sponsor Tilt is significantly positively associated with fund size within each category. Overall, we find strong associations between our proxies for investor awareness and ETF size.

After including all of these additional fund characteristics, we find evidence that investors reduce their investments in dominated ETFs for certain types of funds. Specifically, dominated Index ETFs are about 21% smaller (t-statistic of -2.40) than expected given their fund characteristics, and the subset of dominated Sector ETFs that track a broad-based sector index are about 15% smaller (t-statistic of -1.77). Across the remaining categories, we continue to find statistically significant excess allocations to dominated ETFs after controlling for our full set of fund characteristics. Quasi-Index ETFs remain 61% larger (t-statistic of 5.06) and non-index Sector ETFs are 31% larger (t-statistic of 4.45). The coefficients are even larger for Active and Smart Beta ETFs, with implied excess sizes of 233% (t-statistic of 9.69) and 123% (t-statistic of 10.27), respectively. These findings run counter to our hypothesis that investors will reduce their investments in dominated products because of the existence of dominant alternatives.

Our finding that dominated ETFs attract substantial assets should not overshadow the importance of the financial innovation that created exchange-traded funds. The rise of the ETF market has provided investors with access to a set of index ETFs that offer cheap diversification and high intra-day liquidity. Competition has also contributed to declining fees among these index ETFs. Notwithstanding these benefits, our results indicate that ETF investors may be overpaying because of their investment choices. The rapid expansion of listed ETFs has been accompanied by increasing excess costs to investors in dominated ETFs. Many of these dominated ETFs claim unique strategies despite their high correlations with cheaper alternatives, and investors are making large excess allocations to these funds.

Our findings are consistent with other studies that show that investors may not benefit from the increase in available ETFs. Bhattacharya, Loos, Meyer, and Hackethal (2017) show that German ETF investors display poor timing and selection ability in the broad spectrum of ETF listings relative to choosing low-fee, well-diversified ETFs. Box, Davis, and Fuller (2019) document that existing ETFs experience a decline in liquidity when a new related ETF lists and that an increase in listed ETFs does not create downward pressure on expense ratios, and Khomyn et al. (2020) estimate sizable welfare losses when multiple ETFs compete across the fee-liquidity spectrum due to duplicated fixed costs and network inefficiencies. Ben-David, Franzoni, Kim, and Moussawi (2021) show that specialized ETFs, particularly newly listed ones, earn negative risk-adjusted returns. Increasing costs are also consistent with Hortaçsu and Syverson's (2004) prediction of welfare losses from higher search costs as more funds become available. Investors would benefit from focusing on a small set of cheap, liquid index ETFs.

Our study of dominated products in the ETF market is closely related to a literature that focuses on mutual funds. Elton et al. (2004) and Hortaçsu and Syverson (2004) demonstrate considerable variation in expense ratios across S&P 500 index funds despite their nearly identical portfolios. Elton et al. (2004) and Choi, Laibson, and Madrian (2010) emphasize investor irrationality and Hortaçsu and Syverson (2004) attribute the differences in fees to search costs and non-financial differentiation. Boldin and Cici (2010) attribute most of the losses from higher-fee index funds to retail investors who are influenced by brokers and financial advisors with incentives to guide investors into high-fee funds. Our focus on ETFs, which have no broker incentives, allows us to rule out a similar explanation for dominated ETFs. Cooper et al. (2020) find that fee dispersion has persisted both in index funds and in other types of mutual funds, and they estimate large costs to investing in high-fee funds. We provide complementary evidence of economically large costs in the ETF market that are persisting, and even growing, through time.

2 Data

Section 2.1 describes our data sources for ETF characteristics. Section 2.2 discusses the measures we create to examine the performance and uniqueness of ETFs. Section 2.3 presents information about our sample and summary statistics.

2.1 Data Sources

We focus on the universe of U.S. equity ETFs.[4] We identify ETFs as U.S. equity using Lipper codes from The Center for Research in Security Prices (CRSP), and we remove leveraged ETFs by dropping any fund with a leverage factor from Bloomberg that does not equal one. Our sample period is January 2000 through June 2018, and we measure ETF characteristics quarterly to form the panel data.

Our ETF characteristic data are from Bloomberg and CRSP. We collect daily ETF share prices, net asset values (NAVs), shares outstanding, and trading volumes from both Bloomberg and CRSP. We follow Brown, Davies, and Ringgenberg (2020) and use Bloomberg as the primary data source, and we clean these data by removing anomalies that are not verifiable via CRSP. We collect inception dates from Bloomberg. From CRSP, we use the fund sponsors, expense ratios, internal turnover ratios, and bid-ask spreads.

We classify ETFs into five categories based on their strategies: Index, Quasi-Index, Active, Sector, and Smart Beta. We hand classify each ETF in the sample based on Lipper codes and fund descriptions from ETF.com and ETFDB.com. Sector ETFs are identified using Lipper codes, and these ETFs are further classified into 11 sector classifications.[5] Smart Beta funds are identified as such by ETF.com, and we flag these ETFs as Value, Growth, Small Cap, Momentum, Profitability, Quality, and Low Volatility funds based on their stated strategies.

The set of ETFs that are not identified as Sector or Smart Beta ETFs are assigned to the Index, Quasi-Index, and Active categories using fund descriptions from ETF.com and ETFDB.com. Index ETFs are those designed to closely track an index. Whereas many ETFs track indexes that are specifically designed and constructed for use by the ETF (Huang, Song, & Xiang, 2020), we only classify funds as Index ETFs if they track standard indexes from well-known index providers (CRSP, Morningstar, Russell, or S&P Dow Jones) or an exchange (Nasdaq or New York Stock Exchange).[6]

[4] Our sample includes U.S. equity ETFs with various legal structures, including ETFs that are organized as open-end mutual funds (e.g., the iShares Core S&P 500 ETF, IVV), unit investment trusts (e.g., the SPDR S&P 500 ETF Trust, SPY), and share classes of open-end mutual funds (e.g., the Vanguard 500 Index Fund ETF, VOO).

[5] The U.S. ETF sectors are Basic Materials, Consumer Goods, Consumer Services, Energy MLP, Financial Services, Health and Biotechnology, Industrials, Natural Resources, Real Estate, Science and Technology, Telecommunications, and Utilities. We reclassify Energy MLP ETFs as Natural Resources ETFs because the Energy MLP sector is relatively small and does not have a natural benchmark in the data as described further below.

[6] Several Index ETFs track small-cap indexes such as the S&P 600. These ETFs could reasonably be considered

Quasi-Index ETFs follow relatively straightforward rule-based strategies but do not directly track a previously established index. Examples of Quasi-Index ETFs are the Guggenheim S&P 500 Equal Weight ETF (RSP), the SPDR S&P 500 Buyback ETF (SPYB), and the SPDR S&P 500 High Dividend ETF (SPYD). Active ETFs follow more complicated proprietary strategies or have actively managed portfolios.

Each ETF is assigned a benchmark to create performance and uniqueness measures. The benchmarks depend on the ETF category, and we describe how we use these benchmarks in Section 2.2. For the Smart Beta ETFs, we use daily return data for a set of commonly used factors from the asset pricing literature that mirror the stated Smart Beta strategies. The MKT, SMB, HML, RMW, and CMA factors of Fama and French (2015) and the MOM factor are from Kenneth French's website. The BAB factor of Frazzini and Pedersen (2014) and the QMJ factor of Asness, Frazzini, and Pedersen (2019) are from AQR's website. For the remaining categories we use Vanguard ETFs as benchmarks, and we use daily returns from these ETFs to create our measures. Vanguard was an early entrant into the ETF market, such that the benchmark time series span most of our sample, and the Vanguard ETFs we use track highly diversified indexes with low expense ratios. Vanguard funds are also used in the literature as peer-based benchmarks (e.g., Berk & van Binsbergen, 2015). Each Index ETF is matched to a Large-Cap Balanced, Mid-Cap Balanced, or Small-Cap Balanced Vanguard benchmark ETF. To create benchmarks for Quasi-Index and Active ETFs, we use nine Vanguard ETFs in the two-dimensional style grid of Large-Cap, Mid-Cap, and Small-Cap interacted with Value, Balanced, and Growth as well as the Vanguard High Dividend Yield ETF. We detail the construction of the Quasi-Index and Active benchmarks in Section 2.2.2. The Sector ETFs are each assigned the Vanguard ETF from the same sector as a benchmark.[7]

to be either Small Cap Smart Beta funds or Index funds. We follow the ETF.com classification system to designate these funds as Index ETFs rather than Smart Beta ETFs. ETFs that track, for example, the S&P 500 Value index are classified as Smart Beta by ETF.com, and we follow this classification system.

[7]The full set of Vanguard benchmarks is VAW, VB, VBK, VBR, VCR, VDC, VDE, VFH, VGT, VHT, VIS, VNQ, VO, VOE, VOT, VOX, VPU, VTV, VUG, VV, and VYM.

2.2 Performance and Uniqueness Measures

We create measures of performance and uniqueness. Given that ETFs in different categories can have very different strategies and goals, we adopt category-specific approaches to calculating and interpreting these measures. This section describes measures for each category.

2.2.1 Index ETFs

We calculate performance and uniqueness measures for Index ETFs relative to the benchmark Index ETFs that are described in Section 2.1. Specifically, we use the following regression,

$$R_{i,t} = \alpha_i + \beta_i R_{BENCH,t} + \epsilon_{i,t}, \tag{1}$$

where $R_{i,t}$ is the daily ETF excess return and $R_{BENCH,t}$ is the daily excess return on the benchmark ETF. Each regression uses daily data over the past 12 months. We require that each fund has at least 120 daily return observations during this period to estimate the regression. The alpha from this regression measures abnormal performance relative to the benchmark, which can reflect operational efficiencies and costs for Index ETFs. Uniqueness is calculated as $(1 - R^2)$, such that it is inversely related to the regression R^2 from equation (1). Given that investors in Index ETFs are likely seeking funds that closely track a diversified index, they may prefer a low uniqueness measure for Index ETFs.

2.2.2 Quasi-Index and Active ETFs

Quasi-Index and Active ETFs in our sample follow a variety of strategies that range from simple to complex. To account for this strategy variation, we use a relatively large set of ten Vanguard benchmark ETFs that span many potential strategies. As described in Section 2.1, these ten ETFs include nine funds in the style grid and a dividend yield ETF. For each Quasi-Index or Active ETF, we estimate the regression

$$R_{i,t} = \alpha_i + \sum_{j=1}^{10} \beta_{i,j} R_{BENCH(j),t} + \epsilon_{i,t}. \tag{2}$$

This regression estimates the portfolio of benchmark ETFs that most closely mimics the returns of the ETF under consideration. The R^2 from this regression is informative about the uniqueness of the ETF's strategy (Amihud & Goyenko, 2013), and we calculate the uniqueness measure as $(1 - R^2)$. A Quasi-Index or Active ETF that is nearly perfectly spanned by the benchmarks provides relatively little value in terms of helping to complete the market. More unique ETFs may be desirable to investors for these categories. The regression alpha is informative about the ETF's performance relative to the fitted benchmark.

2.2.3 Sector ETFs

We estimate a regression for each Sector ETF following equation (1) with the sector-specific Vanguard benchmark ETF. Alpha and $(1-R^2)$ from this regression are the performance and uniqueness measures.

2.2.4 Smart Beta ETFs

We design Smart Beta measures to capture an ETF's ability to achieve its desired factor exposures while minimizing additional systematic and idiosyncratic risk exposures. We also assess each ETF's performance relative to its benchmark model. As a first step to forming our measures, we estimate factor models for each Smart Beta ETF. Each factor model regression uses daily data over the past 12 months. For each Smart Beta ETF, we estimate three factor models. The first, and most comprehensive, model includes the full set of eight factors that we consider,

$$\begin{aligned} R_{i,t} =& \alpha_i + \beta_{i,MKT} R_{MKT,t} + \beta_{i,SMB} R_{SMB,t} + \beta_{i,HML} R_{HML,t} + \beta_{i,MOM} R_{MOM,t} \\ &+ \beta_{i,RMW} R_{RMW,t} + \beta_{i,CMA} R_{CMA,t} + \beta_{i,QMJ} R_{QMJ,t} + \beta_{i,BAB} R_{BAB,t} + \epsilon_{i,t}. \end{aligned} \tag{3}$$

The second model—a matched Smart Beta factor model—is a restricted version of equation (3) that only includes the market factor and the factors that are associated with a Smart Beta ETF's reported strategy. For example, for an ETF that claims the Value and Small Cap strategies, we include the MKT, HML, and SMB factors. We include as factors HML for both Value and Growth

ETFs, SMB for Small Cap, MOM for Momentum, RMW for Profitability, QMJ for Quality, and BAB for Low Volatility. The third factor model we consider is the single-factor market model.

The first set of Smart Beta ETF measures decomposes the total variance of ETF returns into four components: (i) Market Risk, (ii) Smart Beta Risk, (iii) Other Factor Risk, and (iv) Idiosyncratic Risk. Market Risk is the square root of the total explained variance from the market model. Smart Beta Risk is the square root of the difference between the explained variances from the matched Smart Beta factor model and the market model, and Other Factor Risk is the square root of the additional explained variance from the full factor model in equation (3) relative to the matched Smart Beta factor model. Finally, Idiosyncratic Risk is the square root of the residual variance from equation (3). Smart Beta Risk measures the ETF's exposure to the desired factors. Idiosyncratic Risk captures return variation that is unrelated to the eight factors and may, thus, be more diversifiable, whereas Other Factor Risk measures the impact of unwanted systematic factor exposures that may be more difficult or costly for ETF investors to diversify away.

The second measure is Factor Purity, which captures the proportion of systematic factor risk that is in line with the Smart Beta ETF's stated objective. To calculate Factor Purity, we use the unadjusted R^2s from the market model, the matched Smart Beta model regression, and the full factor model in equation (3). Factor Purity is defined as

$$\text{Factor Purity} = \frac{\text{Smart Beta Model } R^2 - \text{Market Model } R^2}{\text{Full Model } R^2 - \text{Market Model } R^2}. \tag{4}$$

The numerator represents the difference in explained variance from adding the desired factors to the market model, and the denominator measures the incremental explained variance from considering the additional seven systematic factors relative to the market model. This ratio is bounded by zero and one, with a value of zero indicating that the Smart Beta ETF generates no desired factor exposure and a value of one showing that the fund has no unwanted exposure to additional systematic factors.

Three additional measures are the tracking error, the $(1 - R^2)$ uniqueness measure, and the alpha relative to the matched Smart Beta regression benchmark. Tracking error and uniqueness both represent the extent to which a Smart Beta ETF's returns are spanned by the market and its stated factors, and alpha measures abnormal performance.

2.3 Sample Characteristics

Table 1 displays the number and total market capitalization of ETFs in our full sample for each year as well as information across the five categories. The ETF market originated with a small set of Index and broad-based Sector ETFs. The substantial assets drawn by these ETFs invited competition from new fund sponsors and additional listings from early sponsors, and the Index and Sector categories quickly grew in the early years of the ETF market. In recent years, large numbers of Quasi-Index, Active, more specialized Sector, and Smart Beta ETFs have been listed, whereas the set of Index and broad-based Sector ETFs has been relatively stable.[8] Newer entrants into the Quasi-Index, Active, Sector, and Smart Beta categories have filled the ETF market with a wide variety of stated strategies that promise investors unique exposures. As of the end of our sample, the 39 Index ETFs (9% of listed funds) combine to manage over $1.0 trillion (52% of total market capitalization). The remaining categories contain a multitude of ETFs, most of which are much smaller than the average Index ETF. Recent growth in the Smart Beta ETF market is particularly notable, and the 144 Smart Beta funds (32% of listed funds) collectively manage $428.0 billion (22% of total market capitalization) by the end of the sample period.

Table 2 summarizes ETF characteristics across categories. The table shows counts of ETFs and observations in our sample and the sample means of fund characteristics. The dependent variable in the panel regressions in Section 3.2 is the log of market capitalization. We consider several additional characteristics to better understand the main drivers of ETF size. Fees are measured by the expense ratio. We include bid-ask spread and trading turnover as liquidity measures. The average bid-ask spread is calculated as a percentage of NAV, and trading turnover is defined as the average shares traded divided by the average shares outstanding. In addition to measuring the secondary-market turnover of an ETF with trading turnover, we measure its internal turnover (i.e., how often the fund changes its positions) via the turnover ratio.

We include three measures of investor awareness. ETF age is the number of years since the

[8] Recent ETF market growth is consistent with Betermier, Schumacher, and Shahrad (2020), who find mutual fund proliferation is driven by incumbent firms' efforts to "fill up the style grid."

inception date. Search volume is calculated each quarter as the average monthly Google Keyword Planner volume for the ETF ticker from Keywords Everywhere. Finally, sponsor tilt is an ETF-level variable that captures the effect from 13F institutions' tendencies to hold ETFs from the same sponsor. For a given ETF, sponsor tilt measures the relation across 13F institutions between the holdings in the ETF and investments in same sponsor ETFs (excluding the ETF under consideration).[9]

Table 2 indicates that ETF characteristics generally have monotonic patterns across the Index, Quasi-Index, and Active categories. Index ETFs are larger, cheaper, and more liquid than their Quasi-Index and Active counterparts. They are also less unique and have greater investor awareness. Sector and Smart Beta ETFs are generally similar to Quasi-Index ETFs on these dimensions.

Average expense ratios differ considerably across ETF categories. Much has been made of declining fees in the open-end mutual fund and ETF markets during our sample period. Figure 1 plots quarterly average fees of ETFs within each category. Average Index ETF fees fell from 0.21% in 2000 to 0.13% in 2018, and the average Quasi-Index expense ratio declined from a peak of 0.57% in 2005 to 0.36% in 2018. Fees of Active, Sector, and Smart Beta ETFs, on the other hand, are relatively steady throughout our sample. As such, declines in fees seem isolated to ETFs that follow relatively straightforward strategies.[10]

Table 3 shows additional summary statistics for Sector and Smart Beta ETFs. The most popular sectors as shown in Panel A are Natural Resources (15.6% of observations), Financial Services (14.7%), and Science and Technology (13.3%). Among Smart Beta ETFs, Panel B shows that Value (44.1%), Small Cap (29.4%), and Growth (27.6%) are the most popular strategies. Some Smart Beta ETFs claim to provide exposures to more than one factor (1.75 factors on average). For example, while 75 funds specify Value exposure and 50 funds specify Small Cap exposure, 24 of those ETFs claim both strategies.[11]

Panel C of Table 3 summarizes our Smart Beta variables. The statistics on estimated factor

[9] Detailed variable descriptions are in Table A1 of the appendix.

[10] Box, Davis, and Fuller (2020) also find that average ETF expense ratios are steady over this period even though news coverage has emphasized competition on fees.

[11] Furthermore, within those 24 funds, seven also claim Momentum exposure, seven claim Quality exposure, four claim Low Volatility exposure, two claim Profitability exposure, and one claims Growth exposure.

loadings indicate how successfully Smart Beta ETFs gain exposure to the factors. We provide summary statistics for the estimated betas only for the ETFs that claim a particular strategy. That is, the 2,354 quarterly observations for Value Beta are the estimated HML loadings for the ETFs that claim a Value strategy.[12] Small Cap ETFs, as a whole, are able to provide substantial exposure to SMB with an average beta of 0.76, and 90% of Small Cap ETF quarters have a beta on SMB of 0.49 or larger. Among the remaining factors, Smart Beta ETFs are most successful at creating exposures to the Value (average beta of 0.25), Growth (0.19), and Low Volatility (0.24) factors. Average exposures are smallest for the Momentum (0.09), Profitability (0.05), and Quality (-0.00) factors.

Whereas some average factor exposures are small, this finding does not imply that all funds fail to provide their claimed exposures. The 10th to 90th percentile range is quite large for each of the factors, which indicates substantial variation in the performance of Smart Beta ETFs in gaining factor exposures. For example, the 10th to 90th percentile ranges are 0.01 to 0.47 for Value Beta and -0.14 to 0.35 for Momentum Beta. The highest betas indicate that some ETFs are generating substantial risk exposures, but the low end shows that a significant portion of Smart Beta ETFs are actually producing negative exposures to their claimed factors.[13]

Turning to the performance measures, our decomposition of return volatility into Market Risk, Smart Beta Risk, Other Factor Risk, and Idiosyncratic Risk shows that, on average, funds gain more volatility from their stated factor risk than from other factors. Factor Purity shows that ETFs vary significantly in their abilities to isolate Smart Beta exposures. On average, 56% of the systematic factor risk that an ETF is exposed to is attributable to its stated risk factors, but the 10th percentile is only 9% while the 90th percentile is 95%.[14]

[12] Growth Beta is the negative of the ETF's estimated HML beta.

[13] Consistent with our results, Rubesam and Hwang (2019) demonstrate significant variation in Smart Beta ETF exposures to factor returns and Johansson, Sabbatucci, and Tamoni (2020) find that about one-third of Smart Beta funds cannot be distinguished from market index funds.

[14] Glushkov (2016) also finds that Smart Beta ETFs exhibit unintended factor exposures.

3 Results

Section 3.1 presents information about dominated ETFs and the aggregate costs associated with investments in these funds. Section 3.2 studies allocations to the dominated ETFs with a panel regression design. Section 3.3 discusses our findings.

3.1 Dominated ETFs

We identify dominated ETFs as those for which a highly correlated, lower-fee, higher-liquidity ETF exists. ETFs that are highly correlated with a given ETF are identified based on return correlations calculated using daily returns over the past 12 months, and we require at least 120 days of returns. We use primary correlation thresholds of 95.0%, 97.5%, and 99.0% to classify highly correlated ETFs, but we also consider a range of correlations from 90.0% to 99.5% in some analyses. Fee comparisons are based on a weak inequality of expense ratios. Liquidity comparisons use both quarterly average bid-ask spreads and quarterly average dollar trading volume. An ETF must have a lower bid-ask spread and higher volume than a competing ETF to be classified as more liquid.

As an illustration of our dominated ETF classifications, consider the three largest Index ETFs that track the S&P 500 index during our sample period: IVV, SPY, and VOO. The ETFs are highly correlated such that they are candidates for domination. Still, none is dominated in most of the recent quarters in our sample. Taking the fourth quarter of 2017 as an example, VOO charged the lowest expense ratio at 0.04%, followed by IVV at 0.05% and SPY at 0.09%. Although SPY is the most expensive of the three ETFs, it is also more liquid than IVV and VOO such that it is not dominated by either fund. IVV is cheaper than SPY and more liquid than VOO, and VOO is cheaper than both IVV and SPY. Hence, none is dominated in both fees and liquidity during this quarter. IVV subsequently lowered its expense ratio to 0.04% in the second quarter of 2018, matching the low fee of VOO. In this quarter, VOO is dominated by IVV, which is more liquid and equally cheap, but it is not dominated by SPY, which is more expensive.

ETFs with relatively high fees and low liquidity are often dominated by multiple highly correlated ETFs. When we analyze the costs of dominated ETFs, we compare each dominated ETF with the lowest-fee fund in the set of dominant ETFs.[15] An alternative approach of comparing each dominated ETF to the dominant ETF with the lowest bid-ask spread gives qualitatively similar results.

Figure 2 provides information about ETF classifications as of June 2018. Panel A shows the percentage of ETFs that are classified as dominated as well as the percentages in three categories of ETFs that are not dominated: those with the lowest fees among their group of highly correlated ETFs, those with higher liquidity than related ETFs with lower fees, and those that are unique in the sense that no other ETF is highly correlated. We consider a range of correlations from 90.0% to 99.5% in this figure to demonstrate sensitivity to the threshold. Panel B repeats the analysis with AUM.

Figure 2 demonstrates that significant capital is allocated to a large number of dominated ETFs. For example, using the correlation threshold of 95.0% (97.5%) [99.0%], about 39% (20%) [8%] of ETFs are dominated. These dominated ETFs collectively manage about 46% (27%) [17%] of total ETF assets.

The remaining categories of ETFs in Figure 2 provide insights into the ETFs that are not dominated. Many ETFs are not highly correlated with any other ETFs. About 26% of ETFs are unique at the low correlation threshold of 90.0%, and that figure increases to 78% at the 99.0% threshold. These unique ETFs tend to be small, however, and they collectively manage only 1% of total ETF assets for the 90.0% threshold and 25% at the 99.0% threshold. The ETFs with low fees and high liquidity, on the other hand, are relatively few in number but manage about half of ETF assets at higher correlation thresholds.

Figure 3 shows that the large allocation to dominated ETFs is not unique to June 2018. We plot the time series of the percentage of total ETF AUM that is invested in dominated ETFs for the correlation thresholds of 95.0%, 97.5%, and 99.0%. Dominated assets were relatively low early in the sample period when most ETFs faced little competition from highly correlated ETFs. As the market grew and more competing ETFs were listed, the percentage of assets in dominated funds increased. Since 2005, the average percentage of assets in dominated ETFs using the correlation

[15]If multiple dominant ETFs share the lowest fee, we choose the ETF with the lowest bid-ask spread among this group.

threshold of 95.0% (97.5%) [99.0%] is 42% (27%) [11%]. Large investments in dominated ETFs have persisted in the ETF market.

Table 4 provides insights into dominated and dominant ETFs. Panel A reports the number of unique ETFs and the proportions of ETF-quarter observations for dominated and dominant ETFs that occur within each of the five categories. A striking number of funds are dominated at some point in our sample, as 322 (222) [109] unique ETFs have been dominated by 164 (112) [66] unique ETFs at the 95.0% (97.5%) [99.0%] correlation threshold. At the 95.0% correlation threshold, dominated ETFs tend to fall into the Smart Beta (41%) and Sector (27%) categories, whereas dominant ETFs tend to be Index ETFs (62%).

Panel B of Table 4 compares the characteristics of dominated ETFs versus their dominant counterparts. By construction, dominated ETFs have higher expense ratios, higher bid-ask spreads, and lower volume. Dominated ETFs are significantly smaller than their dominant counterparts on average. Dominated ETFs also have lower values for our three measures of investor awareness, as they are younger and have lower search volume and sponsor tilt measures. To a large degree, these differences in characteristics reflect the fact that many ETFs are dominated by large, well-known Index ETFs.

A possibility for why dominated ETFs attract significant assets is that they could outperform relative to their peers. The only significant difference in Alpha in Panel B of Table 4 shows outperformance by dominant ETFs for the 95.0% correlation threshold. Panel C further compares the return moments of dominated and dominant ETFs for both the lagged quarter and the next quarter. There is no evidence of better performance by dominated ETFs in terms of average returns. Dominated ETFs have significantly higher standard deviations relative to dominant ETFs, and there are no consistent patterns of significant differences in skewness and kurtosis across specifications. Overall, the evidence does not support a conjecture that investors are making allocations to dominated ETFs based on superior performance.

We now quantify the aggregate costs of investing in relatively high-fee, low-liquidity ETFs. We calculate costs by comparing dominated ETFs to their dominant ETFs for the 95.0%, 97.5%, and 99.0% correlation thresholds. The aggregate cost calculations include both direct costs from expense ratios and indirect costs from higher trading costs of less liquid ETFs. The costs are calculated as if each dominated ETF had the same expense ratio and bid-ask spread as its dominant ETF. For each quarter, we calculate the extra fees as one-fourth of the difference in annual expense ratios multiplied by the quarterly average market capitalization. We calculate the extra trading costs as one-half of the difference in bid-ask spreads multiplied by the quarterly volume.

Figure 4 displays aggregate quarterly costs from higher fees and additional trading costs. Panel A shows the time series of excess costs based on the 95.0% correlation threshold. The average annual cost from investing in dominated ETFs is $255 million from higher fees and $105 million from higher trading costs. The aggregate cost from extra fees in the last quarter of our sample is more than three times the average. These large excess costs reflect a broader trend in Figure 4 of increasing potential cost savings from moving to lower-fee ETFs, which mirrors the shift discussed in Section 2.3 toward a more crowded market of ETFs promising unique exposures at the cost of higher fees. Near the peak of the financial crisis, extra trading costs were nearly an order of magnitude higher than extra fees. As such, the evidence suggests that expense ratios drive the cost of dominated ETFs in normal times, but trading costs loom large when volatile markets lead to wide bid-ask spreads and high trading volumes.

Table 5 tabulates the aggregate costs of investing in dominated ETFs. Panel A shows the average annual costs under each correlation threshold. The costs decrease as the correlation threshold increases because fewer ETFs are defined as dominated under stricter criteria. Nonetheless, the average annual costs remain economically large across all thresholds. Panel B reports the annualized costs from the last quarter in our sample period, which provide a better representation of current costs given the growth in the ETF industry. Using the correlation threshold of 95.0% (97.5%) [99.0%], the annualized aggregate additional costs of dominated ETFs in the second quarter of 2018 totaled $847 ($398) [$146] million. Panel C reports that the total additional cost estimates from investing in dominated ETFs during our sample period range from $1.0 billion (using the 99.0% correlation threshold) to $6.7 billion (using the 95.0% correlation threshold), such that costs to investors from suboptimally investing in ETFs are economically large.[16] For completeness,

[16] We note that, given our focus on U.S. equity ETFs, these economic costs provide a lower bound for the overall ETF market that includes fixed income, international, commodity, and other types of ETFs.

Panel B of Figure 4 provides total excess fees and trading costs across correlations from 90.0% to 99.5% to show sensitivity to the threshold.

Finally, Figure 5 plots time series of the average institutional ownership of ETFs across all ETFs and for dominated ETFs using the 95.0%, 97.5%, and 99.0% correlation thresholds.[17] Investments in dominated ETFs could persist if investors are irrational or if they face large search costs that prevent them from making better allocations to ETFs. One could expect ex ante that these explanations would apply more to retail investors than to institutional investors, such that we may observe that dominated funds are held in greater proportions by retail investors.

The results in Figure 5 show that, instead, retail and institutional investors have similar allocations to dominated ETFs. At the 95.0% correlation threshold, institutions have held 34% on average for dominated ETFs during the sample period versus 36% for all ETFs. The time-series patterns are also quite similar across the groups. These findings suggest that if investor irrationality or investment search costs are responsible for investments in dominated ETFs, then institutional and retail investors are similarly affected by these issues.

3.2 Allocation of Capital to ETFs

The results in Section 3.1 show that investors are, in aggregate, making large investments in dominated ETFs. We now turn to ETF-level evidence to study these allocations. We specifically study ETF size to discover whether investors are reducing their allocations to dominated ETFs because dominant alternatives exist. To test this hypothesis, we study the relations between ETF assets and fund characteristics using panel regressions. We focus on ETF size rather than fund flows because the excess costs that investors incur from investing in dominated ETFs depend on the total amount invested in the ETF rather than on the quarterly net flow.[18]

The dependent variable in each regression specification is the ETF's quarter-end Log Market Cap.[19] We include a dominated ETF indicator variable based on the 95.0% correlation threshold

[17] We calculate quarterly holdings of institutional investors by aggregating the Thomson Reuters 13F database and scaling by market capitalization plus short interest. We correct for known errors in the database (e.g., Sias, Turtle, & Zykaj, 2016).

[18] Papers that study fund flows include Clifford, Fulkerson, and Jordan (2014) and Dannhauser and Pontiff (2019).

[19] We use Log Market Cap to study the allocation of capital because the distribution of ETF size is highly skewed, but inferences are robust to using ETF market cap or percentage of total quarterly U.S. equity ETF assets as the

as an independent variable in each specification to estimate the marginal effect of the existence of a dominant competitor after controlling for other fund characteristics. We include quarter fixed effects and cluster standard errors at the quarter level.[20] When we consider economic magnitudes using one-standard-deviation changes in independent variables, the standard deviations account for the fixed effects such that they are interpretable as within-quarter standard deviations across ETFs.

Table 6 begins with an examination of ETFs in the Index, Quasi-Index, Active, Sector, and Smart Beta categories with easily observable fund characteristics that measure fees, liquidity, portfolio turnover, and prior quarter returns. Expenses are strongly negatively related to fund assets within each category, consistent with investors' preferences for lower-fee investment options. Within the Index category, a one-standard-deviation increase in Expense Ratio is associated with a 33% decrease in ETF market cap (t-statistic of -6.46). Other categories have similar magnitudes of effects that range from a 17% decrease in size for Active ETFs to a 42% decrease for the Smart Beta category.

Table 6 also shows a role for liquidity. Within the Index category, indications of greater liquidity from lower Bid-Ask Spread and higher Trading Turnover are associated with larger ETF market cap. A one-standard-deviation improvement in liquidity is associated with an increase in ETF size of 130% (t-statistic of -4.24) for Bid-Ask Spread and 87% (t-statistic of 6.03) for Trading Turnover. Liquidity also appears particularly important for Sector ETFs. Among Quasi-Index and Active ETFs, on the other hand, the liquidity measures are not as consistently associated with ETF size.

We also find that trading in the underlying ETF portfolio is related to fund size. Among Active ETFs, investors allocate more to ETFs that are more active in trading (as measured by Turnover Ratio). A one-standard-deviation increase in Turnover Ratio is associated with an 11% increase in size (t-statistic of 4.98) for these ETFs. In contrast, ETFs with less portfolio turnover attract

dependent variable.

[20]In the appendix, we investigate alternative regression specifications. Our inferences are robust to including ETF fixed effects (Table A2) or ETF family fixed effects (Table A3). We also show robustness to clustering standard errors at the quarter and ETF levels (Table A4). Finally, we show our results are robust to using dominated indicator variables based on the 97.5% (Table A5) and 99.0% (Table A6) correlation thresholds as well as using a continuous measure of excess fees and trading costs for each dominated ETF relative to their dominant ETF (Table A7).

significantly more assets in the other categories.

Recent returns may be mechanically related to ETF size if existing investors are sticky, and return chasing behavior by investors could also produce a relation. We find that ETF size is significantly associated with the prior quarter return for Index and Sector ETFs. A one-standard-deviation increase in return is associated with a 10% increase in market cap for Index ETFs (t-statistic of 2.11) and a 10% increase for Sector ETFs (t-statistic of 4.14).[21] Prior quarter return is not significantly related to size in the remaining categories, such that sticky investors and return chasing do not appear to be first-order determinants of ETF size.

Finally, we test the hypothesis that investors decrease their allocations to dominated ETFs, all else equal, because of the existence of a dominant alternative. This hypothesis implies that the dominated ETF indicator variable should be negatively associated with ETF size. The coefficient estimates in Table 6 are inconsistent with the hypothesized effects. The coefficient for Index ETFs is small and statistically insignificant. The dominated ETF coefficient estimates are significantly positive in the remaining categories, which implies that dominated ETFs are larger than would be expected given the other fund characteristics. The economic magnitudes are large at 531% for Quasi-Index ETFs (t-statistic of 15.58), 817% for Active ETFs (t-statistic of 18.33), 43% for Sector ETFs (t-statistic of 4.51), and 580% for Smart Beta ETFs (t-statistic of 31.64). Dominated ETFs, which are competing for assets against dominant funds, are actually larger than would otherwise be expected.

Given these unexpected findings, we expand the set of ETF characteristics in Tables 7 to 9 in an attempt to explain the excess allocations to dominated ETFs. We specifically include variables related to performance, uniqueness, and investor awareness. The independent variables from Table 6 are used as controls in each of these tables.

Table 7 shows results when we include the performance and uniqueness measures from the benchmark analyses developed in Section 2.2. The Alpha coefficient estimate for Quasi-Index ETFs of 0.86 (t-statistic of 2.87) implies a 10% increase in size for a one-standard-deviation increase in Alpha. The coefficient estimates for Alpha are insignificant for the remaining categories and

[21] The positive relation between prior quarter return and size in Sector ETFs is consistent with the style-level feedback trading documented by Broman (2020).

negative for Index, Sector, and Smart Beta ETFs. These results suggest that past abnormal performance explains relatively little variation in ETF allocations.

We predict that the relation between Uniqueness and ETF size will be negative for Index, Sector, and Smart Beta ETFs, as investors likely desire products in these categories that more closely track their stated strategies. This prediction is supported in the data. A one-standard-deviation decrease in Uniqueness among Index ETFs is associated with a 165% increase in size (t-statistic of -6.54). Sector ETFs show a weaker relation between Uniqueness and ETF market cap with an implied 56% increase in size (t-statistic of -9.50). Finally, Smart Beta ETFs have an implied 73% increase in size (t-statistic of -14.82) with a one-standard-deviation decrease in Uniqueness.

Surprisingly, Table 7 also shows that Uniqueness and size are negatively related among Quasi-Index and Active ETFs. Ex ante, we expect a positive relation because investors in these categories are paying higher fees on average for ETFs with strategies that deviate from straight index investments. Nonetheless, we find economically large increases in size of 184% for Quasi-Index ETFs (t-statistic of -24.11) and 174% for Active ETFs (t-statistic of -19.71) associated with a one-standard deviation decrease in Uniqueness. Given the higher fees for ETFs in these categories relative to their Index ETF peer benchmarks, the finding of a strong negative association between Uniqueness and Log Market Cap is consistent with our observation that many dominated ETFs attract substantial assets. Across all categories, investors display a preference for ETFs that more closely track low-cost alternatives even with the large differences in average fees.

Controlling for the performance and uniqueness variables in Table 7 produces changes in inferences about dominated ETFs in some categories.[22] After accounting for investor preferences for ETFs that more closely track their benchmarks, dominated Index ETFs are about 32% smaller (t-statistic of -2.43) than would otherwise be expected and the Sector ETF coefficient is negative and statistically insignificant. The coefficient estimates remain significantly positive in the remaining categories. Dominated ETFs are excessively large in the Quasi-Index (151% larger, t-statistic of 9.10), Active (230% larger, t-statistic of 8.92), and Smart Beta (306% larger, t-statistic of 20.33)

[22]The sample size decreases from Table 6 to Table 7 because of data requirements for the performance and uniqueness variables. We show in the appendix (Table A8) that inferences for the tests in Table 6 are the same using the sample from Table 7.

categories.

We now consider the additional Smart Beta ETF performance measures introduced in Section 2.2. Table 4 shows that many of the dominated ETFs are in the Smart Beta category. It is possible that investors allocate capital to these funds based on their abilities to provide promised risk exposures despite being dominated by competing ETFs. We therefore examine whether the performance of Smart Beta ETFs in delivering on their promised strategies explains the excess allocations to dominated ETFs.

Table 8 introduces the additional Smart Beta performance measures. We include two panel regression specifications. The first regression includes measures of the Market Risk, Smart Beta Risk, Other Factor Risk, and Idiosyncratic Risk measures and the second includes the Factor Purity and Tracking Error measures. We also include indicators for 2+ Smart Beta Flags and 3+ Smart Beta Flags to capture the effects of pure versus combination Smart Beta strategies. Finally, we include fixed effects for the Smart Beta strategies that ETFs follow (e.g., Value or Momentum).

Table 8 produces four broad takeaways: (i) Smart Beta ETFs are rewarded for implementing their stated strategies well, (ii) investors prefer ETFs that avoid idiosyncratic risk, (iii) investors prefer pure Smart Beta strategies that provide exposure to only one factor, and (iv) Smart Beta performance measures do not explain the large allocations to dominated ETFs.

The conclusion that Smart Beta ETFs are rewarded for gaining desired factor exposures and avoiding unwanted risk is supported by both sets of performance measures. Column (1) shows that Smart Beta ETF size is positively related to Smart Beta Risk with a one-standard-deviation increase corresponding to a 13% (t-statistic of 5.47) increase in market cap. On the other hand, a one-standard-deviation increase in Idiosyncratic Risk is associated with a 38% decrease in market cap (t-statistic of -8.29), and the coefficient on Other Factor Risk is statistically insignificant. Column (2) reports that Factor Purity, which captures the proportion of systematic factor risk that is spanned by the stated strategy, is strongly positively associated with Smart Beta ETF market cap. A one-standard-deviation increase in Factor Purity is associated with a 22% increase in market cap (t-statistic of 8.10), whereas an increase in Tracking Error is associated with a 32% decrease (t-statistic of -8.41). Overall, investors are able to identify and invest in Smart Beta

ETFs that deliver on their promised exposures while maintaining low exposures to other risks.

Investors also prefer Smart Beta ETFs that isolate a single factor. The 2+ Smart Beta Flags coefficient of -0.57 (t-statistic of -11.41) indicates that, all else equal, an ETF with two strategies is expected to be about 44% smaller compared with a single-strategy ETF. The 3+ Smart Beta Flags coefficient indicates an additional negative association with size, such that these ETFs are about 61% smaller than single-strategy funds.

Finally, Table 8 indicates that controlling for measures of Smart Beta ETF performance in delivering desired risk exposures only explains a portion of the excess size of dominated ETFs. Whereas dominated Smart Beta ETFs are 453% larger compared with expectations in Table 6, the two specifications in Table 8 indicate effects of 326% and 350%. These performance measures, thus, do not explain the abnormally large allocations to dominated ETFs.

Tables 7 and 8 show that including ETF performance and uniqueness measures is not sufficient to explain the excess allocations to dominated ETFs. We supplement these variables with ETF characteristics related to investor awareness. Search costs may be important in fund selection (Hortaçsu & Syverson, 2004), and the salience of a particular ETF may explain investor allocations despite the other fund characteristics.

Table 9 introduces three ETF characteristics related to investor awareness.[23] ETF age is likely positively related to investor awareness. Older ETFs initially competed in a less-crowded ETF market and may have been more salient to investors, and investors may remember owning these older ETFs in the past. We find that age is significantly positively associated with ETF size across all categories. The economic magnitudes are large. For example, doubling the age of an Index ETF is associated with a 406% increase in size (t-statistic of 29.59). Age is likely related to several aspects of an ETF, but this variable's strong positive relation with size provides initial evidence that funds with greater investor awareness attract more capital.

The remaining two awareness variables—search volume and sponsor tilt—are also positively and significantly related to ETF size across almost all categories. Internet search volume proxies for investor awareness and attention, and it likely captures retail investor awareness given the sheer

[23] A version of Table 9 that includes the coefficient estimates for all control variables is available in the appendix (Table A9).

number of retail investors and the fact that institutions have alternative platforms for information (e.g., Bloomberg). Sponsor tilt, on the other hand, is designed to measure the effect of institutional investor tendencies to invest in ETFs from the same sponsor. A one-standard-deviation increase in Log Search Volume for Index ETFs, for example, is associated with a 17% increase in market cap (t-statistic of 4.45). Search volume is insignificantly related to size for Active ETFs but positive and significant for the remaining three categories. The coefficient estimate for Sponsor Tilt is significantly positive across all categories. A one-standard-deviation increase in Sponsor Tilt is associated with a 39% increase in size (t-statistic of 8.07) for Index ETFs. Overall, these results indicate that investor awareness is a significant driver of allocations to ETFs. This finding is consistent with a substantial role of search costs in determining allocations in the ETF market.

Table 9 shows that dominated Index ETFs are about 21% smaller (t-statistic of -2.40) than predicted based on our full set of fund characteristics. In contrast, we continue to find significant excess allocations in the Quasi-Index, Active, and Smart Beta categories. Dominated Quasi-Index ETFs remain 61% larger (t-statistic of 5.06). The coefficients are even larger in magnitude for Active and Smart Beta ETFs, with implied excess sizes of 233% (t-statistic of 9.69) and 123% (t-statistic of 10.27), respectively.

Dominated products in the Sector category do not attract significant excess assets. Sector ETFs can be further classified into two groups: those that simply track a sector index (Sector Index ETFs) and those that develop more complex strategies (Sector Active ETFs). In the appendix (Table A10), we estimate the regressions from Table 9 separately for Sector Index ETFs and Sector Active ETFs. We find that dominated Sector Index ETFs are significantly smaller with an effect on size of -15% (t-statistic of -1.77), whereas dominated Sector Active ETFs are 31% larger (t-statistic of 4.45) than would otherwise be expected.

Our analysis of dominated ETFs reveals a distinction between the relatively straightforward Index and Sector Index ETFs and the more specialized Quasi-Index, Active, Sector Active, and Smart Beta ETFs. On the one hand, dominated ETFs that track well-known indexes (whether they are non-sector or sector indexes) receive lower allocations on average. On the other hand, dominated ETFs that promise more complex or active strategies attract significant excess assets. This dichotomy suggests investors are more capable of identifying dominated products when comparisons are simpler.

The specification in Table 9 includes many fund characteristics related to ETF fees, liquidity, turnover, performance, uniqueness, and awareness. In the appendix (Table A11), we introduce additional control variables. ETFs that belong to large fund families may have more assets, so we control for the log of family size (excluding the ETF itself). We also control for whether ETFs allow for in-kind creation and redemption, which may affect tax efficiency, as well as whether a given ETF has a focus on environmental, social, and governance (ESG) issues. Finally, the share price of an ETF can deviate from the NAV during trading, and the creation and redemption mechanism generally maintains a close relation between the two. We include the average absolute premium, where the premium is calculated as the difference between the share price and the NAV scaled by the NAV, as a measure of the magnitude of deviations. Inferences about excess allocations to dominated ETFs are unchanged after including these additional fund characteristics.

3.3 Discussion

The cross-sectional findings from Section 3.2 appear consistent with the time-series observation in Section 3.1 that the aggregate costs of dominated ETFs have been increasing over time. The origins of the ETF market are rooted in Index ETFs. These ETFs provide low-cost diversification, similar to open-end index mutual funds, with added features of intraday trading and tax efficiencies. Investors seem able to identify and avoid dominated products in this category, at least to some degree. The decline in Index ETF fees over time shown in Figure 1 is consistent with a competitive market in which fee-sensitive investors allocate less to Index ETFs that are dominated by cheaper funds.

Recent years are marked by a proliferation of listed ETFs, a variety of more complex stated ETF strategies, and greater potential for closet indexing by ETFs that claim to provide unique exposures. As ETF market complexity increases, investors are allocating more and more capital to ETFs that are dominated by highly correlated, lower-fee, higher-liquidity ETFs. Many of these dominated ETFs are newly listed. Strikingly, 46% of the dominated ETFs from the second half of our sample are immediately dominated by an already existing fund (compared with 11% of dominated ETFs from the first half), and dominated ETFs are dominated by an older ETF in 92% of the ETF-quarter observations. Excess allocations to these dominated ETFs and the lack of a decline in fees for non-Index ETFs are consistent with these funds feeling less pressure to compete on price. Apparent differentiation in strategy may be enough for an ETF to attract assets, even when its returns are closely tracking those of cheaper alternatives. Our results suggest investors would benefit from simplifying their search process and focusing on a set of low-cost index ETFs.

4 Conclusion

This paper studies dominated products in the U.S. equity ETF market. We find that investors collectively make substantial investments in a large number of ETFs that are dominated by highly correlated, lower-fee, higher-liquidity ETFs. We find evidence that investors are able to identify and reduce allocations to dominated Index and Sector Index ETFs. Dominated ETFs in other categories, however, counterintuitively receive excess allocations from investors. This finding persists after controlling for a multitude of fund characteristics relating to fees, liquidity, turnover, performance, uniqueness, and investor awareness.

The aggregate cost to investors of allocations to dominated ETFs is economically large. We estimate the total cost of using high-fee, low-liquidity ETFs in the U.S. equity ETF market to be $1.0 billion to $6.7 billion during our sample period. The cost is increasing over time, and annualized excess cost estimates at the end of our sample period range from $146 million to $847 million.

The structure of the ETF market allows us to rule out some potential explanations for these excess costs. In contrast to other markets that have features that limit investor access to alternative products (e.g., share classes in open-end mutual funds), ETF investors are almost universally able to access both a dominated ETF and its dominant alternative. The lack of broker incentives to guide investors into dominated ETFs rules out agency problems as a primary cause. Investor irrationality and search costs remain as potential explanations, although we do include proxies for investor awareness as controls. We also note that our evidence on institutional ownership of dominated ETFs requires that irrationality or search cost explanations apply equally to retail and institutional investors.

Regardless of the underlying cause, our findings suggest that allocations to dominated ETFs are eroding any potential gains from the increases in the number of available ETFs and the variety of ETF strategies. Given the totality of our evidence, we believe the growing complexity of the ETF market in recent years may not be particularly beneficial for investors. We conclude that investors, whether short-term or long-term oriented, would benefit from isolating their attention to a small set of low-cost, high-liquidity ETFs.

References

Allen, J., Clark, R., & Houde, J.-F. (2014). Price dispersion in mortgage markets. *Journal of Industrial Economics*, *62*(3), 377–416.

Allen, J., Clark, R., & Houde, J.-F. (2019). Search frictions and market power in negotiated-price markets. *Journal of Political Economy*, *127*(4), 1550–1598.

Amihud, Y., & Goyenko, R. (2013). Mutual fund's R^2 as predictor of performance. *Review of Financial Studies*, *26*(3), 667–694.

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, *24*(1), 34–112.

Ben-David, I., Franzoni, F., Kim, B., & Moussawi, R. (2021). Competition for attention in the ETF space. *Working Paper, The Ohio State University*.

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

Betermier, S., Schumacher, D., & Shahrad, A. (2020). The global menu of funds. *Working Paper, McGill University*.

Bhargava, S., Loewenstein, G., & Sydnor, J. (2017). Choose to lose: Health plan choices from a menu with dominated option. *Quarterly Journal of Economics*, *132*(3), 1319–1372.

Bhattacharya, U., Loos, B., Meyer, S., & Hackethal, A. (2017). Abusing ETFs. *Review of Finance*, *21*(3), 1217–1250.

Boldin, M., & Cici, G. (2010). The index fund rationality paradox. *Journal of Banking and Finance*, *34*(1), 33–43.

Box, T., Davis, R., & Fuller, K. (2020). The dynamics of ETF fees. *Financial Analysts Journal*, *76*(1), 11–18.

Box, T., Davis, R. L., & Fuller, K. P. (2019). ETF competition and market quality. *Financial Management*, *48*(3), 873–916.

Broman, M. S. (2020). Naïve style-level feedback trading in passive funds. *Working Paper, Syracuse University*.

Brown, D. C., Davies, S. W., & Ringgenberg, M. C. (2020). ETF arbitrage, non-fundamental demand, and return predictability. Forthcoming in *Review of Finance*.

Brown, J. R., & Goolsbee, A. (2002). Does the Internet make markets more competitive? Evidence from the life insurance industry. *Journal of Political Economy*, *110*(3), 481–507.

Choi, J. J., Laibson, D., & Madrian, B. C. (2010). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies*, *23*(4), 1405–1432.

Christofferson, S. E. K., & Musto, D. K. (2002). Demand curves and the pricing of money management. *Review of Financial Studies*, *15*(5), 1499–1524.

Clifford, C. P., Fulkerson, J. A., & Jordan, B. D. (2014). What drives ETF flows? *Financial Review*, *49*(3), 619–642.

Cooper, M., Halling, M., & Yang, W. (2020). The persistence of fee dispersion among mutual funds. Forthcoming in *Review of Finance*.

Dannhauser, C. D., & Pontiff, J. (2019). Flow. *Working Paper, Villanova University*.

Egan, M. (2019). Brokers versus retail investors: Conflicting interests and dominated products. *Journal of Finance*, *74*(3), 1217–1260.

Elton, E. J., Gruber, M. J., & Busse, J. A. (2004). Are investors rational? Choices among index funds. *Journal of Finance*, *59*(1), 261–288.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25.

Glushkov, D. (2016). How smart are smart beta exchange-traded funds? Analysis of relative performance and factor exposure. *Journal of Investment Consulting*, *17*(1), 50–74.

Green, R. C., Hollifield, B., & Schürhoff, N. (2007). Dealer intermediation and price behavior in the aftermarket for new bond issues. *Journal of Financial Economics*, *86*(3), 643–682.

Gurun, U. G., Matvos, G., & Seru, A. (2016). Advertising expensive mortgages. *Journal of Finance*, *71*(5), 2371–2416.

Hortaçsu, A., & Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *Quarterly Journal of Economics*, *119*(2), 403–456.

Huang, S., Song, Y., & Xiang, H. (2020). The smart beta mirage. *Working Paper, University of Hong Kong*.

Johansson, A., Sabbatucci, R., & Tamoni, A. (2020). Smart beta made smart. *Working Paper,* *Stockholm School of Economics*.

Khomyn, M., Putniņš, T. J., & Zoican, M. (2020). The value of ETF liquidity. *Working Paper, University of Technology Sydney*.

Rubesam, A., & Hwang, S. (2019). Do smart beta ETFs capture factor premiums? A Bayesian perspective. *Working Paper, IÉSEG School of Management*.

Sias, R., Turtle, H. J., & Zykaj, B. (2016). Hedge fund crowds and mispricing. *Management Science*, *62*(3), 764–784.

Figure 1: Average ETF Expense Ratio by Category. The figure plots the average quarterly expense ratio of ETFs in the Index, Quasi-Index, Active, Sector, and Smart Beta categories. The sample period is January 2000 through June 2018.

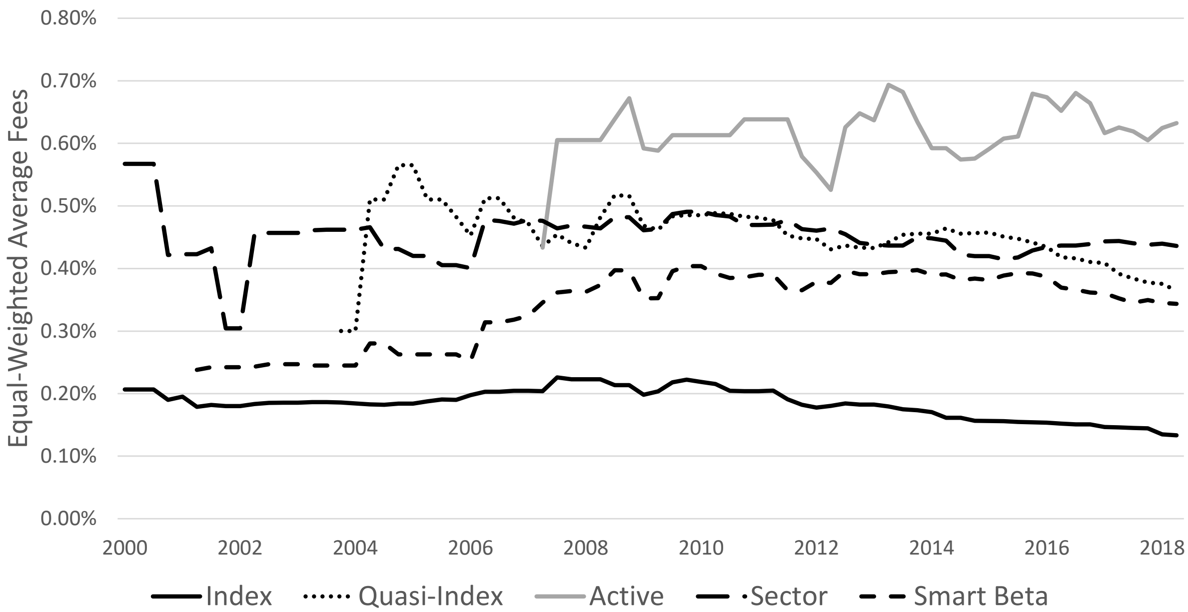
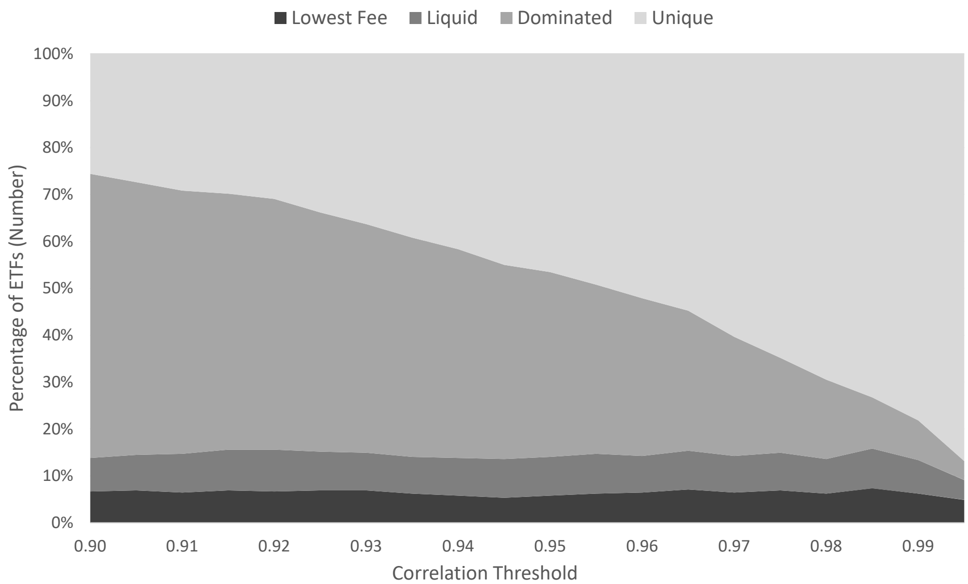
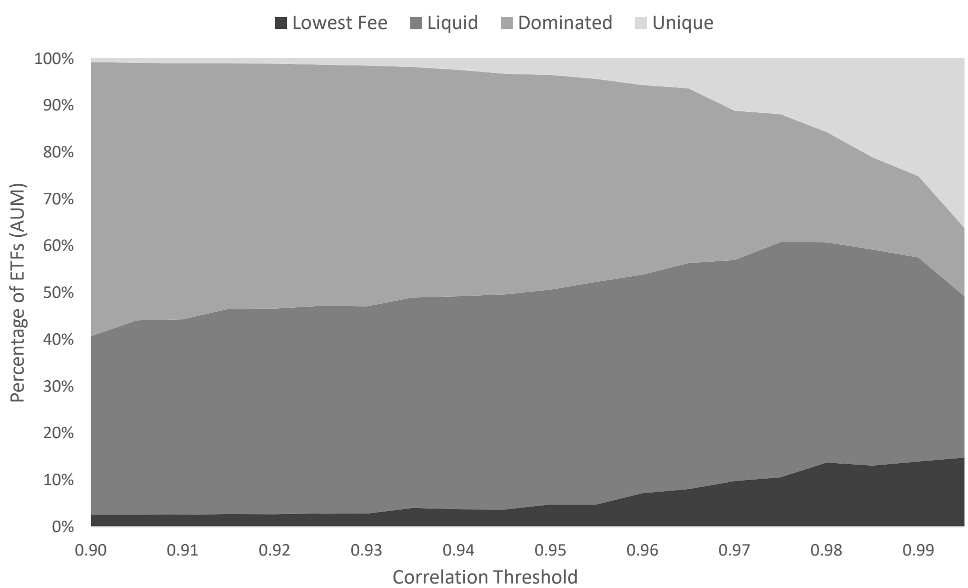


Figure 2: ETFs and Assets in Dominated and Non-Dominated Categories. The figure displays percentages associated with the number of ETFs (Panel A) and cumulative assets under management (Panel B) for four categories of ETFs (lowest fee, liquid, dominated, and unique) as a function of the correlation threshold. Dominated ETFs are defined as those with a dominant ETF that exceeds the return correlation threshold, has a weakly lower expense ratio, has a lower bid-ask spread, and has higher trading volume. Lowest fee ETFs have the lowest fee of all correlated ETFs. Liquid ETFs are correlated with other ETFs, do not have the lowest fee among these ETFs, but are also not dominated. Thus, they provide better liquidity (along at least one dimension) relative to the correlated lowest fee ETF. Unique ETFs have correlations with all other ETFs that are strictly less than the correlation threshold. The sample period is June 2018.



Panel A: Percentage of ETF Types by Number



Panel B: Percentage of ETF Types by AUM

Figure 3: Assets Under Management in Dominated ETFs. The figure plots the total market capitalization of dominated ETFs as a percentage of the total market capitalization of all ETFs in the sample using correlation thresholds of 95.0% (dark solid line), 97.5% (dashed line), and 99.0% (light solid line). Dominated ETFs are defined as those with a dominant ETF that exceeds the return correlation threshold, has a weakly lower expense ratio, has a lower bid-ask spread, and has higher trading volume. The sample period is January 2000 through June 2018.

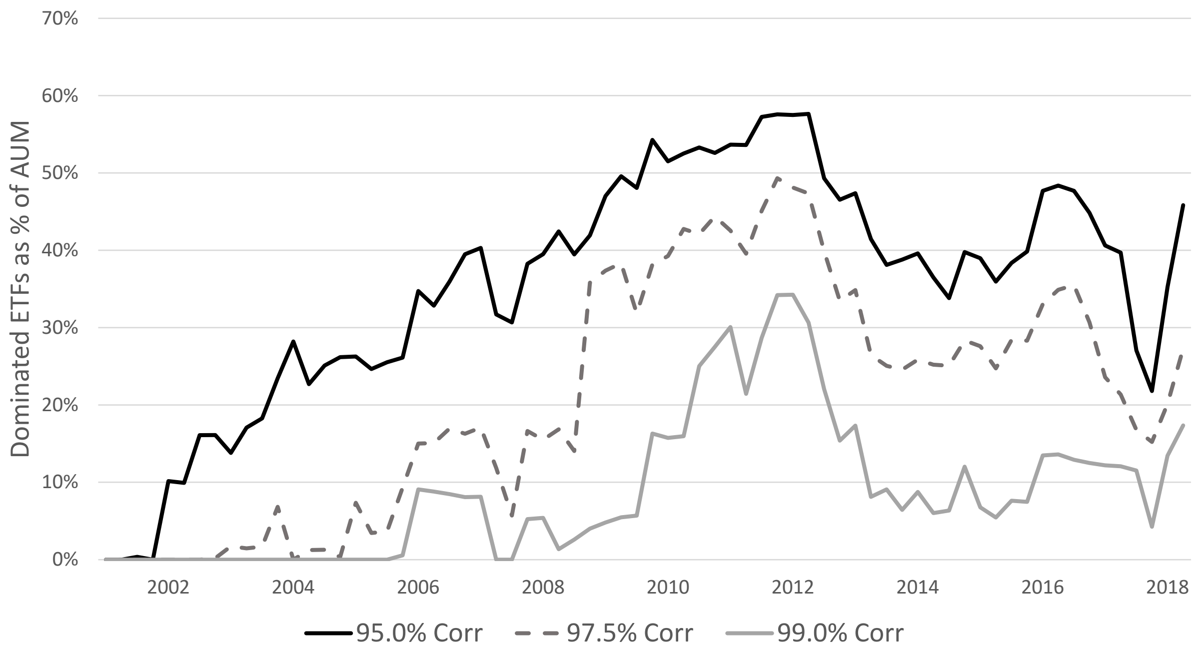
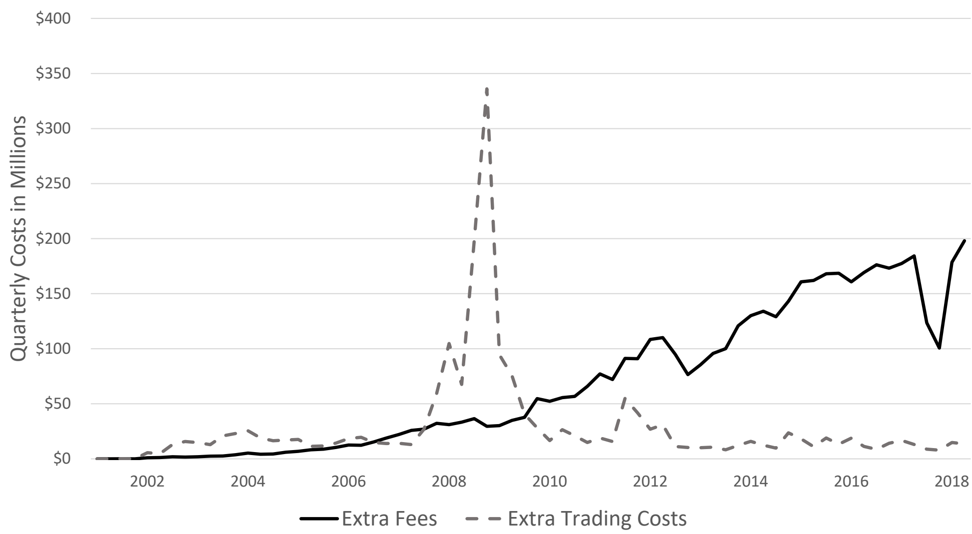
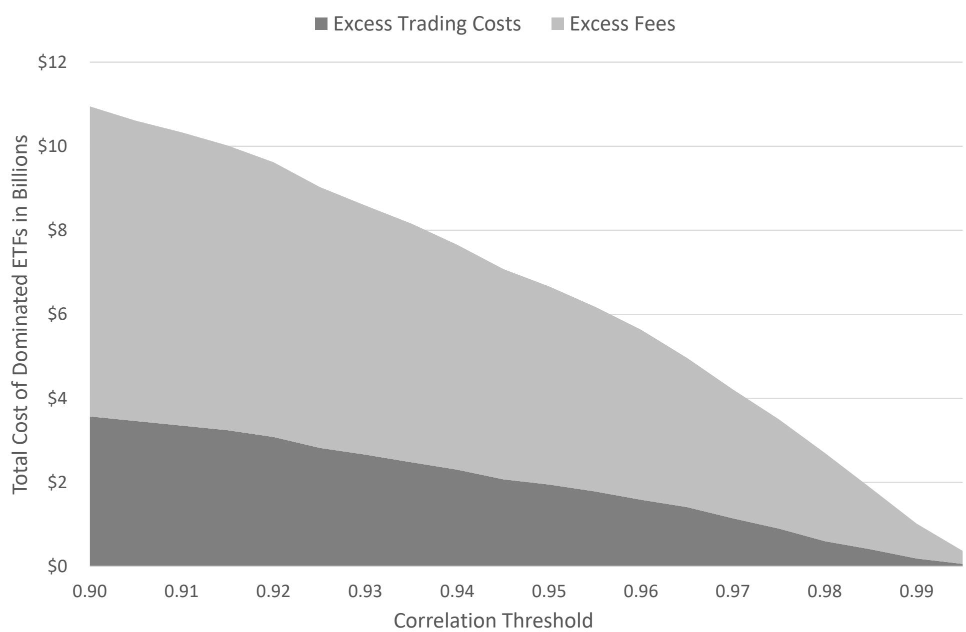


Figure 4: Aggregate Excess Costs of Dominated ETFs. The figure plots the costs of dominated funds. Panel A plots the quarterly cost of dominated ETFs broken down by extra fees (solid line) and extra trading costs (dashed line), based on the 95.0% correlation threshold. Panel B plots the total extra fees and extra trading costs over our full sample as a function of the correlation threshold. Extra trading costs are measured as one-half of the difference in bid-ask spreads between the dominated and dominant ETFs multiplied by the trading volume. Dominated ETFs are defined as those with a dominant ETF that exceeds the return correlation threshold, has a weakly lower expense ratio, has a lower bid-ask spread, and has higher trading volume. The sample period is January 2000 through June 2018.



Panel A: Quarterly Dominated Costs (95.0% Correlation)



Panel B: Total Dominated Costs across Correlations

Figure 5: Institutional Ownership of Dominated ETFs. The figure plots the average institutional ownership of all ETFs in the sample (solid bold line) and of dominated ETFs using correlation thresholds of 95.0% (dashed line), 97.5% (dotted line), and 99.0% (solid line). Dominated ETFs are defined as those with a dominant ETF that exceeds the return correlation threshold, has a weakly lower expense ratio, has a lower bid-ask spread, and has higher trading volume. The sample period is January 2000 through June 2018.

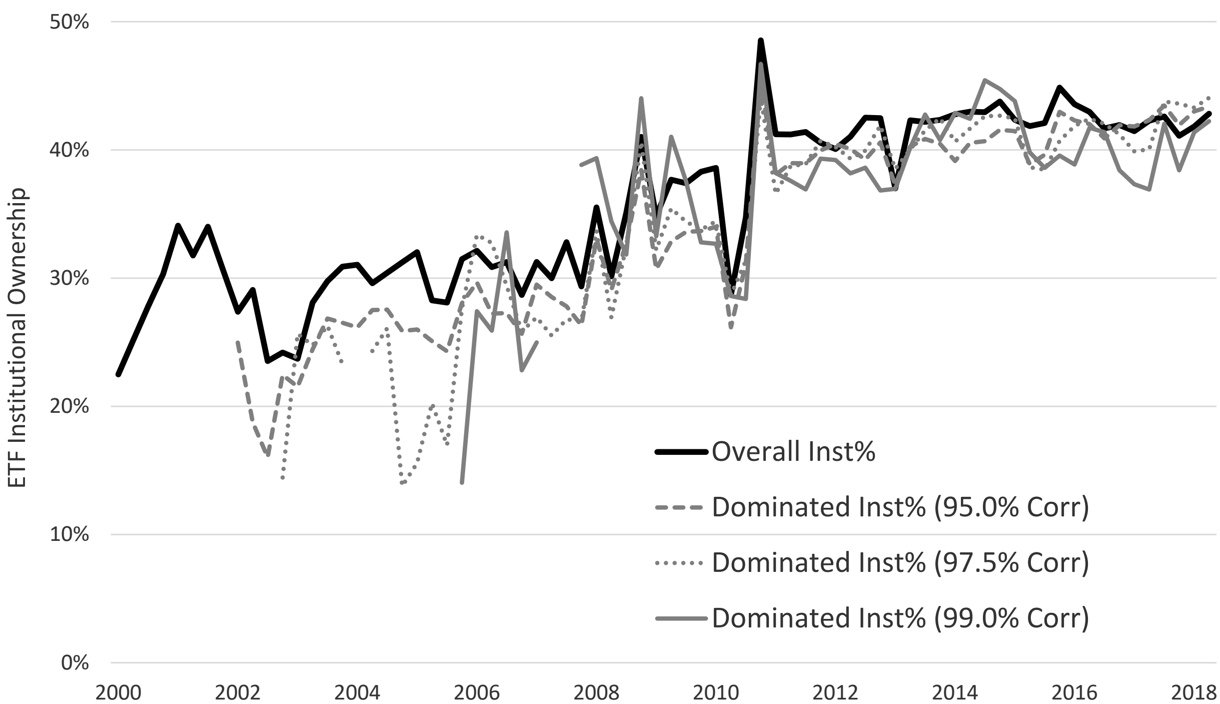


Table 1: Annual ETF Sample

The table reports the number and aggregate market capitalization of ETFs in our sample. ETF classifications are based on Lipper codes and fund descriptions on ETF.com and ETFDB.com. Number and Market Cap are measured at the end of each calendar year, except for 2018 which is measured at the end of June. Market Cap is reported in millions.

	Index ETFs		Quasi-Index ETFs		Active ETFs		Sector ETFs		Smart Beta ETFs		All ETFs	
Year	Number	Market Cap	Number	Market Cap	Number	Market Cap	Number	Market Cap	Number	Market Cap	Number	Market Cap
2000	3	$29,988	0	$0	0	$0	7	$2,259	0	$0	10	$32,247
2001	10	$46,742	0	$0	0	$0	9	$3,469	14	$4,172	33	$54,382
2002	11	$56,407	0	$0	0	$0	26	$5,402	16	$6,965	54	$68,805
2003	12	$76,126	2	$331	0	$0	27	$10,371	16	$13,366	57	$100,195
2004	17	$106,196	2	$732	0	$0	34	$17,656	22	$25,248	75	$149,832
2005	22	$117,744	4	$9,416	0	$0	38	$24,926	30	$33,077	94	$185,164
2006	25	$132,550	10	$11,978	0	$0	57	$35,000	44	$54,270	137	$233,813
2007	28	$201,121	21	$12,006	6	$386	91	$53,846	54	$73,995	200	$341,354
2008	30	$188,103	28	$8,723	10	$230	107	$44,586	69	$55,798	244	$297,440
2009	30	$207,749	31	$12,316	10	$449	115	$63,407	71	$65,789	257	$349,709
2010	33	$235,522	36	$24,781	11	$688	126	$83,719	75	$81,228	281	$425,938
2011	43	$238,173	42	$39,066	8	$802	136	$90,162	86	$83,768	315	$451,970
2012	41	$322,893	49	$50,977	10	$1,203	147	$118,239	100	$106,794	347	$600,106
2013	41	$485,034	43	$81,797	16	$2,912	145	$182,625	108	$159,525	353	$911,893
2014	40	$581,881	52	$103,160	19	$4,224	160	$243,872	119	$213,995	390	$1,147,132
2015	40	$586,250	57	$97,406	29	$4,025	166	$245,604	129	$242,534	421	$1,175,819
2016	40	$739,591	66	$127,953	29	$5,484	176	$280,787	150	$311,833	461	$1,465,648
2017	40	$979,235	71	$155,043	33	$8,689	191	$349,069	152	$405,598	487	$1,897,634
2018	39	$1,020,377	61	$136,924	27	$9,104	178	$352,477	144	$427,979	449	$1,946,861

Table 2: Sample Summary Statistics

The table reports summary statistics for ETFs in our sample split by Index, Quasi-Index, Active, Sector, and Smart Beta ETFs. All variables are defined in Table A1. The sample period is January 2000 through June 2018. Market cap is reported in millions.

	Index	Quasi-Index	Active	Sector	Smart Beta
Number of ETFs	46	92	50	211	170
Number of Observations	2,047	2,057	698	7,046	5,041
Market Cap	$10,262	$1,433	$171	$1,068	$1,555
Expense Ratio	0.180	0.441	0.626	0.450	0.365
Bid-Ask Spread	0.097	0.222	0.456	0.170	0.180
Trading Turnover	2.22	0.53	0.60	1.82	0.67
Turnover Ratio	0.151	0.498	1.505	0.353	0.559
Quarter Return	2.63	2.63	2.38	2.60	2.69
Alpha	−0.041	−0.026	−0.052	−0.009	−0.040
Uniqueness	0.080	0.147	0.246	0.193	0.146
ETF Age	8.38	5.65	3.70	6.94	6.49
Search Volume	638,735	302,295	245,603	553,146	283,815
Sponsor Tilt	0.086	0.035	0.053	0.049	0.073
Dominated	0.518	0.489	0.182	0.267	0.581

Table 3: Sector and Smart Beta Summary Statistics

The table reports summary statistics for Sector and Smart Beta ETFs. Panel A summarizes which sectors are represented, Panel B summarizes which Smart Beta strategies are represented, and Panel C summarizes key Smart Beta variables. All variables are defined in Table A1. The sample period is January 2000 through June 2018.

Panel A: Sector Flags	N	Mean			
Basic Materials	211	0.052			
Consumer Goods	211	0.057			
Consumer Services	211	0.071			
Financial Services	211	0.147			
Health and Biotechnology	211	0.114			
Industrials	211	0.095			
Natural Resources	211	0.156			
Real Estate	211	0.095			
Science and Technology	211	0.133			
Telecommunications	211	0.019			
Utilities	211	0.057			
Panel B: Smart Beta Flags	N	Mean			
Value	170	0.441			
Growth	170	0.276			
Small Cap	170	0.294			
Momentum	170	0.235			
Profitability	170	0.129			
Quality	170	0.153			
Low Volatility	170	0.224			
Total Flags	170	1.753			
Panel C: Smart Beta Summary Statistics	N	Mean	Median	P10	P90
Value Beta	2,354	0.251	0.263	0.006	0.470
Growth Beta	1,937	0.191	0.219	−0.076	0.403
Small Cap Beta	1,426	0.760	0.799	0.487	0.985
Momentum Beta	848	0.090	0.058	−0.138	0.350
Profitability Beta	535	0.050	0.060	−0.249	0.339
Quality Beta	349	−0.004	0.050	−0.317	0.207
Low Volatility Beta	471	0.241	0.217	−0.011	0.548
Market Risk	5,041	1.005	0.859	0.571	1.635
Smart Beta Risk	5,041	0.212	0.153	0.053	0.443
Other Factor Risk	5,041	0.160	0.126	0.066	0.275
Idiosyncratic Risk	5,041	0.340	0.264	0.130	0.595
Factor Purity	5,041	0.555	0.586	0.089	0.946
Tracking Error	5,041	0.380	0.308	0.153	0.656

Table 4: **Dominated ETF Characteristics**

The table displays information on characteristics of dominated and dominant ETFs. Panel A shows the number of dominated and dominant ETFs and the distributions of observations across ETF categories for each correlation threshold. Panel B displays sample means of ETF characteristics for dominated and dominant ETFs and the differences between the sample means, and Panel C shows return moments for the lagged and next quarters. For each correlation threshold, the samples include quarterly observations for dominated ETFs along with the paired dominant ETF. All variables are defined in Table A1, and ***, **, and * indicate statistical significance of differences in sample means at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018. Market cap is reported in millions.

Panel A: Dominated and Dominant ETFs by Category

	95.0% Corr.			97.5% Corr.			99.0% Corr.		
	Dominated	Dominant	Diff.	Dominated	Dominant	Diff.	Dominated	Dominant	Diff.
Unique ETFs	322	164		222	112		109	66	
Observations	7,327	7,327		3,779	3,779		1,378	1,378	
Index	16%	62%	−46%***	20%	58%	−39%***	31%	50%	−19%***
Quasi-Index	14%	4%	9%***	11%	3%	8%***	6%	3%	3%***
Active	2%	0%	2%***	2%	0%	2%***	1%	0%	1%***
Sector	27%	23%	5%***	24%	23%	1%	32%	32%	0%
Smart Beta	41%	11%	30%***	43%	15%	28%***	31%	15%	15%***

Panel B: Dominated and Dominant ETF Characteristics

	95.0% Corr.			97.5% Corr.			99.0% Corr.		
	Dominated	Dominant	Diff.	Dominated	Dominant	Diff.	Dominated	Dominant	Diff.
Market Cap	$2,157	$13,749	−$11,592***	$2,761	$15,012	−$12,251***	$3,288	$17,563	−$14,275***
Expense Ratio	0.34	0.12	0.22***	0.29	0.12	0.17***	0.24	0.12	0.12***
Bid-Ask Spread	0.10	0.04	0.06***	0.07	0.03	0.04***	0.05	0.03	0.03***
Trading Turnover	0.90	2.46	−1.56***	0.85	2.63	−1.78***	0.75	2.93	−2.19***
Turnover Ratio	0.33	0.12	0.21***	0.25	0.12	0.13***	0.15	0.11	0.05***
Alpha	−0.04%	−0.01%	−0.03%**	−0.02%	0.01%	−0.02%	0.04%	0.01%	0.03%
Uniqueness	0.06	0.03	−0.03***	0.04	0.02	−0.02***	0.02	0.01	−0.01***
ETF Age	7.77	9.00	−1.24***	8.66	9.84	−1.19***	9.09	10.64	−1.56***
Log Search Volume	9.89	11.21	−1.32***	10.11	11.60	−1.49***	10.12	11.67	−1.55***
Sponsor Tilt	0.07	0.18	−0.10***	0.08	0.16	−0.08***	0.10	0.14	−0.05***

Panel C: Dominated and Dominant ETF Return Moments

	95.0% Corr.			97.5% Corr.			99.0% Corr.		
	Dominated	Dominant	Diff.	Dominated	Dominant	Diff.	Dominated	Dominant	Diff.
Quarter t Return Mean	2.64	2.67	−0.03	2.85	2.89	−0.03	3.40	3.43	−0.03
Quarter t Return Std. Dev.	8.90	8.42	0.48***	8.86	8.66	0.20***	8.25	8.16	0.09***
Quarter t Return Skewness	−0.80	−0.82	0.02	−0.88	−0.91	0.04	−0.77	−0.80	0.03
Quarter t Return Kurtosis	5.27	5.17	0.10	5.39	5.26	0.13*	4.84	4.76	0.08
Quarter $t+1$ Return Mean	3.00	2.99	0.00	3.33	3.36	−0.04	3.60	3.61	−0.01
Quarter $t+1$ Return Std. Dev.	8.59	8.20	0.39***	8.17	8.06	0.11***	7.74	7.67	0.07*
Quarter $t+1$ Return Skewness	−0.82	−0.91	0.09***	−0.77	−0.83	0.05	−0.80	−0.86	0.06*
Quarter $t+1$ Return Kurtosis	5.60	5.62	−0.01	5.46	5.40	0.06	5.39	5.34	0.05

Table 5: Dominated ETF Excess Costs

The table displays the excess costs of investors using dominated ETFs relative to using the dominant counterparts. Panel A displays the average annual costs from 2001 through June 2018, Panel B displays the annualized costs for the second quarter of 2018, and Panel C shows the total costs from our sample period of January 2000 through June 2018. Extra fees are calculated using the difference in expense ratios between the dominated and dominant ETF pairs multiplied by the dominated ETF's average market capitalization during the quarter. Extra trading costs are calculated using one-half the difference in bid-ask spreads between the dominated and dominant ETF pairs multiplied by the dominated ETF's annual volume. Costs are reported in millions.

Panel A: Average Annual Costs			
	95.0% Corr.	97.5% Corr.	99.0% Corr.
Extra Fees	\$255	\$141	\$45
Extra Trading Costs	\$105	\$49	\$10
Total Average Annual Costs	\$360	\$189	\$55
Panel B: Annualized Costs for Q2 2018			
	95.0% Corr.	97.5% Corr.	99.0% Corr.
Extra Fees	\$792	\$371	\$137
Extra Trading Costs	\$54	\$27	\$9
Total Annualized Costs	\$847	\$398	\$146
Panel C: Total Costs			
	95.0% Corr.	97.5% Corr.	99.0% Corr.
Extra Fees	\$4,715	\$2,600	\$827
Extra Trading Costs	\$1,949	\$901	\$187
Total Costs	\$6,664	\$3,501	\$1,014

Table 6: ETF Size Determinants by Category

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−3.65***	−2.27***	−0.71***	−2.52***	−2.36***
	(−6.46)	(−13.17)	(−5.35)	(−11.22)	(−21.26)
Bid-Ask Spread	−6.94***	−1.36***	−0.33***	−4.30***	−1.29***
	(−4.24)	(−5.99)	(−6.43)	(−5.45)	(−6.14)
Trading Turnover	0.13***	−0.22***	−0.06**	0.09***	0.03
	(6.03)	(−3.97)	(−2.09)	(7.20)	(1.20)
Turnover Ratio	−1.20***	−0.09***	0.02***	−0.34***	−0.12**
	(−3.50)	(−3.53)	(4.98)	(−10.09)	(−2.65)
Quarter Return	0.04**	0.00	−0.00	0.01***	0.00
	(2.11)	(0.19)	(−0.08)	(4.14)	(0.17)
Dominated	0.13	1.84***	2.22***	0.36***	1.92***
	(1.03)	(15.58)	(18.33)	(4.51)	(31.64)
Observations	2,047	2,057	698	7,046	5,041
Adjusted R^2	0.423	0.445	0.368	0.443	0.549
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table 7: Performance and Uniqueness

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. Additional controls are Expense Ratio, Bid-Ask Spread, Trading Turnover, Turnover Ratio, and Quarter Return. All variables are defined in Table A1. Standard errors are clustered at the quarterly level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Alpha	−0.94	0.86***	0.38	−0.02	−0.00
	(−1.55)	(2.87)	(1.36)	(−0.18)	(−0.32)
Uniqueness	−9.99***	−6.45***	−5.93***	−2.59***	−3.72***
	(−6.54)	(−24.11)	(−19.71)	(−9.50)	(−14.82)
Dominated	−0.38**	0.92***	1.19***	−0.03	1.40***
	(−2.43)	(9.10)	(8.92)	(−0.37)	(20.33)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.531	0.527	0.517	0.483	0.577
Additional Controls	Yes	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table 8: Smart Beta Performance

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics for Smart Beta ETFs. Additional controls are Expense Ratio, Bid-Ask Spread, Trading Turnover, Turnover Ratio, Quarter Return, and Alpha. All variables are defined in Table A1. Standard errors are clustered at the quarterly level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap
Market Risk	0.26**	
	(2.31)	
Smart Beta Risk	0.79***	
	(5.47)	
Other Factor Risk	0.07	
	(0.27)	
Idiosyncratic Risk	−1.82***	
	(−8.29)	
Factor Purity		0.66***
		(8.10)
Tracking Error		−1.45***
		(−8.41)
2+ Smart Beta Flags	−0.57***	−0.62***
	(−11.41)	(−13.50)
3+ Smart Beta Flags	−0.37***	−0.55***
	(−5.13)	(−7.27)
Dominated	1.45***	1.50***
	(17.69)	(19.52)
Observations	5,041	5,041
Adjusted R^2	0.592	0.592
Additional Controls	Yes	Yes
Factor Fixed Effects	Yes	Yes
Quarter Fixed Effects	Yes	Yes
Cluster	Quarter	Quarter
ETF Type	Smart Beta	Smart Beta

Table 9: Investor Awareness

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. Additional controls are Expense Ratio, Bid-Ask Spread, Trading Turnover, Turnover Ratio, Quarter Return, Alpha, and Uniqueness. All variables are defined in Table A1. Standard errors are clustered at the quarterly level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Log ETF Age	2.34***	1.56***	1.02***	1.48***	1.13***
	(29.59)	(22.65)	(9.69)	(20.28)	(23.54)
Log Search Volume	0.05***	0.11***	−0.01	0.01***	0.08***
	(4.45)	(8.85)	(−1.13)	(3.34)	(18.71)
Sponsor Tilt	2.45***	0.56**	1.54***	2.22***	2.79***
	(8.07)	(2.49)	(4.56)	(15.05)	(11.19)
Dominated	−0.24**	0.48***	1.20***	0.02	0.80***
	(−2.40)	(5.06)	(9.69)	(0.42)	(10.27)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.775	0.667	0.574	0.626	0.693
Additional Controls	Yes	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

A Variable Definitions and Additional Results

Table A1: Variable Definitions

This table contains the definitions and descriptions of the variables used in the paper.

Variable	Definition
Market Cap	Share price times shares outstanding at quarter end (Sources: Bloomberg and CRSP).
Expense Ratio	The annual expense ratio (Source: CRSP).
Bid-Ask Spread	The mean of the daily bid-ask spread as a percentage of NAV (Source: CRSP).
Trading Turnover	The mean of the daily trading volume in the ETF divided by shares outstanding (Source: Bloomberg).
Turnover Ratio	The annual turnover ratio for the ETF portfolio (Source: CRSP).
Quarter Return	The ETF return for the quarter (Sources: Bloomberg and CRSP).
Alpha	Alpha from the benchmark regression described in Section 2.2 multiplied by 252 (Sources: CRSP, Kenneth French, AQR).
Uniqueness	$(1 - R^2)$ from the benchmark regression described in Section 2.2 (Sources: CRSP, Kenneth French, AQR).
ETF Age	Number of years since fund inception (Source: Bloomberg).
Search Volume	Average monthly Google Keyword Planner search volume (Source: keywordseverywhere.com).
Sponsor Tilt	The target-ETF-share-weighted average of the abnormal sponsor holdings of the 13F institutions that own an ETF. Abnormal sponsor holdings are calculated by subtracting the 13F market share of each sponsor from the sponsor's portfolio weights in each 13F institution (excluding the target ETF) (Source: Thomson Reuters 13F).
Dominated	Indicator equal to one if the ETF has at least a 95.0% correlation in daily returns over the last year with another ETF that has a weakly lower expense ratio, lower bid-ask spread, and higher trading volume (Source: CRSP).
Dominated Cost	The cost, in millions of dollars, of the extra fees and trading costs of the dominated ETF relative to its dominant ETF, where dominated ETFs are defined with the 95.0% correlation threshold. Extra fees are based on one-quarter of the difference in annual expense ratios and extra trading costs are calculated using one-half the difference in bid-ask spreads.
In-Kind Creation	Indicator equal to one if the ETF allows for in-kind creation and redemption (Source: Bloomberg).
ESG	Indicator equal to one if the ETF claims to be focused on environmental, social, and governance issues (Source: ETF.com and ETFDB.com).

Table A1: continued from previous page

Variable	Definition
Average Absolute Premium	The mean of the daily absolute premium, where the premium is calculated as the difference between the price and NAV as a percentage of NAV (Sources: Bloomberg and CRSP).
Institutional Ownership	Quarterly institutional ownership for ETFs as a percentage of available shares. Available shares are shares outstanding plus short interest (Source: Thomson Reuters 13F and Compustat).
Family Size	Sum of market cap at quarter end for all ETFs that share a fund sponsor excluding the current ETF (Sources: Bloomberg and CRSP).
Value	Indicator equal to one if the ETF claims to be a Value ETF (Source: ETF.com and ETFDB.com).
Growth	Indicator equal to one if the ETF claims to be a Growth ETF (Source: ETF.com and ETFDB.com).
Small Cap	Indicator equal to one if the ETF claims to be a Small Cap ETF (Source: ETF.com and ETFDB.com).
Momentum	Indicator equal to one if the ETF claims to be a Momentum ETF (Source: ETF.com and ETFDB.com).
Profitability	Indicator equal to one if the ETF claims to be a Profitability ETF (Source: ETF.com and ETFDB.com).
Quality	Indicator equal to one if the ETF claims to be a Quality ETF (Source: ETF.com and ETFDB.com).
Low Volatility	Indicator equal to one if the ETF claims to be a Low Volatility ETF (Source: ETF.com and ETFDB.com).
Market Risk	Square root of ETF return variance attributable to exposure to market risk.
Smart Beta Risk	Square root of ETF return variance attributable to exposure to factors associated with claimed strategies.
Other Factor Risk	Square root of ETF return variance attributable to exposure to factors not associated with claimed strategies.
Idiosyncratic Risk	Square root of ETF return variance that is not attributable to systematic factor exposures.
Factor Purity	Proportion of variance that is attributable to systematic factor exposures that is contributed by claimed strategies.
Tracking Error	Residual standard deviation from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.

Table A1: continued from previous page

Variable	Definition
Total Flags	The sum of the Value, Growth, Small Cap, Momentum, Profitability, Quality, and Low Volatility indicators (Source: ETF.com and ETFDB.com).
2+ Smart Beta Flags	Indicator equal to one if Total Flags is two or greater (Source: ETF.com and ETFDB.com).
3+ Smart Beta Flags	Indicator equal to one if Total Flags is three or greater (Source: ETF.com and ETFDB.com).
Value Beta	The estimated HML beta from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.
Growth Beta	The negative of the estimated HML beta from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.
Small Cap Beta	The estimated SMB beta from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.
Momentum Beta	The estimated MOM beta from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.
Profitability Beta	The estimated RMW beta from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.
Quality Beta	The estimated QMJ beta from the Smart Beta regression that includes factors associated with claimed Smart Beta strategies.
Low Volatility Beta	The estimated BAB beta from the Smart Beta regression that includes factors associated with claimed smart beta strategies.

Table A2: ETF Fixed Effects

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics with ETF and quarter fixed effects. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−3.43***	1.13***	−0.14	−0.88**	0.48**
	(−3.83)	(4.30)	(−0.12)	(−2.24)	(2.01)
Bid-Ask Spread	−0.71***	−0.11	−0.01	−0.89***	−0.24***
	(−2.81)	(−1.43)	(−0.45)	(−4.15)	(−2.75)
Turnover Ratio	0.03	0.26***	0.01	−0.38***	−0.17***
	(0.14)	(2.67)	(1.45)	(−7.95)	(−4.42)
Trading Turnover	0.00	−0.05*	0.01	0.00	−0.01
	(0.33)	(−1.90)	(0.70)	(1.60)	(−0.34)
Quarter Return	0.00	−0.01	−0.00	0.01***	−0.00
	(0.39)	(−0.92)	(−0.43)	(5.45)	(−0.16)
Alpha	−0.60***	0.39**	0.89***	0.25***	0.01***
	(−5.28)	(2.22)	(3.29)	(5.01)	(4.06)
Uniqueness	−1.74***	−2.44***	−1.75***	−1.72***	−2.02***
	(−5.75)	(−9.58)	(−4.32)	(−9.94)	(−9.06)
Log ETF Age	1.51***	0.41***	0.47*	1.12***	1.13***
	(15.96)	(2.80)	(1.68)	(11.08)	(15.73)
Log Search Volume	0.04***	0.12***	0.04*	0.06***	0.03***
	(4.17)	(6.29)	(1.82)	(4.81)	(5.38)
Sponsor Tilt	0.25*	1.15***	1.68***	2.36***	−0.28
	(1.85)	(4.77)	(3.43)	(12.86)	(−1.50)
Dominated	0.13***	0.26***	0.13	0.02	0.36***
	(4.84)	(4.67)	(1.10)	(0.65)	(8.96)
Observations	1,783	2,007	693	6,691	5,041
Adjusted R^2	0.974	0.934	0.902	0.906	0.940
Additional Controls	No	No	No	No	No
ETF Fixed Effects	Yes	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A3: ETF Family Fixed Effects

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics with ETF family and quarter fixed effects. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−2.49***	−0.47**	−0.16	−1.65***	−3.22***
	(−7.89)	(−2.54)	(−0.30)	(−10.39)	(−16.36)
Bid-Ask Spread	−3.41***	−0.26***	−0.01	−2.06***	−0.59***
	(−2.78)	(−3.24)	(−0.14)	(−4.48)	(−3.75)
Turnover Ratio	−0.52***	0.00	−0.00	−0.66***	−0.29***
	(−3.05)	(0.30)	(−0.20)	(−13.91)	(−7.11)
Trading Turnover	0.08***	0.04	−0.01	0.03***	0.07***
	(8.40)	(0.85)	(−0.47)	(4.85)	(3.21)
Quarter Return	0.01	−0.01	−0.01	0.01***	0.00
	(0.38)	(−0.48)	(−0.78)	(4.52)	(0.11)
Alpha	−1.14***	1.02***	0.46**	0.24***	−0.00
	(−4.13)	(3.74)	(2.06)	(2.94)	(−0.23)
Uniqueness	−6.19***	−3.79***	−3.24***	−1.09***	−3.01***
	(−7.23)	(−11.29)	(−6.28)	(−7.96)	(−14.43)
Log ETF Age	2.86***	1.39***	1.42***	1.59***	1.18***
	(22.13)	(19.47)	(7.72)	(30.65)	(17.79)
Log Search Volume	0.07***	0.12***	−0.00	0.03***	0.10***
	(5.98)	(9.75)	(−0.15)	(7.14)	(18.61)
Sponsor Tilt	1.36***	−0.07	1.17**	2.01***	1.60***
	(4.79)	(−0.20)	(2.64)	(12.01)	(6.19)
Dominated	−0.22***	0.40***	1.02***	−0.06	0.60***
	(−2.88)	(5.03)	(6.20)	(−1.44)	(8.47)
Observations	1,783	2,007	694	6,691	5,041
Adjusted R^2	0.808	0.746	0.751	0.716	0.750
Additional Controls	No	No	No	No	No
ETF Family Fixed Effects	Yes	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A4: ETF Clustering

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. Standard errors are clustered at the quarter and ETF levels, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−2.82*	−2.19***	0.37	−1.88***	−1.82***
	(−1.85)	(−3.25)	(0.78)	(−3.52)	(−4.56)
Bid-Ask Spread	−2.83	−0.32*	−0.06	−2.60***	−0.68***
	(−1.60)	(−1.80)	(−1.38)	(−3.24)	(−2.90)
Turnover Ratio	−0.10	−0.07	0.03***	−0.30	−0.19*
	(−0.10)	(−1.46)	(4.72)	(−1.26)	(−1.73)
Trading Turnover	0.08***	−0.02	−0.04**	0.06***	0.08
	(3.97)	(−0.24)	(−2.31)	(3.60)	(1.56)
Quarter Return	0.00	−0.01	−0.01	0.01***	−0.01
	(0.12)	(−0.47)	(−0.44)	(3.29)	(−0.66)
Alpha	−1.16**	0.53	0.38	0.17	−0.00
	(−2.26)	(1.29)	(0.81)	(0.97)	(−0.47)
Uniqueness	−7.89***	−4.28***	−4.75***	−0.67	−2.58***
	(−4.29)	(−6.16)	(−5.67)	(−1.20)	(−4.10)
Log ETF Age	2.34***	1.56***	1.02***	1.48***	1.13***
	(8.59)	(4.91)	(3.43)	(10.91)	(7.66)
Log Search Volume	0.05	0.11***	−0.01	0.01	0.08***
	(1.08)	(3.09)	(−0.29)	(0.58)	(5.13)
Sponsor Tilt	2.45***	0.56	1.54*	2.22***	2.79***
	(2.81)	(0.37)	(1.97)	(4.64)	(5.96)
Dominated	−0.24	0.48*	1.20***	0.02	0.80***
	(−1.16)	(1.86)	(3.33)	(0.19)	(5.24)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.775	0.667	0.574	0.626	0.693
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	ETF-Quarter	ETF-Quarter	ETF-Quarter	ETF-Quarter	ETF-Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A5: 97.5% Correlation Threshold for Dominated ETFs

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. The Dominated indicator variable uses the 97.5% correlation threshold. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−2.99***	−2.10***	0.41**	−1.87***	−1.66***
	(−8.71)	(−18.92)	(2.17)	(−15.20)	(−14.04)
Bid-Ask Spread	−2.73**	−0.28**	−0.05*	−2.60***	−0.68***
	(−2.26)	(−2.53)	(−1.89)	(−4.36)	(−4.64)
Turnover Ratio	−0.12	−0.07***	0.03***	−0.30***	−0.18***
	(−0.46)	(−4.33)	(6.77)	(−6.33)	(−3.68)
Trading Turnover	0.08***	−0.02	−0.04***	0.06***	0.08***
	(8.91)	(−0.55)	(−3.28)	(6.36)	(3.48)
Quarter Return	0.00	−0.01	−0.01	0.01***	−0.01
	(0.07)	(−0.57)	(−0.64)	(3.38)	(−0.84)
Alpha	−1.12***	0.60**	0.38*	0.17**	−0.00
	(−3.37)	(2.33)	(1.92)	(2.05)	(−0.63)
Uniqueness	−7.56***	−4.81***	−5.66***	−0.67***	−3.30***
	(−9.14)	(−19.57)	(−16.84)	(−4.16)	(−13.35)
Log ETF Age	2.34***	1.55***	1.01***	1.48***	1.14***
	(30.98)	(23.25)	(10.22)	(20.37)	(22.79)
Log Search Volume	0.05***	0.11***	−0.01*	0.01***	0.08***
	(4.70)	(9.83)	(−1.80)	(3.19)	(18.43)
Sponsor Tilt	2.47***	0.61**	1.26***	2.22***	2.57***
	(8.07)	(2.54)	(3.69)	(15.15)	(11.42)
Dominated	−0.11	0.47***	1.14***	0.05	0.82***
	(−1.11)	(8.22)	(7.25)	(1.04)	(12.96)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.774	0.665	0.556	0.626	0.698
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A6: 99.0% Correlation Threshold for Dominated ETFs

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. The Dominated indicator variable uses the 99.0% correlation threshold. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−3.03***	−2.12***	0.14	−1.86***	−1.76***
	(−8.96)	(−18.85)	(0.83)	(−15.29)	(−13.85)
Bid-Ask Spread	−2.74**	−0.27**	−0.01	−2.59***	−0.64***
	(−2.29)	(−2.33)	(−0.32)	(−4.37)	(−4.61)
Turnover Ratio	−0.10	−0.09***	0.03***	−0.29***	−0.21***
	(−0.37)	(−5.28)	(7.55)	(−6.33)	(−4.18)
Trading Turnover	0.09***	−0.02	−0.04***	0.06***	0.11***
	(9.19)	(−0.42)	(−3.20)	(6.36)	(4.24)
Quarter Return	0.00	−0.01	−0.01	0.01***	−0.01
	(0.07)	(−0.52)	(−0.75)	(3.40)	(−0.95)
Alpha	−1.07***	0.59**	0.37*	0.18**	−0.00
	(−3.31)	(2.28)	(1.84)	(2.06)	(−0.68)
Uniqueness	−7.29***	−5.08***	−6.15***	−0.62***	−3.91***
	(−9.49)	(−20.22)	(−18.40)	(−4.03)	(−16.14)
Log ETF Age	2.34***	1.58***	1.02***	1.48***	1.26***
	(31.15)	(21.55)	(9.42)	(20.61)	(31.04)
Log Search Volume	0.06***	0.12***	−0.01	0.01***	0.08***
	(4.91)	(10.29)	(−1.20)	(3.11)	(18.75)
Sponsor Tilt	2.48***	0.54**	1.40***	2.23***	2.78***
	(7.84)	(2.27)	(3.91)	(15.30)	(11.46)
Dominated	−0.00	0.41***	1.64***	0.21***	0.68***
	(−0.02)	(4.31)	(6.93)	(3.94)	(7.64)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.773	0.661	0.538	0.627	0.681
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A7: Continuous Measure for Dominated ETFs

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. The Dominated Cost characteristic measures the excess quarterly dollar cost from additional fees and trading costs for each dominated ETF relative to its dominant ETF. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−3.22***	−2.01***	0.24	−1.93***	−1.89***
	(−9.58)	(−21.43)	(1.61)	(−15.93)	(−15.53)
Bid-Ask Spread	−2.55**	−0.28**	−0.06**	−2.61***	−0.53***
	(−2.17)	(−2.62)	(−2.27)	(−4.35)	(−3.78)
Turnover Ratio	−0.06	−0.07***	0.03***	−0.29***	−0.13**
	(−0.23)	(−4.60)	(7.07)	(−6.41)	(−2.61)
Trading Turnover	0.08***	−0.04	−0.04***	0.05***	−0.01
	(9.33)	(−1.10)	(−3.19)	(6.33)	(−0.33)
Quarter Return	−0.00	−0.01	−0.01	0.01***	−0.01
	(−0.06)	(−0.52)	(−0.48)	(3.46)	(−0.83)
Alpha	−1.05***	0.71***	0.21	0.17**	−0.00
	(−3.15)	(2.77)	(1.05)	(2.05)	(−0.57)
Uniqueness	−7.20***	−4.61***	−5.14***	−0.64***	−3.80***
	(−9.88)	(−20.80)	(−15.93)	(−3.75)	(−15.34)
Log ETF Age	2.27***	1.30***	1.00***	1.47***	1.00***
	(28.34)	(17.44)	(8.95)	(19.80)	(19.13)
Log Search Volume	0.06***	0.11***	−0.01*	0.01***	0.07***
	(4.86)	(9.25)	(−1.99)	(3.00)	(18.23)
Sponsor Tilt	2.44***	1.96***	0.95**	2.19***	2.41***
	(7.87)	(6.71)	(2.19)	(15.54)	(10.76)
Dominated Cost	0.07***	0.43***	2.30***	0.07	0.34***
	(3.09)	(17.91)	(10.27)	(1.23)	(7.96)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.776	0.724	0.591	0.627	0.725
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A8: ETF Size with Constant Sample

The table repeats the regression specifications in Table 6 with the sample used in Tables 7 to 9. The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−2.85***	−2.11***	−0.71***	−2.48***	−2.36***
	(−5.75)	(−11.64)	(−5.33)	(−10.85)	(−21.26)
Bid-Ask Spread	−8.81***	−1.36***	−0.33***	−4.45***	−1.29***
	(−4.20)	(−6.00)	(−6.43)	(−5.09)	(−6.14)
Trading Turnover	0.13***	−0.22***	−0.06**	0.09***	0.03
	(7.43)	(−3.92)	(−2.10)	(7.10)	(1.20)
Turnover Ratio	−1.09**	−0.09***	0.02***	−0.34***	−0.12**
	(−2.65)	(−3.72)	(4.99)	(−10.29)	(−2.65)
Quarter Return	0.02	0.00	−0.00	0.01***	0.00
	(0.76)	(0.26)	(−0.12)	(4.27)	(0.17)
Dominated	0.20	1.81***	2.21***	0.34***	1.92***
	(1.41)	(15.37)	(18.15)	(3.90)	(31.64)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.455	0.432	0.369	0.449	0.549
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A9: ETF Size Complete Specification

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−2.82***	−2.19***	0.37**	−1.88***	−1.82***
	(−7.59)	(−19.94)	(2.31)	(−13.30)	(−12.89)
Bid-Ask Spread	−2.83**	−0.32***	−0.06**	−2.60***	−0.68***
	(−2.28)	(−2.95)	(−2.44)	(−4.35)	(−4.71)
Turnover Ratio	−0.10	−0.07***	0.03***	−0.30***	−0.19***
	(−0.41)	(−4.43)	(6.64)	(−6.33)	(−4.35)
Trading Turnover	0.08***	−0.02	−0.04***	0.06***	0.08***
	(8.49)	(−0.43)	(−3.17)	(6.47)	(3.36)
Quarter Return	0.00	−0.01	−0.01	0.01***	−0.01
	(0.10)	(−0.45)	(−0.49)	(3.38)	(−0.64)
Alpha	−1.16***	0.53*	0.38*	0.17**	−0.00
	(−3.38)	(1.88)	(1.71)	(2.05)	(−0.74)
Uniqueness	−7.89***	−4.28***	−4.75***	−0.67***	−2.58***
	(−8.87)	(−17.92)	(−12.62)	(−3.75)	(−10.03)
Log ETF Age	2.34***	1.56***	1.02***	1.48***	1.13***
	(29.59)	(22.65)	(9.69)	(20.28)	(23.54)
Log Search Volume	0.05***	0.11***	−0.01	0.01***	0.08***
	(4.45)	(8.85)	(−1.13)	(3.34)	(18.71)
Sponsor Tilt	2.45***	0.56**	1.54***	2.22***	2.79***
	(8.07)	(2.49)	(4.56)	(15.05)	(11.19)
Dominated	−0.24**	0.48***	1.20***	0.02	0.80***
	(−2.40)	(5.06)	(9.69)	(0.42)	(10.27)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.775	0.667	0.574	0.626	0.693
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta

Table A10: Sector Splits

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector Index ETFs, Column (5) focuses on non-Index Sector ETFs, and Column (6) focuses on Smart Beta ETFs. Additional controls include Expense Ratio, Bid-Ask Spread, Trading Turnover, Turnover Ratio, Quarter Return, Alpha, and Uniqueness. All variables are defined in Table A1. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap	(6) Log Market Cap
Log ETF Age	2.34***	1.56***	1.02***	2.04***	1.03***	1.13***
	(29.59)	(22.65)	(9.69)	(22.75)	(14.04)	(23.54)
Log Search Volume	0.05***	0.11***	−0.01	0.00	0.01**	0.08***
	(4.45)	(8.85)	(−1.13)	(0.07)	(2.26)	(18.71)
Sponsor Tilt	2.45***	0.56**	1.54***	0.70	1.86***	2.79***
	(8.07)	(2.49)	(4.56)	(1.38)	(10.23)	(11.19)
Dominated	−0.24**	0.48***	1.20***	−0.16*	0.27***	0.80***
	(−2.40)	(5.06)	(9.69)	(−1.77)	(4.45)	(10.27)
Observations	1,783	2,007	696	2,051	4,640	5,041
Adjusted R^2	0.775	0.667	0.574	0.748	0.502	0.693
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector Index	Sector Active	Smart Beta

Table A11: Additional ETF Characteristics

The table displays quarterly panel regressions of Log Market Cap on ETF characteristics. Column (1) focuses on Index ETFs, Column (2) focuses on Quasi-Index ETFs, Column (3) focuses on Active ETFs, Column (4) focuses on Sector ETFs, and Column (5) focuses on Smart Beta ETFs. All variables are defined in Table A1. All Index ETFs in the sample allow for in-kind creation and redemption, and no Index, Sector, or Smart Beta ETFs have an ESG focus. We omit these variables from the corresponding regressions. Standard errors are clustered at the quarter level, t-statistics are shown below the estimates in parentheses, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is January 2000 through June 2018.

	(1) Log Market Cap	(2) Log Market Cap	(3) Log Market Cap	(4) Log Market Cap	(5) Log Market Cap
Expense Ratio	−2.76***	−1.67***	0.62***	−1.94***	−1.87***
	(−8.08)	(−13.51)	(4.09)	(−13.80)	(−12.67)
Bid-Ask Spread	−3.22***	−0.30***	−0.05	−2.59***	0.00
	(−3.68)	(−3.11)	(−0.81)	(−4.34)	(0.03)
Turnover Ratio	−0.32	−0.02	0.03***	−0.32***	−0.21***
	(−1.33)	(−1.39)	(6.89)	(−6.49)	(−4.94)
Trading Turnover	0.07***	0.02	−0.02*	0.06***	0.08***
	(8.51)	(0.41)	(−1.93)	(6.62)	(3.10)
Quarter Return	0.01	−0.01	−0.00	0.01***	−0.00
	(0.34)	(−0.43)	(−0.31)	(3.54)	(−0.32)
Alpha	−1.15***	0.77***	0.56***	0.18**	−0.01
	(−3.50)	(2.92)	(2.71)	(2.18)	(−1.07)
Uniqueness	−7.01***	−3.97***	−4.11***	−0.66***	−2.12***
	(−7.55)	(−17.77)	(−11.29)	(−3.71)	(−9.99)
Log ETF Age	2.57***	1.20***	1.17***	1.49***	1.16***
	(30.05)	(13.72)	(9.90)	(21.92)	(23.37)
Log Search Volume	0.07***	0.12***	−0.00	0.01***	0.08***
	(5.30)	(10.49)	(−0.36)	(2.91)	(20.70)
Sponsor Tilt	1.79***	0.12	1.73***	2.16***	2.55***
	(5.46)	(0.53)	(5.49)	(14.66)	(10.39)
Log Family Size	−0.29***	0.15***	−0.00	−0.01	−0.03**
	(−6.49)	(14.28)	(−0.01)	(−1.01)	(−2.39)
In Kind Creation		0.08	0.68***	1.03***	0.72***
		(0.82)	(4.45)	(10.62)	(4.81)
ESG		−0.64***	2.00***		
		(−7.04)	(7.95)		
Average Absolute Premium	−1.14	0.19***	−0.08	−0.02***	−2.37***
	(−0.92)	(5.81)	(−0.78)	(−5.89)	(−5.03)
Dominated	−0.14	0.62***	1.31***	0.01	0.74***
	(−1.43)	(6.76)	(10.42)	(0.25)	(9.21)
Observations	1,783	2,007	696	6,691	5,041
Adjusted R^2	0.794	0.694	0.591	0.631	0.705
Additional Controls	No	No	No	No	No
Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter
ETF Type	Index	Quasi-Index	Active	Sector	Smart Beta