Obfuscation in Mutual Funds

Ed deHaan
edehaan@uw.edu
Foster School of Business
University of Washington

Yang Song
songy18@uw.edu
Foster School of Business
University of Washington

Chloe Xie
chloexie@mit.edu
Sloan School of Management
MIT

Christina Zhu
chrzhu@wharton.upenn.edu
The Wharton School
University of Pennsylvania

September 29, 2020

Working draft.

Abstract

Mutual funds hold 32% of the U.S. equity market and comprise 58% of retirement savings, yet retail investors consistently make poor choices when selecting funds. Theory suggests that poor choices are partially due to mutual fund managers creating unnecessarily complex disclosures and fee structures to keep investors uninformed and obfuscate poor performance. An empirical challenge in investigating this "strategic obfuscation" theory is isolating manipulated complexity from complexity arising from inherent differences across funds. We examine obfuscation among S&P 500 index funds, which have largely the same regulations, risks, and gross returns but can charge widely different fees. Using bespoke measures of complexity designed for mutual funds, we find evidence consistent with funds attempting to obfuscate high fees. Our study improves our understanding of why investors make poor mutual fund choices, and of how price dispersion persists among homogeneous index funds. We also discuss insights for mutual fund regulation and the academic literature on corporate disclosures.

Thanks to the *JAE* editors, Beth Blankespoor, Matthias Breuer, Alper Darendeli, Zachary Kaplan, Mark Lang, Brian Miller, Olivia Mitchell, Yini Wang, Rachel Zhang, Frank Zhou, several Fidelity professionals, and workshop participants at Georgia Tech, Texas A&M, UT Austin, Wharton, the Miami Winter Conference, and the FARS Midyear Meeting for helpful comments. We thank Dengsheng Chen, Adam Greene, Andri Hail, Douglas King, Shibao Liu, Claudia Peng, Jia Teo, and Jun Wu for data assistance. An Online Appendix is available upon request.

1. Introduction

Over 9,000 mutual funds holding $21.3 trillion in assets were traded on U.S. exchanges during 2019. Mutual funds hold 32% of the total U.S. equity market value and comprise 58% of investors' retirement savings (Investment Company Institute 2020). Despite the popularity of mutual funds, many studies find that they underperform and that retail investors consistently make poor choices when selecting funds.[1] Investor advocates argue that poor mutual fund choices are due in part to overly complex disclosures and fee structures that make it difficult to understand and compare funds, and that unnecessary complexity persists despite decades of regulatory efforts (e.g., SEC 1998; 2009a, 2014, 2018; 2020; Tucker and Xia 2020). Theory suggests that complexity persists because it is part of a strategy to obfuscate unfavorable information and extract rents from retail investors (Carlin 2009). Even small amounts of rent extraction, given the size of the mutual fund market, could have significant implications in aggregate. We empirically investigate whether mutual funds create unnecessarily complex disclosures and fee structures to obfuscate weak net performance.

An econometric challenge in investigating strategic obfuscation in mutual funds is controlling for variation in non-discretionary complexity caused by differences across funds. We mitigate this concern by investigating S&P 500 index funds, which have largely homogeneous gross investment returns and risks but charge different fees so have heterogeneous net returns. For example, Schwab's S&P 500 fund charged 2 basis points (bps) in 2019 while Deutsche's charged up to 506 bps, despite earning nearly identical pre-expense returns (31.46% and 31.47%). Thus, S&P 500 funds provide a setting to examine how disclosures and fee structures

[1] As a few examples, see the following for evidence that mutual funds' net-of-fees performance is worse than that of benchmark portfolios, and that investors often choose high-fee funds even when similar low-fee funds are available: Jensen (1968), Malkiel (1995), Gruber (1996), French (2008), Fama & French (2010), Elton et al. (2004), Frazzini & Lamont (2008), Choi et al. (2010), Evans & Fahlenbrach (2012), and Del Guercio & Reuter (2014).

vary across funds with weaker versus stronger net performance (i.e., due to differences in fees), while holding constant many drivers of non-discretionary complexity.

1.1. Theory and predictions

Our predictions come from Carlin (2009), which models strategic obfuscation among securities like index funds. In Carlin (2009), complex disclosures and fee structures make it difficult for retail investors to understand and compare fees across otherwise identical funds. Some investors become informed and purchase the cheapest fund. Other investors find learning to be too costly, so they remain uninformed and invest randomly.[2] The fraction of uninformed investors is determined endogenously by aggregate processing costs across funds, and funds face a trade-off between capturing more investors and charging higher fees. As depicted in Figure 1, a mixed-strategy equilibrium emerges in which some funds choose low fees and simple disclosures and fee structures, and other funds choose the opposite.

Carlin (2009, p279) describes two complementary methods that high-fee funds use to increase complexity and obfuscate high fees. The first, which increases "narrative complexity," is to use unnecessarily bad writing to make disclosures less readable. The second method, which increases "structural complexity," is to create complex intra-fund structures that make it hard for investors to compare funds and identify the fees they must pay. For example, low-fee funds might have a single class with one annual fee, while high-fee funds have multiple share classes with a combination of fees (see Figure 2).

Based on Carlin (2009), our two predictions are that high-fee funds have greater narrative

[2] Uninformed investors may be aware that they could be over-paying, but expected cost savings from a cheaper fund do not exceed the requisite processing costs. Prior research finds empirical evidence of random allocation of capital by uninformed investors. For example, Huberman & Jiang (2006) find that pension plan participants allocate contributions evenly across the funds offered. This finding supports the "1/n strategy" of naïve diversification described by Benartzi & Thaler (2001) and assumed in Carlin (2009). Also, investors do not reevaluate their fund purchases in Carlin's one-shot game, which is consistent with evidence in prior literature that mutual fund investors do not learn from investment performance (e.g., Goetzmann & Peles 1997).

complexity and greater structural complexity. These predictions are not causal in that complexity causes high fees, but instead the association between fees and complexity is a joint outcome of funds' strategy choice. The major sources of tension in our predictions are that mutual fund disclosures are heavily regulated and that S&P 500 funds are especially simple and standardized compared to other funds, so strategic obfuscation is plausibly unrealistic. Thus, variation in complexity may not exist, or could be benign.[3]

1.2. Data and variable measurement

Our sample spans 1994-2017 and includes 38 S&P 500 index mutual funds. We measure funds' total fees including annual fees and amortized one-time charges (*Fees*).

We examine narrative complexity within funds' prospectuses because research finds that a majority of retail investors use prospectuses to learn about mutual funds (see Section 6.1). Further, we show that websites such as Fidelity copy text directly from prospectuses, so investors can be affected by prospectus readability even if they do not read the filing itself.

We operationalize narrative complexity using four measures of readability. We develop two custom measures of prospectus readability based on guidance from practitioners and the SEC. *FundsinFiling* is the number of unique funds that managers include in a single prospectus (e.g., S&P 500 fund, Russell 3000 fund, etc.). *Repetition* measures the degree to which the summary section of the prospectus exactly repeats language from the details section. We also use two standard narrative complexity proxies, *Length* and *WordsPerSentence*, measured for both the entire prospectus and just in the summary expense disclosure.

We measure structural complexity based on the fund's number of share classes and types

[3] Views differ on whether studies of corporate disclosure find strong support for strategic obfuscation (see Section 7.2). Regardless, it is unclear that findings from corporate disclosures would generalize to mutual funds. Fund disclosures provide less latitude for manipulation and are carefully scrutinized by FINRA, the SEC, and investor advocate groups, so obfuscation is more likely to be detected, especially for homogeneous index funds.

and tiers of fees, combined into a principal component *Structural_Complexity*.

1.3. Analysis and findings

Our sample funds have an average tracking error of only 3.5bps, which indicates that they earn pre-expense returns that closely mirror the S&P 500 index.[4] In contrast, the within-year standard deviation of fees is 51bps, consistent with funds charging different fees. These differences are substantial; e.g., our data indicate that retail investors paid an extra $358M in 2017 alone by holding high-fee versions of S&P 500 index funds.[5] Our data also reveal substantial within-year variation in narrative and structural complexity.

Our analyses are based on OLS and robust regressions with year fixed effects to eliminate common temporal variation in index returns, risks, regulations, and many other non-discretionary fund characteristics. We find evidence consistent with both of our predictions: funds with greater narrative complexity and structural complexity also have higher *Fees*. We continue to find a positive association between *Fees* and narrative complexity just within funds' objective statement and equity risk disclosures, which are specific disclosures that should not be confounded by differences in structural complexity or other unobservable strategies. These findings are consistent with fund managers using both narrative and structural complexity to obfuscate high fees. Numerous additional analyses support our main inferences.

We also examine the role of marketing, which we expect to be more effective when investors are uninformed.[6] We find positive associations between marketing efforts and both narrative complexity and structural complexity, consistent with aggressive marketing being a

[4] Tracking error, defined as the deviation between a fund's gross returns and the S&P 500 index, is an inverse measure of how well the fund tracks the S&P 500 index. Like in Elton et al. (2019), the S&P 500 funds in our sample have small tracking errors. High tracking errors are more common for funds tracking indices with illiquid securities.

[5] Over 30 years, $358M compounds to $6.9B in additional retirement savings.

[6] Traditional advertising is highly restricted for mutual funds, so marketing in recent years is often done by compensating advisors for recommending and selling funds to their clients (e.g., commissions).

complementary strategy for high-fee funds.

Finally, we descriptively analyze the broader mutual fund market. We find that the parent companies of high-fee S&P 500 funds offer fewer other funds and a similar number of other financial products as the parents of low-fee S&P 500 funds. While our tests of S&P 500 funds have better identification, we continue to find a positive association between fees and complexity in parents' broader fund portfolios. These results suggest that fund strategy choices are made at the parent level, and that our findings for S&P 500 funds likely generalize to the broader market.

1.4. Limitations and caveats

While we believe that our tests are sufficiently well-specified to support our inferences, Section 6 discusses potential confounds for readers' consideration. Also, we leave it to future research to devise tests to estimate the impact of obfuscation on investors' choices, in isolation from complementary strategies.[7]

Finally, our findings do not necessarily mean that managers *consciously* create complexity to obfuscate high fees. Alchian (1950) argues that profit-maximizing strategic behaviors can evolve unintentionally through experimentation and mimicry. While it is possible that the managers do not understand that complexity increases investors' processing costs, our findings suggest that high-fee funds have not embraced the SEC's repeated efforts to reduce mutual fund complexity. Moreover, an argument that managers do not recognize the externalities of their behaviors would reinforce the need for strong regulations and enforcement.

1.5. Contributions

Using largely homogeneous index fund fees as an experimental setting, we find evidence

[7] Blankespoor et al. (2020, p90) note that influential studies on strategic disclosure often stop short of examining effects on trading (e.g., Li 2008; Guay et al. 2016): "Some studies go a step further and examine whether strategies are effective at influencing disclosure processing [or] market outcomes, [but…] it can be challenging to control for other actions managers may take along with the strategic disclosure and that also affect outcomes of interest."

that mutual funds create unnecessarily complex disclosures and fee structures to obfuscate unfavorable information. While our tests specifically focus on high fees and index funds, there is little reason to think that active funds do not also use complexity to obfuscate high fees and other drivers of weak performance. For example, active funds might obfuscate how active they are, which prior literature finds is a predictor of performance (e.g., Cremers & Petajisto 2009).

Our findings contribute to the literature investigating why retail investors make poor mutual fund choices. Investment in low-performing funds is attributed to frictions including high search costs and low financial literacy.[8] Our study complements papers examining specific forms of structural complexity used to strategically exploit these frictions. Closely related to our study is Edelen et al. (2012), which finds that funds obfuscate distribution fees by bundling them with broker commissions. Also, Adams et al. (2012) finds that index funds with weak governance have more share classes, higher fees, and higher variation in fees across classes. In addition to examining additional dimensions of structural complexity, our paper finds that funds also manipulate narrative complexity, even within specific disclosures that should be orthogonal to the effects of structural complexity.[9,10]

We also contribute to the nascent literature specifically examining index funds, and help explain how index funds violate the law of one price (Hortaçsu & Syverson 2004; Carlin 2009).

Finally, our findings should inform ongoing regulatory efforts to improve fund disclosures, and provide insights for the corporate disclosure literature. We discuss these

[8] As a few examples: Hortaçsu & Syverson (2004); Sirri & Tufano (1998); Alexander et al. (1998); Barber et al. (2005); Choi et al. (2010).
[9] Descriptive analyses in Philpot & Johnson (2007) find variation in Flesch reading scores across a small sample of 60 disclosures for active mutual funds. Our results conflict with the associations in Philpot & Johnson (2007) in some regards, likely due to our better-specified research design.
[10] Our results are also relevant to the broader literature on mutual fund reporting. For example, research finds that funds (e.g., closet indexers) misrepresent their investment styles and holdings (e.g., Chan et al. 2002; Wermers 2012; Chen et al. 2020), and it is possible that closet indexers also use narrative complexity to obfuscate their styles. Ge & Zheng (2006) and Agarwal et al. (2015) examine fund quarterly versus semiannual reporting.

contributions in our Conclusion.

2. Background information on mutual funds and disclosures

2.1. Mutual fund classes and fee structures

Figure 2 illustrates how mutual funds are structured. Funds are created and sold by institutions such as Schwab. A single fund can be subdivided into share classes, which can have different combinations of fees and may be available only to certain types of investors (e.g., institutions). All classes share the same asset pool so have the same gross returns and risks. Subdividing a fund into classes is discretionary, and many funds have only one class.

Fund classes can have a variety of fees. Most charge an annual management fee. Other annual charges include marketing and service fees ("12b-1" fees), maintenance fees, and "other" fees. Some funds offer waivers or reimbursements if certain conditions are met. Funds can also charge one-time fees when investors buy shares ("front loads") or sell shares ("rear loads"). Front loads can vary depending on the amount purchased. Rear loads can vary by both the amount sold and holding period. The cutoffs for tiers of front and back loads are called "breaks."

Appendix B provides an example of differences in classes and fees between two S&P 500 funds in 2019. In Example 1, Schwab has one share class with one annual fee of 2 bps. The prospectus shows that a $10,000 investment will incur fees of $2 over one year, $6 over three years, etc. In Example 2, Deutsche has five classes with identical assets but different fees, and details from later in the prospectus disclose that classes R6 and S are restricted to certain investors (see our Online Appendix). The cheapest unrestricted class from Deutsche is estimated to cost 232 bps over a one-year holding period (or 59 bps ignoring loads) while the most expensive costs 506 bps, but later details show that actual fees can differ substantially.

Studies have long questioned the motives for structural complexity (e.g., Herman 1963;

Ferris & Chance 1987). Despite funds' claims that multiple share classes and fees allow them to cater to different clienteles, research has found little evidence that clients benefit from multi-class or multi-fee structures.

2.2 Mutual fund disclosures

Mutual fund prospectuses provide information on fund objectives, investments, costs, risks, historical performance, and other details. Funds typically update their prospectuses several times per year. Fund issuers have a choice to file separate prospectuses for each of their funds or to combine multiple funds in one prospectus.

Since 2010, prospectuses begin with a summary section to provide "key information in plain English in a clear and concise format" (SEC 2009a, p1). Summary sections can be disseminated independently from the full prospectus, and are sometimes referred to as "Summary Prospectuses." Summary section information must also be provided on funds' websites, and the information on third-party websites such as Fidelity is frequently copied directly from prospectus summary sections.[11] Thus, investors are likely to see summary section information even if they do not access the prospectus itself.

The examples in Appendix B are from prospectus summary sections. While both funds track the S&P 500 index, differences in their disclosures are immediately apparent. For example, Schwab uses 14 words to describe its objective:

> *"The fund's goal is to track the total return of the S&P 500 Index."*

And Deutsche uses 60 words:

> *"The fund seeks to provide investment results that, before expenses, correspond to the total return of common stocks publicly traded in the United States, as represented by the Standard & Poor's 500 Composite Stock Price Index (S&P 500 Index). The fund invests for capital appreciation, not income; any dividend and interest income is incidental to the*

[11] In Section OA 8 of our Online Appendix we provide results showing that four leading brokerage websites exactly copy funds objective statements from the funds' prospectuses.

pursuit of its objective."

These exact objective statements appear on Fidelity.com. Across the whole prospectus, Schwab averages 24 words per sentence and totals 120,700 words, while Deutsche averages 32 words per sentence and totals 177,271 words.

3. Sample and Variable Construction

We briefly discuss our sample assembly here and provide details in our Online Appendix (OA 2). Appendix A has further details on variable definitions.

3.1 Sample construction

We examine S&P 500 index mutual funds, which are the original type of index fund. We do not expand our sample to other funds because pooling heterogeneous funds would undermine the identification strengths of analyzing funds with homogeneous investments and risks, and because the process of identifying index funds and matching with prospectuses is largely manual. That said, Section 6.5 provides descriptive analyses of the broader mutual fund market.

Our sample starts with S&P 500 index fund share classes that are available to retail investors from 1994 through 2017. We match observations with their most recent prospectus and require the complexity data discussed below, and then aggregate monthly data to the fund-year level. Our sample for tests of full prospectuses includes 458 fund-years and 38 unique funds.[12] Constructing our sample of summary expense disclosures uses XBRL that was not required before 2011, so our tests of summary disclosures are limited to 123 fund-years and 28 funds.

3.2 Measuring Fund Fees

[12] Our samples are aggregated to the fund-level, which explains why they are smaller than those of prior studies of index funds in which analyses are performed at the class-level (e.g., Hortaçsu & Syverson 2004). We must aggregate to the fund-level because prospectuses are typically not available at the class-level. Also, our sample excludes S&P 500 ETFs because ETFs are subject to different reporting requirements. Finally, our sample excludes index mutual funds that are available only to or through institutions, as well as non-U.S. funds. See the Online Appendix (OA 2) for further discussion, and for robustness tests related to how we identify retail share classes.

We follow Gil-Bazo & Ruiz-Verdú (2009) in measuring a fund's total annual cost of ownership, *Fees,* including all annual fees and annualized loads. We amortize loads over seven years. For funds with multiple retail classes, we use the maximum cost across classes.[13]

3.3 Measuring Narrative Complexity

We measure narrative complexity using four measures of readability. Less readable prospectuses should demand greater processing efforts (Blankespoor et al. 2020).

Our first proxy, *FundsinFiling,* is the number of funds included in the prospectus. Investor groups and the SEC have noted that "multiple fund prospectuses contribute substantially to prospectus length and complexity, which act as barriers to investor understanding" (SEC 2009a). We therefore expect that higher *FundsinFiling* implies greater narrative complexity. *FundsinFiling* is available only after 2006, when the SEC required filers to electronically identify separate funds in their filings (SEC 2005).[14]

Our second readability proxy is *Repetition* between the summary section and the rest of the prospectus. The SEC has noted: "the Summary Section is intended to summarize the key information that is important to an investment decision, with more detailed information presented elsewhere… [T]he repetition of substantially the same—or identical—information […] highlights that a fund has not provided a summary [but rather] unnecessary duplication of information" (SEC 2014). We calculate *Repetition* as the percent of sentences in the summary section that are repeated from the rest of the prospectus.

Our third proxy is based on document length. The SEC has repeatedly raised concerns

[13] Fees includes marketing costs, which we separately investigate in Section 6.5. Our Online Appendix (OA 4) discusses robustness tests using several different assumptions for constructing *Fees.*

[14] Analyses in Section 6.4 find that the parent institutions of high-fee S&P 500 index funds offer fewer other mutual funds than do the parent institutions of low-fee S&P 500 index funds. Thus, finding that high-fee S&P 500 funds have more *FundsinFiling* does not appear to be driven by the number funds offered by the parent institution.

that long prospectuses are difficult to understand, and has even considered imposing page limits (SEC 2009a; 2014). Two potential measures of document length are the counts of words and characters, the latter of which has the advantage of also capturing quantitative information (e.g., Li 2008; You and Zhang 2009; Loughran & McDonald 2014). In our various samples the two are highly correlated, so we combine them into a *Length* principal component.[15]

Our fourth proxy, *WordsPerSentence*, captures writing clarity (Loughran & McDonald 2014). The SEC has raised concerns that prospectuses use overly complex language, and in 2009 introduced regulations intending "to improve mutual fund disclosure by providing investors with key information in plain English in a clear and concise format" (SEC 2009a).

We measure *Length* and *WordsPerSentence* for the whole fund prospectus and specifically within the summary expense disclosure (variables *Length_ExpDisc* and *WordsPerSent_ExpDisc*). Given that the summary expense disclosure is intended to be investors' primary source of information about fees, we expect that managers aiming to obfuscate high fees will do so by manipulating this text. *Length_ExpDisc* and *WordsPerSent_ExpDisc* are highly correlated (Table 2), so when used in the same model we combine them into a *Wordiness_ExpDisc* principal component.

3.4 Measuring Structural Complexity

We use five characteristics to measure structural complexity, and combine them into a *Structural_Complexity* principal component. The characteristics used in our *Structural_Complexity* measure are discretionary; i.e., we are unable to document any regulatory requirements for a fund to have these characteristics.

[15] All principal component combinations herein generate one component with an eigenvalue greater than one. Our Online Appendix (OA 3) discusses robustness tests using additional measures of narrative complexity such as file size (Loughran & McDonald 2016) and BOG score (Bonsall et al. 2017).

First, we use the number of retail share classes within a fund (*ShareClasses*). A critique of multi-class funds is that they make it difficult for investors to comparison shop.

Our next two characteristics are indicator variables relating to front loads and 12b-1 fees. Front loads and 12b-1 fees are controversial because few investors understand what they are (NASD 2003; Barber et al. 2005; Beshears et al. 2009), and these fees has been the focus of recent SEC investigations (Wall Street Journal 2019a). A particular concern with 12b-1 fees is that some funds advertise "no load" while charging high 12b-1 fees. The indicator variable *FrontLoad* identifies firms with front loads, and *NoLoad_12b1* is equal to one for funds without a front load but that do have a 12b-1 fee.

Our final two measures, *FrontLoadBreaks* and *RearLoadBreaks,* are counts of the numbers of breakpoints that determine the levels of the front and rear loads. The more breakpoints, the more complex the structure.

4. Summary statistics and descriptive information

Table 1 Panel A provides summary statistics. The rightmost column tabulates the residual standard deviation in each variable after it is orthogonalized to year fixed effects. We highlight a few key details. First, there is considerable variation in *Fees*, with the interquartile range being 20 to 115 bps.[16] Moreover, the standard deviation of *Fees* is virtually unchanged between the raw data and orthogonalized data, indicating that dispersion in *Fees* between funds persists over time. Second, the mean tracking error is small at 3.4 bps. Third, we observe substantial and temporally persistent dispersion in our measures of structural and narrative complexity.

Table 1 Panel B provides descriptive information after dividing the sample into low- and

[16] We assume a seven-year holding period to determine loads and amortize one-time charges, so our *Fees* estimates are lower than the one-year costs estimated by Deutsche in Appendix B. Deutsche estimates that annual costs for Class A will average roughly (754 bps / 5 =) 151 bps over five years and (1,132 bps / 10 =) 113 bps over 10 years.

high-fee funds based on the lowest 20% and remaining 80% of *Fees* by year. Class-level data in rows (i) and (ii) show that low-fee funds tend to have small differences in average *Fees* between their least and most expensive classes (14.5 bps and 16.4 bps, respectively), while the inter-class difference in *Fees* is larger within high-fee funds (55.5 bps and 84.2 bps). Rows (iii) through (viii) find similar patterns for most of the fee components. Also, high-fee funds tend to be significantly more expensive than low-fee funds across the class-level fee components. The lower rows of Table 1 Panel B find that low-fee funds are larger than high-fee funds, as predicted by Carlin (2009), and that low-fee funds also tend to be older but update their prospectuses as often as high-fee funds.[17]

Table 2 finds that our five measures of structural complexity are highly correlated, which motivates our use of the *Structural_Complexity* principal component in most analyses. The four narrative complexity measures also tend to be positively correlated, especially in the expense disclosure sample in Panel B.

Table 3 provides insights about the economic magnitude of the effects of mutual fund price dispersion on investor wealth. Table 3 presents total assets and fees for sample funds with the lowest 20% of fees versus the remaining 80%. The average fees are weighted by assets at the fund-class level to estimate the aggregate fees paid by retail investors.[18] In 2017, investors in low-fee funds paid 13.8 bps (column ii) versus 63.7 bps paid by investors in high-fee funds (column iv). Column (vi) shows that this fee dispersion equated to $358M in 2017; i.e., all else equal, retail investors would have paid $358M less had they held low-fee instead of high-fee

[17] In Carlin (2009), the low-fee funds capture the entire share of informed investors plus an equal share of uninformed investors, while the high-fee funds capture only their share of the uninformed investors. Thus, fund size is endogenous and low-fee funds should be larger than high-fee funds. See Section 6.2 for analyses explicitly incorporating fund size.

[18] These data cannot be directly compared to Table 1 because average fees are weighted by class-level assets.

S&P 500 funds. Over 30 years, $358M would compound to $6.9B in additional savings.[19]

Since our study uses S&P 500 funds to examine a question that is relevant to the broader mutual fund market, we also estimate what our results imply for the total market.[20] Column (vii) of Table 3 shows the $358M high-low fund fee differential in 2017 equates to 0.077 bps of the total assets of our sample funds. Thus, 0.077 bps of total market assets is an estimate of the extra fees that retail investors paid in 2017 for high-fee versions of similar mutual funds. Investment Company Institute (2018) reports that retail investors held $16.7 trillion of mutual funds in 2017. Applying the 0.077 bps rate to the total holdings indicates that retail investors might have saved ($16.7T x 0.077 bps =) $12.9B in 2017 had they held low-fee versions of similar mutual funds. While this estimate is an obviously rough approximation, it may be a low estimate if price dispersion among similar active funds is greater than dispersion among S&P 500 index funds.

To be clear, we are not implying that obfuscation alone causes these disparities in investor wealth. The purpose of these analyses is to highlight that differences in fees between funds have significant effects. Thus, obfuscation is likely economically important even if it is responsible for only a portion of total excess fees paid.

5. Primary analyses

As depicted in Figure 1, Carlin (2009) predicts that fund fees and complexity are simultaneous outcomes of the manager's choice of fund strategy. Thus, complexity does not cause high fees or vice versa.

As we cannot observe managers' strategic choices, we cannot perform typical regressions in which outcome variables *Y* (in our case, high fees and complexity) are regressed on

[19] Calculated assuming that the $358M is invested in S&P 500 funds earning the average S&P 500 dividend-reinvested gross return of 10% less the 2017 low-fee rate of 13.8 bps: $358M*e^(9.862%*30years) = $6.9B.
[20] These extrapolations are in the spirit of French (2008) and are admittedly rough approximations.

independent variables *X* (the manager's strategic choice). Instead, our empirical strategy is to test whether the two outcome variables from Carlin (2009) are associated with one another in the way predicted by the model. If managers aim to obfuscate high fees with complex disclosures and fee structures, then high fees and complexity should be positively associated.

Figure 3 provides visual evidence of the associations between *Fees* and complexity. The panes plot the average of each complexity measure (vertical axes) by quintile of *Fees* (horizontal axis). For brevity we only plot combined *Structural_Complexity*. The plots show positive associations between *Fees* and complexity, especially in the expense disclosure sample.

5.1 Regression tests

We formally test the associations between fees and complexity using OLS regressions. Because fees and complexity are both outcome variables, either can be the left-side variable in an OLS model. We include fees as the left-side variable so that we can investigate multiple complexity measures on the right-side at the same time:[21]

$$Fees_{i,y} = \beta_1 Complexity_{i,y} + \Sigma\beta_y Year_y + \varepsilon_{i,y} \quad (1)$$

Fees is the total annual cost of ownership for fund *i* in year *y*. *Complexity* is one or more of our complexity variables. β_1 estimates the statistical association between complexity and fees, which we predict is positive. $Year_y$ fixed effects eliminate common temporal trends in the S&P 500 return, risks, regulations, and many other non-discretionary characteristics of the index funds, as well as common temporal trends in fees and complexity. Standard errors are clustered by fund. We do not cluster by year because some tests include as few as seven years of data. Given our

[21] OLS regressions are often used to test for causal relations between *X* and *Y*, but there is nothing necessarily causal about OLS estimates. OLS tests for statistical associations between *X* and *Y* and is appropriate for our purposes. Our Online Appendix (OA 5) presents regressions in which *Fees* is the right-side variable, all of which produce qualitatively similar results. "Qualitatively similar" means that the sign and significance of the coefficient of interest remains unchanged at the 10% level.

small sample sizes, winsorizing is potentially ineffective in mitigating the effects of extreme observations. Thus, we also perform robust regressions instead of OLS (Leone et al. 2019).

Table 4, Panel A presents results of model (1) for narrative complexity in the full prospectus. Columns (i) through (iv) examine each of *FundsinFiling*, *Repetition*, *Length*, and *WordsPerSentence*. The sample size changes because *FundsinFiling* is available post-2006 and *Repetition* is available post-2010. *FundsinFiling* and *Repetition* have significantly positive associations with *Fees*. *Length* and *WordsPerSentence* have insignificant associations with *Fees*. Columns (v) and (vi) examine the readability measures together and find that coefficients on *FundsinFiling* and *Repetition* are positive and significant using OLS and using robust regression.[22]

Panel B presents stronger results for the expense disclosures, which is expected because the summary expense disclosures discuss the fees that managers likely aim to obfuscate. Columns (i) through (iv) find that each of our narrative complexity measures is positively associated with *Fees*. Columns (v) and (vi) present results considering all readability measures simultaneously. As discussed, we combine *Length_ExpDisc* and *WordsPerSent_ExpDisc* into *Wordiness_ExpDisc* because the variables are highly correlated. We find significantly positive coefficients on all measures using OLS and robust regression.

Overall, the results in Table 4 are consistent with managers using narrative complexity to obfuscate high fees. Results indicate that a one standard deviation increase in the expense disclosure narrative complexity measures is associated with 28 – 45 bps higher fees.

Results for structural complexity are tabulated in Table 5. For brevity we present robust

[22] Unless otherwise noted, all regressions that include multiple complexity measures have variance inflation factors below 10. For brevity we only tabulate and discuss robust regression results for the most complete models that include all complexity measures.

regression results only for the combined *Structural_Complexity* measure. All measures of structural complexity are highly significant. These results are consistent with managers obfuscating high fees with structural complexity, and indicate that a one standard deviation increase in structural complexity for the expense disclosure sample is associated with 35 – 52 bps higher fees.

5.2. Analyses within specific disclosures

It is possible that our narrative complexity proxies also capture the effects of structural complexity or other unobservable strategies. If so, it would be unclear whether their positive associations with *Fees* are due to managers manipulating narrative complexity *per se*. *Length* and possibly *WordsPerSentence* are affected by structural complexity. It is less clear why structural complexity or unobservable strategies would affect *FundsinFiling* and *Repetition* (e.g., structural complexity should not force issuers to also include more funds in a prospectus), but the possibility remains. We address this concern by examining specific disclosure items that should be unaffected by differences in structural complexity or unobservable strategies.[23]

We first examine funds' objective statement disclosures. S&P 500 funds ostensibly have the same objective, and it seems unlikely that differences in fund structure, marketing, parent characteristics, or other unobservable tactics should directly affect the narrative complexity of how funds describe their objective. Yet, the examples in Appendix B indicate that Deutsche's objective statement is less readable than Schwab's, and results in Table 6 Panel A find a significantly positive association between the objective disclosure *Length* and fund *Fees*. These results are consistent with high-fee funds having more narratively complex objective statements.

[23] Our intuition is that funds use complexity throughout prospectuses to increase processing costs and keep investors uninformed. For example, if investors do not understand disclosures that come before fees in the prospectus (e.g., funds' objective statements), they plausibly conclude that understanding the entire prospectus is too difficult, and therefore stop reading.

We find a positive but insignificant association for *WordsPerSentence*.

We next examine funds' "equity risk" disclosures. Equity risk is the generic risk that prices rise and fall, and most funds disclose equity risk as a risk factor in both their summary and detailed prospectus sections. It seems unlikely that equity risk is affected by structural complexity or other strategies, so these disclosures provide a relatively clean setting to examine how well funds summarize information between the prospectus summary and details.

We manually collect the equity risk disclosures for the most recent year for each of the 28 funds in our post-2010 sample. As explained in our Online Appendix (OA 2), we are able to locate stand-alone equity risk disclosures for 25 funds. Examples for Schwab and Deutsche are in Appendix C. Based on our reading, the summary risk disclosures do not discuss issues relating to structural complexity or other factors that could confound our tests.

We measure summarization between the detailed and summary disclosures based on the difference in character count, such that a larger value of *Summarization* indicates greater summarization (i.e., lower narrative complexity). We create an indicator, *Summarization_Zero*, if the disclosure is identical between the two sections (has high narrative complexity).[24] We also examine *Length* and *WordsPerSentence* of the summary disclosures.

Results in Table 6 Panel B find significant associations between fees and all measures of narrative complexity. High-fee funds have less summarization (column i), are more likely to have exact repetition (column ii), and have longer and wordier summary disclosures (columns iii and iv). In sum, Table 6 results indicate that high-fee funds use narrative complexity as part of an

[24] *Summarization* and *Summarization_Zero* are intended to capture a similar construct as *Repetition* but are better specified than *Repetition* for these particular tests. *Repetition* identifies nearly identical sentences and can compare blocks of text that may not discuss exactly the same topics. Analyses within specific disclosures involve small blocks of text that we know discuss the same topic between the summary and detailed sections, so *Summarization* and *Summarization_Zero* are better suited for this context.

obfuscation strategy, in disclosures that should substantively unaffected by structural complexity or unobserved differences across funds.

5.3. Discussion

Our main results provide compelling evidence that high-fee funds have more complex disclosures and fee structures than low-fee funds. Here we discuss potential confounds.

One validity threat is that a correlated omitted variable in model (1) causes the positive association between *Fees* and *Complexity* for reasons other than strategic obfuscation. Model (1) intentionally does not control for fund characteristics such as size or tracking error because: i) while they likely affect *Fees*, they are unlikely to also affect *Complexity* in the way necessary to cause omitted variable bias; and ii) some characteristics are endogenous outcomes of the pricing strategy and therefore over-control the model. We also do not include variables that capture potential complementary strategies taken by high-fee funds. Still, Section 6 presents results controlling for fund characteristics, and discusses marketing and other complementary strategies.

A related validity threat is that our proxies for narrative and structural complexity could be compromised if they unintentionally capture other constructs. Results in Section 5.2 address the concern that our narrative complexity measures are confounded by structural complexity or other strategies. Section 6 investigates other confounds, including those related to funds' parent companies and the possibility that our structural complexity measures are driven by funds catering to investors' preferences about horizon and customer service.

6. Additional Analyses and Discussion

This section summarizes additional descriptive analyses and discussions. The Online Appendix contains additional details and robustness tests.

6.1 Do Retail Investors Actually Read Mutual Fund Prospectuses?

An assumption underlying Carlin (2009) and our empirical tests is that retail investors attempt to learn from fund disclosures. Our Online Appendix (OA 8) discusses evidence to support this assumption, and here we provide a brief overview. First, existing studies find that retail investors attempt to learn from fund disclosures, and the primary barrier is that disclosures are difficult to understand (e.g., SEC 2012; Alexander et al. 1998; Investment Company Institute 2006). Second, the SEC's ongoing focus on improving disclosure readability indicates that regulators believe that investors want to learn from prospectuses. Third, many studies find evidence that retail investors read corporate 10-Q/K filings (e.g., Miller 2010; Lawrence 2013), so it seems likely that many retail investors read at least the summary prospectuses of mutual funds that comprise the majority of savings. Finally, because investment websites frequently copy text directly from prospectuses, investors can be affected by narrative complexity even if they do not read the prospectus itself.[25]

6.2 Are High Fees Actually Discretionary?

Here we investigate Carlin's (2009) assumption that managers choose to set fees above marginal costs to extract rents from investors. The null is that the complexity we observe is driven by funds trying to obfuscate high fees that are beyond their control.[26]

Our approach is to decompose *Fees* into its expected and excess components, and then test whether the excess component is positively associated with obfuscation. The assumption is that the expected component of fees is due to fundamentals, and the excess component is discretionary. If managers use complexity to obfuscate discretionarily high fees, then we should

[25] Analyses in OA8 of our Online Appendix compare the objective statements on four leading brokerage websites to the funds' prospectuses. When the brokerage offers the fund, the objective statements are direct excerpts in nearly all cases. The only exception is the omission of a parenthetical statement for one fund.

[26] While finding evidence of discretionarily high fees would further support Carlin (2009), failing to find evidence of discretionary fees would not invalidate our inferences that fund managers aim to obfuscate high fees with complex disclosures. That is, even if high fees are non-discretionary, narrative and structural complexity are still discretionary and can be used by managers to obfuscate high fees.

observe a positive association between *Complexity* and excess fees.[27]

Operating costs are likely the primary driver of funds' non-discretionary fees; i.e., funds with higher costs must charge higher fees to cover those costs. Elton et al. (2019) explain that two primary determinants of operating costs for index funds are fund size and turnover. Larger funds can spread fixed costs over more assets under management, resulting in lower average costs. Funds that must trade more often to track the S&P 500 index (i.e., have higher turnover) are less efficient and have higher operating costs. Thus, we model non-discretionary fees based on fund size and turnover and include tracking error for completeness. A regression of *Fees* on these determinants generates an r-squared of 31%.

We test the association between excess fees and complexity using a single-step regression to avoid biases from two-step procedures (Chen et al. 2018):

$$Fees_{i,y} = \beta_1 Complexity_{i,y} + \beta_2 Size_{i,y} + \beta_3 Turnover_{i,y} + \beta_4 TE_{i,y} + \Sigma\beta_y Year_y + \varepsilon_{i,y} \quad (2)$$

A caveat to model (2) is that size is endogenous in Carlin (2009), such that low-fee funds should be larger than high-fee funds. Using size to model non-discretionary expenses likely over-controls the model and biases β_1 toward zero.[28]

With this caveat, Table 7 tabulates results of model (2). In general, the narrative complexity coefficients are smaller after including controls, but most remain significant in Panel B. Table 8 investigates excess fees and structural complexity and again finds smaller but significant results. Overall, these results support Carlin's (2009) theory that funds use complexity

[27] This approach is similar to studies of abnormal accruals (e.g., Dechow et al. 2011) or excess compensation (e.g., Core et al. 2008).

[28] Excluding fund size from model (1) should not cause omitted variable bias. To generate omitted variable bias, fund size must affect fees and correlate with complexity for reasons other than the fund's pricing strategy. While fund size likely affects fees, there is little reason to think that fund size should necessarily affect fund managers' choices about *Repetition* or *FundsinFiling* in model (1), nor should fund size affect choices within specific disclosures (Section 5.2).

to obfuscate discretionarily high fees.

6.3 Pre/Post-2009 SEC Regulation Changes

Tests in the Online Appendix (OA 7) investigate the effects of 2009 regulations that introduced summary prospectuses and likely lowered investors' costs of learning from disclosures. They also likely motivated funds to improve their disclosures. Weaknesses of these tests are that the regulations coincided with the financial crisis and the pre-2009 trends may not be parallel, so our results should be interpreted with caution.

First, we find that, relative to low-complexity funds, high-complexity funds improved the narrative complexity of their prospectuses in the post-regulation period. These findings are consistent with the regulations emphasizing the SEC's increased focus on narrative complexity and motivating funds to improve their disclosures. Another important takeaway from these results is to help corroborate that our measures of narrative complexity are discretionary.

Second, a downward shock to narrative complexity should cause more investors to become informed and, all else equal, increase the market share captured by funds that follow a low-fee/complexity strategy. Assuming there are frictions to immediately switching strategies, we expect that high-fee/complexity funds lose market share in the short term. We find that, relative to low-complexity funds, high-complexity funds shrink and are more likely to close in the post-regulation period. These findings are consistent with the regulations improving narrative complexity and reducing rent extraction by high-complexity funds.

6.4 Do High-Fee Fund Parent Institutions Offer More Fund Options or Financial Services?

We investigate the breadth of products offered by parent institutions to address four concerns. First, some S&P 500 funds might charge high fees because customers like that the parent companies offer a wider variety of funds and products. Second, some S&P 500 funds

might charge high fees because they bundle or cross-sell funds with other products. Third, the parents of high-fee funds could be more complex institutions, and somehow this institutional complexity affects the complexity of fund prospectuses. Fourth, *FundsinFiling* might be higher for high-fee funds simply because the parent companies offer more funds.

We identify 8,154 funds issued by the parents in our sample over 2000 – 2017. We start in 2000 because that is when CRSP began providing consistent parent identifiers. Sample construction procedures are detailed in our Online Appendix (OA 2). We calculate *Fees* and other variables for each fund-year and then average each measure to the parent-year level.

Table 9 finds that the parents of low-fee S&P 500 funds tend to issue funds that are cheaper (row i) and larger (row ii) than the parents of high-fee S&P 500 funds. Rows (iv) through (vi) of Table 9 find that the parents of high-fee S&P 500 funds offer *fewer* passive funds and a similar number of active funds. Row (vii) finds that high-fee S&P 500 fund parents also offer a smaller variety of fund categories (e.g., large-cap and small-cap).

Row (viii) summarizes how many other (non-mutual fund) retail financial products are offered by parent companies. As detailed in the Online Appendix (Table OA2), we consider seven standard products: bank accounts; mortgages and loans; credit cards; life and long-term care insurance; auto and property insurance; trading and brokerage platforms; and ETFs. This test uses only one observation per parent company of funds in our post-2010 sample. We find that the parents of low-and high-fee S&P 500 funds offer a similar variety of financial products.

Together, these results are inconsistent with concerns related to parent product offerings.

6.5 Is High-Fee & High-Complexity a Parent-Level Strategy?

This section investigates whether the parent companies of S&P 500 funds follow a consistent obfuscation strategy for all of their funds. Ex ante, we suspect that funds' strategies

are determined at the parent-level due to economies of scale and synergies across funds, and because combining multiple funds into one prospectus requires parent-level coordination.

Parent-level tests are not our primary analyses because they have several weaknesses relative to our tests of S&P 500 funds. First, they pool a wide variety of funds and therefore lack the identification strengths of examining homogeneous index funds. Second, *FundsinFiling* and *Repetition* require manual data collection, which is impracticable for a large sample. Third, it is impracticable to manually verify our matches of funds to prospectuses like we do for our main sample, so our other narrative complexity measures are noisier.

We investigate the association between the narrative complexity of S&P 500 funds and their parent-level averages using the following regression:

$$Parent\ Complexity_{i,y} = \beta_1\ S\&P500\ Fund\ Complexity_{i,y} + \Sigma\beta_y\ Year_y + \varepsilon_{i,y} \quad (3)$$

Columns (i) and (ii) of Table 10 Panel A investigate narrative complexity in the full prospectus, and columns (iii) and (iv) investigate narrative complexity in the expense disclosure. These results indicate that S&P 500 fund *Length* and *WordsPerSentence* are highly correlated with parent-level averages, consistent with narrative complexity being a parent-level strategy. Similar results are observed in Panel B for structural complexity.

We next examine the association between parent-level fees and complexity. Table 10 Panel C examines narrative complexity and finds results that are similar to those for S&P 500 funds. Specifically, columns (i) through (iv) find no association between *Fees* and *Length* or *WordsPerSentence* in the full prospectuses, but columns (v) through (viii) find significantly positive associations in the expense disclosures. Panel D finds consistently positive associations between parent-level fees and parent-level structural complexity. Together, the results in Table 10 indicate that the high-fee and high-complexity strategy we observe for S&P 500 funds is

likely a parent-level choice. These results also suggest that our findings for S&P 500 funds likely generalize to the broader mutual fund market.

6.6 Analysis of Complementary Marketing & Customer Service Strategies

For two reasons, aggressive marketing is a natural complement to obfuscation; i.e., is an additional outcome of high-fee fund managers' strategy choice (see Figure 1). First, marketing efforts are likely more effective when high processing costs inhibit investors' independent research. Second, since investors ultimately foot the bill for fund's marketing efforts, funds can perform more marketing when investors are uninformed about fees.[29]

The existence of a complementary marketing strategy does not, by itself, undermine our main inferences that fund managers manipulate narrative and structural complexity as part of an obfuscation strategy. However, because marketing costs are included in *Fees*, aggressive marketing could pose a validity threat in model (1) if marketing efforts also directly affect our empirical proxies for complexity. We argue that there is little reason to think that marketing affects most of our narrative complexity proxies, especially in our tests of fund objective and equity risk disclosures. Similarly, it seems unlikely that marketing fully drives our structural complexity proxies.

For descriptive purposes, analyses in our Online Appendix (OA 6) investigate the role of marketing for high-fee funds. We replace *Fees* in model (1) with our estimates of funds' marketing efforts, and find significantly positive associations with most of our measures of both narrative and structural complexity. These results are consistent with complexity and marketing being complements. We also remove marketing fees from *Fees* in model (1), and find results

[29] Studies provide somewhat mixed evidence on whether fund marketing impairs investment decisions (e.g., Jain & Wu 2000; Huhmann & Bhattacharya 2005; Gallaher et al. 2015; Sirri & Tufano 1998; Khorana & Servaes 2012).

similar to those in Tables 4 and 5.[30] Because customer service fees are included in our estimate of marketing fees, these results also reduce concerns that investors buy high-fee S&P 500 funds because they provide better service.[31]

6.7 Analysis of Investor Horizon Preferences

Analyses in this section address the concern that structural complexity, rather than obfuscating fees, is used to provide savings over certain investment horizons. Different horizon preferences could justify multi-class funds; e.g., Appendix B shows that Deutsche's Class A estimated costs are higher than Class C's over one year but lower over 10 years. However, if all of the classes of high-fee multi-class funds are more expensive than those of low-fee funds, then structural complexity would not provide savings over certain horizons. For example, Deutsche's Classes A and C are both far more expensive than Schwab's single class in Appendix B, regardless of the holding period.

We provide two results that are inconsistent with structural complexity providing savings over certain horizons. First, Table 1 Panel B provides fee information after dividing the sample into low- and high-fee funds by year. Row (i) shows that the most expensive class within each low-fee fund has *Fees* of 16.4 bps, while row (ii) shows that the least expensive class within each high-fee fund has *Fees* of 55.5 bps. Thus, the cheapest classes from high-fee funds charge an average of 39.1 bps more than the most expensive classes from low-fee funds, and tests in our Online Appendix (OA 10) find that this difference is statistically significant. Second, results in

[30] We do not exclude marketing fees in our main *Fees* measure because it funds likely use narrative and structural complexity to obfuscate the fact that investors must pay these fees. Including marketing fees in *Fees* is standard within the literature.

[31] This finding is consistent with prior literature that does not find support for the hypothesis that high-fee funds offer extra benefits such as helpful investment advice or customer service (Anagol & Kim 2012). See Jain & Wu (2000), Cronqvist (2006), and Gallaher et al. (2015) for additional studies of mutual fund advertising.

the Online Appendix (Tables OA14 and OA15) calculate *Fees* based on three- and five-year holding periods, and find results that are very similar to our main results.

6.8 Discussion of Other Unobservable Variation across S&P 500 Index Funds

Our setting and analyses hold constant many non-discretionary factors that affect complexity. However, unobservable differences across funds still exist, and readers should consider potential confounds when evaluating our findings.

One unobserved source of variation is other price dispersion strategies, such as market segmentation whereby fund classes are only available from specific institutions, or the effects of media coverage (e.g., Hortaçsu & Syverson 2004; Solomon et al. 2014). Funds' exploitation of these frictions may be a complementary strategy to obfuscation, similar to what we observe for marketing. While we are unaware of specific reasons why these complementary strategies should confound our complexity measures, the possibility remains.

Another unobserved source of variation is securities lending. As discussed in our Online Appendix (OA 9), research finds little evidence that lending affects the fees or tracking errors of S&P 500 funds, nor should lending affect most of our complexity measures.

A final unobserved source of variation is manager effort: lazy managers might put little effort into reducing fees and constructing clear disclosures. There are several reasons why this explanation seems less likely than strategic obfuscation. First, while Carlin (2009) shows that obfuscation allows high-fee funds to exist in a competitive market, we are unaware of theory that shows how laziness would allow such an equilibrium to persist. Second, the costs of reducing narrative complexity seem small relative to the size of S&P 500 funds and to the potential cost of SEC enforcement for not complying with disclosure regulations.[32] Finally, it is plausible that it

[32] For example, while not specific to prospectuses, the SEC recently fined mutual fund issuers (including Deutsche) over $125M for failing to adequately disclose 12b-1 fees (Wall Street Journal 2019b)

takes more effort to run a structurally complex fund than a simple fund. Still, we cannot rule out this alternate explanation.

7. Summary and Concluding Discussion

We examine "strategic obfuscation" in mutual funds; i.e., whether fund managers attempt to obfuscate unfavorable information via unnecessarily complex disclosures and fee structures. Our tests examine homogeneous S&P 500 index funds and use a within-year research design to hold constant many non-discretionary drivers of complexity. Consistent with theory in Carlin (2009), we find that funds with higher fees have greater narrative complexity (i.e., less readable disclosures) and structural complexity (i.e., more complicated fee structures), both of which increase investors' processing costs. These findings are consistent with funds attempting to use complexity to obfuscate high fees and extract rents from retail investors. That we find obfuscation and excessive fees among homogeneous S&P 500 index funds is especially striking because their disclosures are heavily regulated, and because conventional wisdom is that index funds are a cheap way to obtain a diverse portfolio.[33]

Section 1.5 discusses the implications of our findings for the mutual funds literature. We conclude by discussing our implications for regulations and the corporate disclosure literature.

7.1. Regulatory implications

Our study should inform ongoing regulatory efforts to improve mutual fund Disclosures. Tucker and Xia (2020) find that mutual funds use complex language, despite the SEC's plain English initiatives. Although fund issuers argue that variation in prospectus readability is driven by innate factors, our results indicate that unreadable disclosures are part of

[33] As an example from Warren Buffett: "What I advise here is essentially identical to certain instructions I've laid out in my will […] Put 10% of the cash in short-term government bonds and 90% in a very low-cost S&P 500 index fund" (Berkshire Hathaway 2013).

a strategy to extract rents from retail investors. These findings should be of interest to the SEC as it develops and enforces regulations to improve fund disclosures (SEC 2020). In particular, the SEC should perhaps revisit its decisions not to limit multiple-fund prospectuses or the length of summary sections (SEC 2014).

Our findings also raise questions about whether the SEC should tighten regulations on complex share classes and fee structures. While funds argue that multiple share classes with different fees help them cater to a range of clients, we find that structurally complex funds are consistently more expensive than simpler funds. In particular, our results are consistent with concerns that funds use 12b-1 fees, loads, and load breaks to extract rents from investors. While the SEC has raised concerns about these fees and prosecuted advisors for lack of transparency, it has so far not prohibited them (e.g., Armstrong & Vickers 2012; Wall Street Journal 2019b; SEC 2019).

Finally, our findings are relevant to the SEC's 2019 "Regulation Best Interest," which requires advisors to only recommend financial products that are in a client's best interest. Section II.C.2 of the Regulation substantially limits advisors' responsibility when making recommendations across mutual funds from different companies. Given that we find large differences in fees across S&P 500 funds, the SEC should perhaps consider raising advisors' responsibilities at least for highly similar index funds.

To be clear, any changes in regulations or enforcement require cost-benefit analyses that are beyond the scope of this paper. Still, our results should be valuable inputs to those analyses.

7.2. Insights for the corporate disclosure literature

Our findings provide several insights relevant to studies on corporate disclosure readability (e.g., Li 2008; Guay et al. 2016; Bonsall et al. 2017; Dyer et al. 2017). First, we

provide new evidence that managers manipulate narrative complexity to obfuscate weak performance. Blankespoor et al. (2020, p92) conclude that isolating disclosure choices from non-discretionary characteristics is a fundamental challenge, and that "findings about whether disclosure characteristics are strategic versus non-discretionary remain mixed" (also see Leuz & Wysocki 2016). We use a novel approach to address non-discretionary complexity, and our findings help triangulate and support existing evidence of obfuscation. Given that fund managers use narrative complexity to obfuscate performance for entities as simple as index funds, it seems highly plausible that managers use narrative complexity to obfuscate poor firm performance.

Second, corporate disclosure studies often treat structural complexity as non-discretionary and control for it when analyzing narrative complexity (e.g., controlling for reporting segments). However, our results show that funds use narrative and structural complexity together. It seems plausible that corporate managers also manipulate structural and narrative complexity as two parts of the same obfuscation strategy. Future research can endeavor to investigate whether the corporate structural and narrative complexity are jointly determined, and their joint effects on market outcomes.

Finally, our study has potential implications for the corporate opacity literature. Similar to index funds, it seems plausible that corporate managers partially choose the poor performance they obfuscate.[34] For example, a manager may shirk and obfuscate the impact on performance via unclear disclosures.[35] If performance and complexity are jointly determined in mutual funds,

[34] Many studies on corporate opacity are silent about whether poor performance is discretionary (e.g., Li 2008). Merkley (2014) finds that longer disclosures help inform investors about poor performance, indicating that performance is non-discretionary. Guay et al. (2016) find that financial statement complexity is non-discretionary, implying that any associated poor performance is also non-discretionary. Performance is explicitly exogenous in Asay et al. (2018). Lo et al. (2017) find that corporate managers use complexity to mask earnings management, providing some evidence that performance and obfuscation are codetermined.

[35] Studies have shown that disclosures can serve as a governance mechanism and reduce agency problems (Bushman & Smith 2001). Disclosure quality is related to effective governance and investment efficiency (Karamanou & Vafeas 2005; Biddle et al. 2009; Cheng et al. 2013). Also see Bebchuk & Fried (2003).

it seems likely that they are joint outcomes of corporate managers' strategies.

References

Adams, J. C., Mansi, S. A., & Nishikawa, T. (2012). "Are mutual fund fees excessive?" *Journal of Banking & Finance* 36: 2245-2259.

Agarwal, V., Mullally, K. A., Tang, Y., & Yang, B. (2015). "Mandatory portfolio disclosure, stock liquidity, and mutual fund performance." *The Journal of Finance* 70: 2733-2776.

Alchian, A. A. (1950). "Uncertainty, evolution, and economic theory." *Journal of Political Economy* 58: 211-221.

Alexander, G., Jones, J., & Nigro, P. (1998). "Mutual fund shareholders: characteristics, investor knowledge, and sources of information." *Financial Services Review* 7: 301-316.

Anagol, S., & Kim, H. (2012). "The impact of shrouded fees: Evidence from a natural experiment in the Indian mutual funds market." *American Economic Review* 102: 576-593.

Armstrong, M., & Vickers, J. (2012). "Consumer protection and contingent charges." *Journal of Economic Literature* 50: 477-93.

Asay, H. S., Libby, R., & Rennekamp, K. M. (2018). "Do features that associate managers with a message magnify investors' reactions to narrative disclosures?" *Accounting, Organizations and Society* 68: 1-14.

Barber, B., Odean, T., & Zheng, L. (2005). "Out of sight, out of mind: The effects of expenses on mutual fund flows." *Journal of Business* 78: 2095-2120.

Bebchuk, L., & Fried, J. (2003). "Executive compensation as an agency problem." *Journal of Economic Perspectives* 17: 71-92.

Benartzi, S., & Thaler, R. H. (2001). "Naive diversification strategies in defined contribution saving plans." *American Economic Review* 91: 79-98.

Berkshire Hathaway. (2013). Shareholder letter. Available at: https://www.berkshirehathaway.com/letters/letters.html. Last retrieved February 10, 2020.

Beshears, J., Choi, J., Laibson, D., & Madrian. B. C. (2009) "How does simplified disclosure affect individuals' mutual fund choices?" In D. A. Wise, ed. *Explorations in the Economics of Aging*. Chicago: University of Chicago Press.

Biddle, G. C., Hilary, G., & Verdi, R. S. (2009). "How does financial reporting quality relate to investment efficiency?" *Journal of Accounting and Economics* 48: 112-131.

Blankespoor, B., deHaan, E., & Marinovic, I. (2020). "Disclosure processing costs, investors' information choice, and equity market outcomes: A review." *Journal of Accounting and Economics*, forthcoming.

Bonsall IV, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics* 63: 329-357.

Bushman, R. M., & Smith, A. J. (2001). "Financial accounting information and corporate governance." *Journal of Accounting and Economics* 32: 237-333.

Carlin, B. I. (2009). "Strategic price complexity in retail financial markets." *Journal of Financial Economics* 91: 278-287.

Chan, L. K., Chen, H. L., & Lakonishok, J. (2002). On mutual fund investment styles. *The Review of Financial Studies* 15: 1407-1437.

Chen, H., Cohen, L., & Gurun, U. (2020). "Don't Take Their Word For It: The Misclassification of Bond Mutual Funds." *Working paper*, available at SSRN: https://ssrn.com/abstract=3474557.

Chen, W., Hribar, P., & Melessa, S. (2018). "Incorrect inferences when using residuals as dependent variables." *Journal of Accounting Research* 56: 751-796

Cheng, M., Dhaliwal, D., & Zhang, Y. (2013). "Does investment efficiency improve after the disclosure of material weaknesses in internal control over financial reporting?" *Journal of Accounting and Economics* 56: 1-18.

Choi, 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: 1405-1432.

Core, J. E., Guay, W., & Larcker, D. F. (2008). "The power of the pen and executive compensation." *Journal of Financial Economics* 88: 1-25.

Cremers, K. M., & Petajisto, A. (2009). "How active is your fund manager? A new measure that predicts performance." *The Review of Financial Studies* 22: 3329-3365.

Cronqvist, H. (2006). "Advertising and portfolio choice." *Working paper*, available at SSRN: https://ssrn.com/abstract=920693.

Dechow, P., Ge, W., Larson, C.R., & Sloan, R.G. (2011). "Predicting material accounting restatements." *Contemporary Accounting Research* 28: 17-82.

Del Guercio, D., & Reuter, J. (2014). "Mutual fund performance and the incentive to generate alpha." *Journal of Finance* 69: 1673-1704.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). "The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation." *Journal of Accounting and Economics* 64: 221-245.

Edelen, R. M., Evans, R. B., & Kadlec, G. B. (2012). "Disclosure and agency conflict: Evidence from mutual fund commission bundling." *Journal of Financial Economics* 103: 308-326.

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

Elton, E.J., Gruber, M., & de Souza, A. (2019). "Passive mutual funds and ETFs: Performance and comparison." *Journal of Banking & Finance* 106: 265-275.

Evans, R. B., & Fahlenbrach, R. (2012). "Institutional investors and mutual fund governance: Evidence from retail–institutional fund twins." *The Review of Financial Studies* 25: 3530-3571.

Fama, E.F., & French, K.R. (2010). "Luck versus skill in the cross-section of mutual fund returns." *Journal of Finance* 65: 1915-1947.

Ferris, S. & Chance, D. (1987). "The Effect of 12b-1 Plans on Mutual Fund Expense Ratios." *Journal of Finance* 42: 1077-1082.

Frazzini, A., & Lamont, O. A. (2008). "Dumb money: Mutual fund flows and the cross-section of stock returns." *Journal of Financial Economics* 88: 299-322.

French, K. (2008). "Presidential Address: The Cost of Active Investing." *The Journal of Finance* 63: 1537-1573.

Gallaher, S. T., Kaniel, R., & Starks, L. T. (2015). "Advertising and mutual funds: From families to individual funds." *Working Paper*, available at SSRN: https://ssrn.com/abstract=1362535.

Ge, W., & Zheng, L. (2006). "The frequency of mutual fund portfolio disclosure." *Working Paper*, available at SSRN: https://ssrn.com/abstract =557186.

Gil-Bazo, J., & Ruiz-Verdú, P. (2009). "The relation between price and performance in the mutual fund industry." *Journal of Finance* 64: 2153-2183

Goetzmann, W.N., & Peles, N. (1997). "Cognitive Dissonance and Mutual Fund Investors." *The Journal of Financial Research* XX: 145-158.

Gruber, M.J. (1996). "Another puzzle: the growth in actively managed mutual funds." *Journal of Finance* 51: 783-810.

Guay, W., Samuels, D., & Taylor, D. (2016). "Guiding through the fog: Financial statement complexity and voluntary disclosure." *Journal of Accounting and Economics* 62: 234-269

Herman, E. (1963). "Mutual fund management fee rates." *The Journal of Finance* 18: 360-376.

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

Huberman, G., & Jiang, W., (2006). "Offering versus choice in 401(k) Plans: Equity exposure and number of funds." *The Journal of Finance* 61: 763-801.

Huhmann, B. A., & Bhattacharyya, N. (2005). "Does mutual fund advertising provide necessary investment information?" *International Journal of Bank Marketing* 23: 296-316.

Investment Company Institute (2006). "Understanding investor preferences for mutual fund information." Available at: https://www.idc.org/pdf/rpt_06_inv_prefs_full.pdf. Last retrieved March 23, 2019.

Investment Company Institute (2018). "Investment company fact book." Available at: https://www.ici.org/pdf/2018_factbook.pdf. Last retrieved March 23, 2019.

Investment Company Institute (2020). "Investment company fact book." Available at: https://www.ici.org/pdf/2020_factbook.pdf. Last retrieved September 15, 2020.

Jensen, M.C. (1968). "The performance of mutual funds in the period 1945-1964." *Journal of Finance* 23: 389-416.

Jain, P. C. & Wu, J. S. (2000). "Truth in mutual fund advertising: Evidence on future performance and fund flows." *The Journal of Finance* 55: 937-958.

Karamanou, I., & Vafeas, N. (2005). "The association between corporate boards, audit committees, and management earnings forecasts: An empirical analysis." *Journal of Accounting Research* 43: 453-486

Khorana, A., & Servaes, H. (2012). "What drives market share in the mutual fund industry?" *Review of Finance* 16: 81-113.

Lawrence, A. (2013). "Individual investors and financial disclosure." *Journal of Accounting and Economics* 56: 130-147.

Leone, A. J., Minutti-Meza, M., & Wasley, C.E. (2019). "Influential observations and inference in accounting research." *The Accounting Review* 94: 337-364.

Leuz, C., & Wysocki, P.D. (2016). "The economics of disclosure and financial reporting regulation: evidence and suggestions for future research" *Journal of Accounting Research* 54: 525-622.

Li, F. (2008). "Annual report readability, current earnings, and earnings persistence." *Journal of Accounting and Economics* 45: 221-247.

Lo, K., Ramos, F., & Rogo, R. (2017). "Earnings management and annual report readability." *Journal of Accounting and Economics* 63: 1-25

Loughran, T., & McDonald, B. (2014). "Measuring readability in financial disclosures." *Journal of Finance* 69: 1643-1.

Loughran, T., & McDonald, B. (2016). "Textual analysis in accounting and finance: A survey." *Journal of Accounting Research* 54: 187-230.

Malkiel, B.G. (1995). "Returns from investing in equity mutual funds 1971 to 1991." *Journal of Finance* 50: 549-572.

Merkley, K. J. (2014). "Narrative disclosure and earnings performance: Evidence from R&D disclosures." *The Accounting Review* 89: 725-757.

Miller, B. P. (2010). "The effects of reporting complexity on small and large investor trading." *The Accounting Review* 85: 2107-2143.

NASD (2003). "NASD Investor Literacy Research." Available at: https://www.finra.org/sites/default/files/InvestorDocument/p011459.pdf. Last retrieved June 3, 2019.

Philpot, J., & Johnson, D. T. (2007) "Mutual fund performance and fund prospectus clarity." *Journal of Financial Services Marketing* 11: 211-216.

Securities and Exchange Commission (1998). "A plain English handbook: How to create clear SEC disclosure documents." Available at: https://www.sec.gov/pdf/handbook.pdf. Last retrieved January 23, 2020.

Securities and Exchange Commission (2005). "Rulemaking for EDGAR system." Available at: https://www.sec.gov/rules/final/33-8590.pdf. Last retrieved April 17, 2019.

Securities and Exchange Commission (2009a). "Enhanced disclosure and new prospectus delivery option for registered open-end management investment companies." Available at: https://www.sec.gov/rules/final/2009/33-8998.pdf. Last retrieved March 15, 2019.

Securities and Exchange Commission (2009b). "Interactive data for mutual fund risk/return summary." Available at: https://www.sec.gov/rules/final/2009/33-9006.pdf. Last retrieved March 15, 2019.

Securities and Exchange Commission (2012). "Study regarding financial literacy among investors." https://www.sec.gov/news/studies/2012/917-financial-literacy-study-part1.pdf. Last retrieved April 9, 2019.

Securities and Exchange Commission (2014). "Guidance update: Guidance regarding mutual fund enhanced disclosure." Available at: https://www.sec.gov/investment/im-guidance-2014-08.pdf. Last retrieved March 11, 2019.

Securities and Exchange Commission (2018). "SEC modernizes the delivery of fund reports and seeks public feedback on improving fund disclosure." Available at: https://www.sec.gov/news/press-release/2018-103. Last retrieved March 23, 2019.

Securities and Exchange Commission (2019). "Regulation Best Interest: The Broker-Dealer Standard of Conduct." Available at: https://www.sec.gov/rules/final/2019/34-86031.pdf. Last retrieved May 29, 2019.

Securities and Exchange Commission (2020). "Proposed Rule: Tailored Shareholder Reports, Treatment of Annual Prospectus Updates for Existing Investors, and Improved Fee and Risk Disclosure for Mutual Funds and Exchange-Traded Funds; Fee Information in Investment Company Advertisements." Available at: https://www.sec.gov/rules/proposed/2020/33-10814.pdf. Last retrieved August 29, 2020.

Sirri E.R., & Tufano, P. (1998). "Costly search and mutual fund flows," *Journal of Finance* 53: 1589-1622.

Solomon, D., Soltes, H. E., & Sosyura, D. (2014). "Winners in the spotlight: Media coverage of fund holdings as a driver of flows." *Journal of Financial Economics* 113: 53-72.

Tucker, A. M., & Xia, Y. (2020). "The promises & perils of plain English mutual fund disclosures." *Working Paper*, available at SSRN: https://ssrn.com/abstract=3436952.

Wall Street Journal (2019a). "Dozens of advisers face claims of overcharging for mutual funds." Available at: http://www.wsj.com/articles/dozens-of-advisers-face-claims-of-overcharging-for-mutual-funds-11548936000. Last retrieved March 15, 2019

Wall Street Journal (2019b). "Firms to pay $125 million to clients over fee-disclosure practices." Available at: http://www.wsj.com/articles/firms-to-pay-125-million-to-clients-over-fee-disclosure-practices-11552335611. Last retrieved March 15, 2019

Wermers, R. (2012). "Matter of Style: The Causes and Consequences of Style Drift in Institutional Portfolios." *Working paper*, available at SSRN: https://ssrn.com/abstract=2024259.

You, H., & Zhang, X. J. (2009). "Financial reporting complexity and investor underreaction to 10-K information." *Review of Accounting Studies* 14: 559-586.

Appendix A: Variable Definitions

This table provides variable definitions and data sources. All variables are constructed to be on a fund-calendar year basis. Unless otherwise noted, variables based on monthly data are first calculated on a fund-month basis and then averaged to the annual level. Continuous variables are winsorized at 1% and 99%. To facilitate comparisons across coefficients, all variables are standardized to have a mean (standard deviation) of zero (one) when used as independent variables in regression tests.

Variable Name	Description	Source
Fees	Total annual ownership cost charged to the fund's retail investors in basis points, multiplied by 100 to be in percentage points. The total annual ownership cost includes annual fees, the annualized rear load (including contingent deferred sales charges), and the annualized front load. The annual expense ratio includes 12b-1 fees and may include waivers and reimbursements. We use an expected seven-year holding period to determine and amortize any front and rear loads. We use the maximum front load for classes with tiered front loads. For funds with multiple classes, we use the maximum cost across retail share classes. Robustness tests using alternate *Fees* assumptions are discussed in the Online Appendix (OA 4).	CRSP
FrontLoad	Indicator variable set to 1 if any share class of the fund, in a given year, has a front load.	CRSP
FrontLoadBreaks	Number of breakpoints in the front load. Calculated at the fund level as the maximum number of breakpoints across share classes.	CRSP
FundsinFiling	Log of the number of funds in the Form 485 filing. This measure is only available post-2006, when the SEC required investment company filers to electronically identify the separate funds in their filings.	SEC Edgar filings
Length	First principal component of combining the word count and character count from the filing. Standardized to have a mean (standard deviation) of 0 (1).	SEC Analytics Suite
Length_ExpDisc	First principal component of combining the word count and the character count from the expense disclosure within the prospectus summary disclosure. Calculated at the annual level as the mean for each expense disclosure obtained from the tag "ExpenseNarrativeTextBlock" in the TXT files of the SEC Mutual Fund Prospectus Risk/Return Summary Data Sets for Form 485 filings related to the fund over the year. Standardized to have a mean (standard deviation) of 0 (1).	SEC XBRL submissions
Length_ObjDisc	First principal component of combining the word count and the character count from objective disclosure within the prospectus summary disclosure. Calculated at the annual level as the mean for each objective disclosure obtained from the tag "ObjectivePrimaryTextBlock" in the TXT files of the SEC Mutual Fund Prospectus Risk/Return Summary Data Sets for Form 485 filings related to the fund over the year. Standardized to have a mean (standard deviation) of 0 (1).	SEC XBRL submissions
Length_RiskDisc	First principal component of combining the word count and the character count from the equity risk disclosure within the prospectus summary disclosure. Standardized to have a mean (standard deviation) of 0 (1).	Mutual fund prospectuses
NoLoad_12b1	Indicator variable set to 1 if any share class of the fund, in a given year, has no front load but has a 12b-1 fee.	CRSP
RearLoadBreaks	Number of breakpoints in the rear load, also known as the Contingent Deferred Sales Charge. Calculated at the fund level as the maximum number of breakpoints across share classes.	CRSP

Repetition	Fraction of sentences in the summary section of Form 485 that are repeated in the rest of the document. A sentence is coded as "repeated" if the cosine similarity between it and any of the sentences in the rest of the document is 90% or greater (Merkley 2014). *Repetition* is available after 2010 when summary sections with XBRL tagging were first required.	SEC XBRL submissions, SEC Edgar filings
ShareClasses	The number of unique share classes of the fund.	CRSP
Size	Log of the fund's total net assets (i.e., assets under management) in millions, inclusive of all retail share classes. First, calculated at the monthly level as the summed net assets across all classes. Then, calculated at the annual level as the mean of the monthly sums. Missing share class level net assets are set to 0.	CRSP
Structural_Complexity	First principal component of combining the following five variables: *ShareClasses*, *FrontLoadBreaks*, *RearLoadBreaks*, *FrontLoad*, and *NoLoad_12b1*. Standardized to have a mean (standard deviation) of 0 (1).	CRSP, SEC Analytics Suite
Summarization	Log of 1 plus the difference between the character count of the equity risk disclosure in the details of the prospectus and the character count of the equity risk disclosure in the summary section of the prospectus.	Mutual fund prospectuses
Summarization_Zero	Indicator variable set to 1 if *Summarization* is 0.	Mutual fund prospectuses
TE	Annual tracking error, in basis points. Consistent with the common academic definition (e.g., Cremers and Petajisto 2009), the steps to calculate *TE* are as follows: 1) calculate the gross monthly return (return + 1/12 of the annual expense ratio); 2) calculate the monthly difference between the gross return and the S&P 500 index return; 3) calculate the standard deviation of the differences within each calendar year.	CRSP, Bloomberg
Turnover	Minimum of aggregated sales or aggregated purchases of securities, divided by the average 12-month Total Net Assets of the fund.	CRSP
Wordiness	First principal component of combining *WordsPerSentence* and *Length* as provided by SEC Analytics Suite. Standardized to have a mean (standard deviation) of 0 (1).	SEC Analytics Suite
Wordiness_ExpDisc	First principal component of combining *WordsPerSent_ExpDisc* and *Length_ExpDisc*, calculated just for the expense disclosure within the prospectus summary disclosure. We calculate *WordsPerSent_ExpDisc* and *Length_ExpDisc* ourselves because SEC Analytics Suite does not have measures for filing subsections.	SEC XBRL submissions
WordsPerSentence	Average number of words per sentence, calculated as the number of words in the prospectus divided by the total number of sentence termination characters after removing those associated with headings and abbreviations.	SEC Analytics Suite
WordsPerSent_ExpDisc	Average number of words per sentence specifically in the expense disclosure of the prospectus summary section. Calculated at the annual level as the mean of this value for each expense disclosure obtained from the tag "ExpenseNarrativeTextBlock" in the TXT files of the SEC Mutual Fund Prospectus Risk/Return Summary Data Sets for Form 485 filings related to the fund over the year.	SEC XBRL submissions

Appendix B: Mutual Fund Disclosure Examples

The examples below present the first two items from the summary sections of the prospectuses of two S&P 500 index funds in 2019. As required by the SEC, the first two items in each prospectus contain summary information on the fund objectives and fees. Example 1 from Schwab has a single class, a single annual fee, no waivers or contingencies, and no loads. Example 2 from Deutsche has multiple classes with various combinations of fees and expenses. Our Online Appendix (Appendix OA 1) presents seven additional pages of information from Deutsche's prospectus that are needed to understand the classes and expenses summarized in Example 2. Schwab's and Deutsche's gross returns in 2019 were 31.46% and 31.47%, respectively, compared to the S&P 500 gross return of roughly 31.48%.

These examples are from 2019, which is outside of our sample period. We provide example disclosures for 2019 because historical disclosures are only available in raw text format and less useful for illustrative purposes. While the specific numbers are slightly different in 2017, the classes, fee structures, and inferences are the same.

Example 1: Excerpts from Schwab's S&P 500 index fund prospectus

Schwab® S&P 500 Index Fund

Ticker Symbol: SWPPX

Investment Objective

The fund's goal is to track the total return of the S&P 500® Index.

Fund Fees and Expenses

This table describes the fees and expenses you may pay if you buy and hold shares of the fund. This table does not reflect any brokerage fees or commissions you may incur when buying or selling fund shares.

Shareholder Fees (fees paid directly from your investment)	
	None
Annual Fund Operating Expenses (expenses that you pay each year as a % of the value of your investment)	
Management fees	0.02
Other expenses	None
Total annual fund operating expenses[1]	**0.02**

[1] The information in the table has been restated to reflect current fees and expenses.

Example

This example is intended to help you compare the cost of investing in the fund with the cost of investing in other funds. The example assumes that you invest $10,000 in the fund for the time periods indicated and then redeem all of your shares at the end of those time periods. The example also assumes that your investment has a 5% return each year and that the fund's operating expenses remain the same. The figures are based on total annual fund operating expenses after any expense reduction. The example does not reflect any brokerage fees or commissions you may incur when buying or selling fund shares. Your actual costs may be higher or lower.

Expenses on a $10,000 Investment

1 Year	3 Years	5 Years	10 Years
$2	$6	$11	$26

Example 2: Excerpts from Deutsche's S&P 500 index fund prospectus

Note: details available later in Deutsche's prospectus and presented in our Online Appendix show that classes R6 and S are not available to most retail investors.

Deutsche S&P 500 Index Fund

INVESTMENT OBJECTIVE

The fund seeks to provide investment results that, before expenses, correspond to the total return of common stocks publicly traded in the United States, as represented by the Standard & Poor's 500 Composite Stock Price Index (S&P 500® Index).

The fund invests for capital appreciation, not income; any dividend and interest income is incidental to the pursuit of its objective.

The fund is a feeder fund that invests substantially all of its assets in a "master portfolio," the Deutsche Equity 500 Index Portfolio (the "Portfolio"), which will invest directly in securities and other instruments. The Portfolio has the same investment objective and strategies as the fund. References to investments by the fund may refer to actions undertaken by the Portfolio.

FEES AND EXPENSES OF THE FUND

These are the fees and expenses you may pay when you buy and hold shares. You may qualify for sales charge discounts if you and your immediate family invest, or agree to invest in the future, at least $100,000 in Class A shares in Deutsche funds or if you invest at least $250,000 in Class T shares in the fund. More information about these and other discounts and waivers is available from your financial advisor and in Choosing a Share Class (p. 34), Sales Charge Waivers and Discounts Available Through Intermediaries (Appendix B, p. 74) and Purchase and Redemption of Shares in the fund's Statement of Additional Information (SAI) (p. II-16).

SHAREHOLDER FEES (paid directly from your investment)

	A	T	C	R6	S
Maximum sales charge (load) imposed on purchases, as % of offering price	4.50	2.50	None	None	None
Maximum deferred sales charge (load), as % of redemption proceeds	None	None	1.00	None	None
Account Maintenance Fee (annually, for fund account balances below $10,000 and subject to certain exceptions)	$20	None	$20	None	$20

ANNUAL FUND OPERATING EXPENSES

(expenses that you pay each year as a % of the value of your investment)

	A	T	C	R6	S
Management fee	0.05	0.05	0.05	0.05	0.05
Distribution/service (12b-1) fees	0.24	0.25	0.99	None	None
Other expenses[1]	0.30	0.30	0.26	0.35	0.29
Total annual fund operating expenses[2]	0.59	0.60	1.30	0.40	0.34
Fee waiver/expense reimbursement	0.00	0.00	0.00	0.05	0.00
Total annual fund operating expenses after fee waiver/expense reimbursement	0.59	0.60	1.30	0.35	0.34

[1] "Other expenses" for Class T are based on estimated amounts for the current fiscal year.

[2]The table and Example below reflect the expenses of both the fund and the Portfolio.

The Advisor has contractually agreed through April 30, 2019 to waive its fees and/or reimburse fund expenses, including expenses of the Portfolio allocated to the fund, to the extent necessary to maintain the fund's total annual operating expenses (excluding certain expenses such as extraordinary expenses, taxes, brokerage, interest and acquired fund fees and expenses) at a ratio no higher than 0.35% for Class R6. The agreement may only be terminated with the consent of the fund's Board.

EXAMPLE

This Example is intended to help you compare the cost of investing in the fund with the cost of investing in other mutual funds. The Example assumes that you invest $10,000 in the fund for the time periods indicated and then redeem all of your shares at the end of those periods. The Example also assumes that your investment has a 5% return each year and that the fund's operating expenses (including one year of capped expenses in each period for Class R6) remain the same. Although your actual costs may be higher or lower, based on these assumptions your costs would be:

Years	A	T	C	R6	S
1	$ 508	$310	$ 232	$ 36	$ 35
3	631	437	412	123	109
5	764	576	713	219	191
10	1,155	981	1,568	500	431

You would pay the following expenses if you did not redeem your shares:

Years	A	T	C	R6	S
1	$ 508	$310	$ 132	$ 36	$ 35
3	631	437	412	123	109
5	764	576	713	219	191
10	1,155	981	1,568	500	431

Appendix C: Examples of "Equity Risk" Disclosures

Below is the text of the "equity risk" risk factor disclosures from Deutsche's and Schwab's 2017 prospectuses. The first column is the equity risk discussion copied from the prospectus summary sections. The second column is the equity risk discussion copied from the prospectus detailed sections.

	Text from Prospectus Summary Sections	Text from Prospectus Detail Sections
Schwab	The prices of equity securities rise and fall daily. These price movements may result from factors affecting individual companies, industries or the securities market as a whole. In addition, equity markets tend to move in cycles, which may cause stock prices to fall over short or extended periods of time.	The prices of equity securities rise and fall daily. These price movements may result from factors affecting individual companies, industries or the securities market as a whole. Individual companies may report poor results or be negatively affected by industry and/or economic trends and developments. The prices of securities issued by such companies may suffer a decline in response. In addition, equity markets tend to move in cycles, which may cause stock prices to fall over short or extended periods of time.
Deutsche	When stock prices fall, you should expect the value of your investment to fall as well. Stock prices can be hurt by poor management on the part of the stock's issuer, shrinking product demand and other business risks. These may affect single companies as well as groups of companies. The market as a whole may not favor the types of investments the fund makes, which could adversely affect a stock's price, regardless of how well the company performs, or the fund's ability to sell a stock at an attractive price. There is a chance that stock prices overall will decline because stock markets tend to move in cycles, with periods of rising and falling prices. Events in the US and global financial markets, including actions taken by the US Federal Reserve or foreign central banks to stimulate or stabilize economic growth, may at times result in unusually high market volatility which could negatively affect performance. Further, geopolitical and other events, including war, terrorism, economic uncertainty, trade disputes and related geopolitical events have led, and in the future may lead, to increased short-term market volatility, which may disrupt securities markets and have adverse long-term effects on US and world economies and markets. To the extent the fund invests in a particular capitalization or sector, the fund's performance may be affected by the general performance of that particular capitalization or sector.	When stock prices fall, you should expect the value of your investment to fall as well. Stock prices can be hurt by poor management on the part of the stock's issuer, shrinking product demand and other business risks. These may affect single companies as well as groups of companies. The market as a whole may not favor the types of investments the fund makes, which could adversely affect a stock's price, regardless of how well the company performs, or the fund's ability to sell a stock at an attractive price. There is a chance that stock prices overall will decline because stock markets tend to move in cycles, with periods of rising and falling prices. Events in the US and global financial markets, including actions taken by the US Federal Reserve or foreign central banks to stimulate or stabilize economic growth, may at times result in unusually high market volatility which could negatively affect performance. Further, geopolitical and other events, including war, terrorism, economic uncertainty, trade disputes and related geopolitical events have led, and in the future may lead, to increased short-term market volatility, which may disrupt securities markets and have adverse long-term effects on US and world economies and markets. To the extent the fund invests in a particular capitalization or sector, the fund's performance may be affected by the general performance of that particular capitalization or sector.

Figure 1: Fund Fee and Complexity Strategies

This figure depicts index fund managers' strategic choices based on Carlin (2009). In Carlin (2009), index fund managers choose strategies within a mixed-strategy Nash equilibrium. We simplify this figure to two strategies for illustrative purposes. Under the "simple" strategy, the manager chooses low fees and simplicity. The "complex" strategy is to choose high fees and complexity.

Uninformed investors cannot understand disclosures so invest randomly across both simple and complex strategy funds. Informed investors are able to understand disclosures and identify the cheapest funds, so they only invest in the simple strategy funds. Because simple strategy funds get their equal share of uninformed investors plus all of the informed investors, they are larger than complex strategy funds. The fraction of uninformed investors is determined endogenously by aggregate complexity across all funds. The model is competitive in that all funds earn equal profits in equilibrium.

This figure adds a third likely outcome from managers' strategic choice: a complementary marketing strategy. The dotted lines indicate that marketing is not explicitly considered in Carlin (2009). However, it seems likely that managers complement their high-complexity strategy by engaging in aggressive marketing; e.g., paying brokers to steer uninformed investors into the high-fee fund. Moreover, marketing should be more effective when targeting uninformed investors who cannot independently evaluate funds. Managers choosing the low-complexity strategy primarily target informed investors, so they should have weaker marketing incentives.

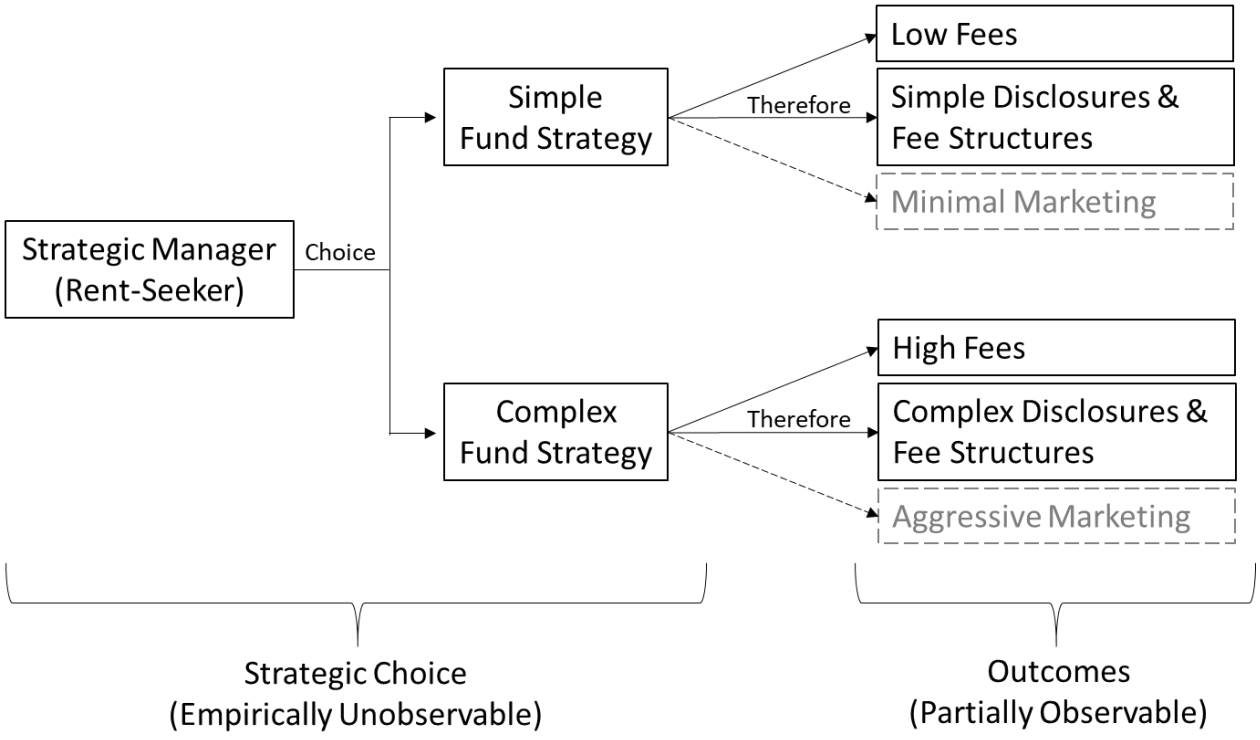


Figure 2: How Funds are Structured

This figure illustrates the structure of a typical mutual fund. Mutual funds are issued by financial institutions such as Schwab. A parent institution can issue many different funds. Each fund can be subdivided into classes with different tickers. All classes within a fund have the same investments and same gross returns, but can have different types and levels of fees. Most fees are paid on an ongoing basis and are expressed in percentage points. "Loads" are typically one-time charges upon the purchase or sale of the fund. Load amounts can differ depending on factors such as the amount purchased or sold, or the length the investor has held the fund. Waivers can reduce fees and loads if certain conditions are met.

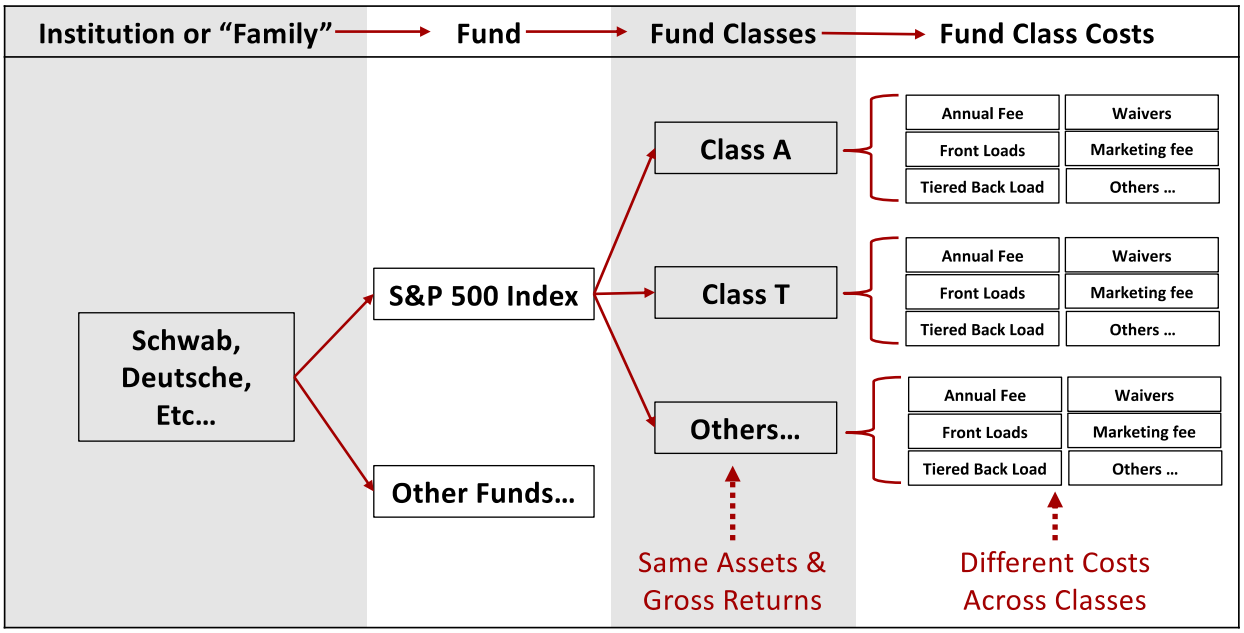
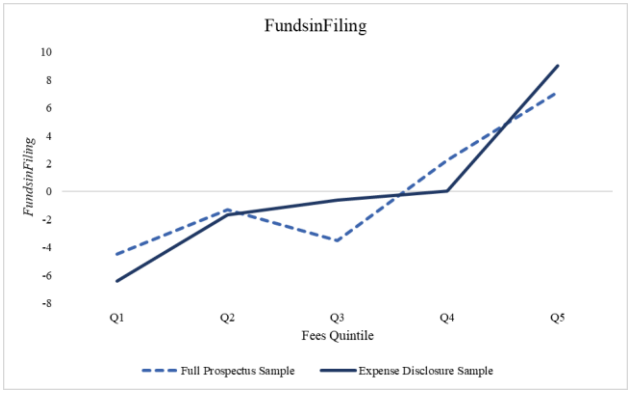
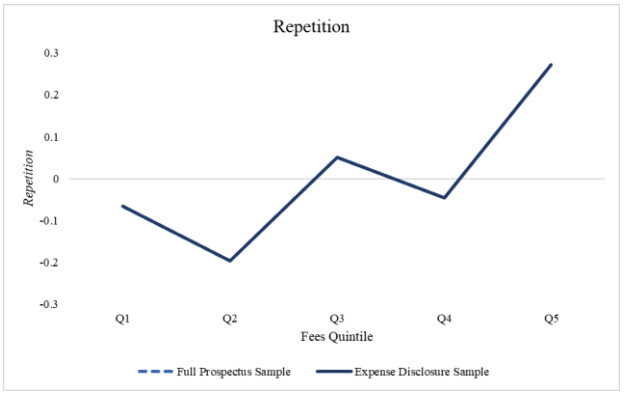
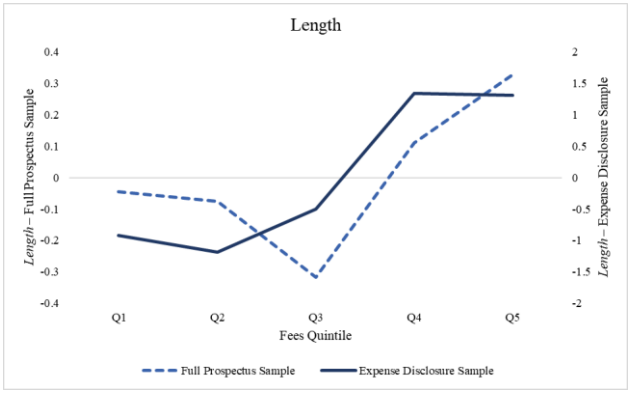


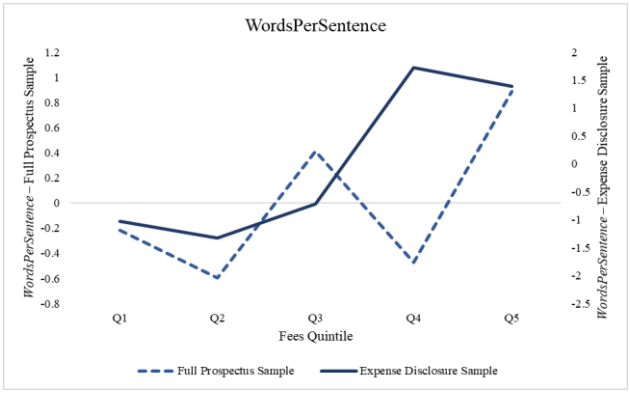
Figure 3: Graphs of Complexity by Quintiles of Fees

This figure depicts the average of each of our narrative and structural complexity measures (on the vertical axes) for each quintile of *Fees* (horizontal axis). *Fees* and complexity measures are orthogonalized to year fixed effects to remove time trends. The dotted line measures complexity within the full prospectus sample while the solid line is the expense disclosure sample. The two lines overlap exactly for *Repetition* because the variable is the same in both samples by construction. For brevity, we present only the principal component of our five structural complexity measures, *Structural_Complexity*. All variables are defined in Appendix A.









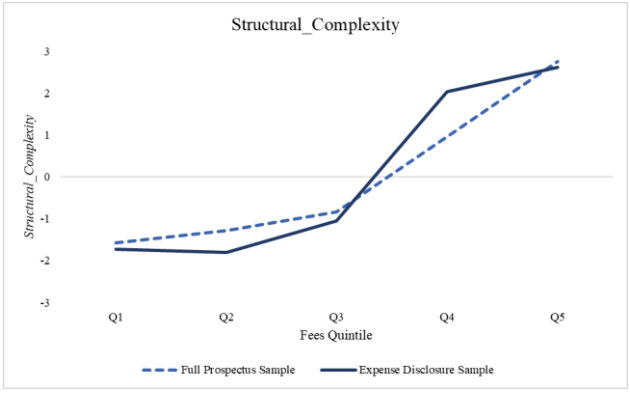


Table 1: Summary Statistics and Descriptive Information

This table presents summary statistics. Panel A presents summary statistics for the pooled sample. To facilitate interpretation, we present summary statistics for non-standardized variable values. The rightmost column tabulates the residual standard deviation in each variable after it is orthogonalized to the year fixed effects used in our regressions. All variables are defined in Appendix A and are winsorized at 1% and 99%. Panel B presents the averages of variables for fee details and fund characteristics, for funds with fees in the lowest 20% of our sample ("low-fee funds") and the remaining 80% of our sample ("high-fee funds"). *t*-statistics are in parentheses. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Variable summary statistics

	N	Mean	P25	Median	P75	Std. Dev.	Residual Std. Dev.
Fund variables							
Fees	458	0.689	0.200	0.554	1.150	0.526	0.510
TE	452	3.445	1.481	2.467	3.885	3.628	3.368
Size	458	7.06	5.60	7.11	7.91	1.75	1.69
Turnover	458	0.117	0.040	0.060	0.100	0.237	0.235
Structural complexity variables							
ShareClasses	458	2.578	1.000	2.000	3.000	1.925	1.819
FrontLoadBreaks	458	1.821	0.000	0.000	5.000	2.704	2.661
RearLoadBreaks	458	1.283	0.000	1.000	1.000	1.787	1.772
FrontLoad	458	0.298	0.000	0.000	1.000	0.456	0.445
NoLoad_12b1	458	0.566	0.000	1.000	1.000	0.495	0.484
Structural_Complexity	458	0.000	-0.798	-0.419	1.031	1.000	0.976
Full prospectus narrative variables							
FundsinFiling	286	2.059	1.225	2.120	2.767	0.964	0.942
Repetition	123	0.311	0.074	0.229	0.511	0.259	0.255
Length	458	0.000	-0.700	-0.225	0.347	1.000	0.945
WordsPerSentence	458	25.387	23.931	25.525	27.123	3.064	2.957
Expense disclosure narrative variables							
FundsinFiling	123	2.235	1.386	2.383	2.866	0.950	0.904
Repetition	123	0.311	0.074	0.229	0.511	0.259	0.255
Length_ExpDisc	123	0.000	-1.119	0.409	0.830	1.000	0.996
WordsPerSent_ExpDisc	123	24.929	20.000	24.000	28.333	6.224	6.190
Wordiness_ExpDisc	123	0.000	-0.495	-0.488	0.485	1.000	0.991

Panel B: Additional descriptive information, by low-fee versus high-fee

		Average for Low-Fee Funds	Average for High-Fee Funds	Difference (Low – High)	t-stat
	Fee details at the fund-class level				
(i)	Max. annualized *Fees* across all classes	0.164	0.842	-0.678***	(-7.28)
(ii)	Min. annualized *Fees* across all classes	0.145	0.555	-0.410***	(-6.67)
(iii)	Max. 12b-1 fee across all classes	0.021	0.385	-0.365***	(-4.76)
(iv)	Min. 12b-1 fee across all classes	0.019	0.062	-0.043**	(-2.23)
(v)	Max. front load across all classes (not annualized)	0.000	1.492	-1.492***	(-3.71)
(vi)	Min. front load across all classes (not annualized)	0.000	0.046	-0.046	(-1.34)
(vii)	Max. rear load across all classes (not annualized)	0.254	1.089	-0.836***	(-2.81)
(viii)	Min. rear load across all classes (not annualized)	0.000	0.000	0.000	(0.00)
	Fund characteristics				
(ix)	*Size*	8.463	6.648	1.816**	(2.35)
(x)	Age (in years)	18.456	10.470	7.985**	(2.33)
(xi)	Prospectus update frequency (per year)	3.777	4.321	-0.544	(-1.20)

Table 2: Correlations

This table presents Pearson (Spearman) correlations above (below) the diagonal. P-values are presented in parentheses. Panel A presents results for the full prospectus sample, and Panel B presents results for the expense disclosure sample. All variables are defined in Appendix A and winsorized at 1% and 99%. To facilitate comparisons, all variables except *Fees* in this table and the remaining tables are standardized to have a mean (standard deviation) of zero (one).

Panel A: Pearson-Spearman correlations, Full Prospectus Sample

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	*Fees*		0.042	-0.267	0.429	0.678	0.842	0.644	0.861	0.618	0.874	0.389	0.446	0.177	0.080
			(0.367)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.089)
2	*TE*	0.231		-0.202	0.072	-0.061	-0.017	-0.005	-0.012	0.055	-0.013	0.021	0.028	-0.096	-0.004
		(0.000)		(0.000)	(0.126)	(0.196)	(0.726)	(0.910)	(0.801)	(0.247)	(0.786)	(0.732)	(0.760)	(0.041)	(0.939)
3	*Size*	-0.443	-0.593		-0.081	-0.041	-0.178	-0.169	-0.177	0.057	-0.131	-0.120	-0.226	-0.056	-0.199
		(0.000)	(0.000)		(0.084)	(0.380)	(0.000)	(0.000)	(0.000)	(0.224)	(0.005)	(0.043)	(0.012)	(0.229)	(0.000)
4	*Turnover*	0.037	0.178	-0.218		0.027	0.149	0.045	0.205	0.168	0.143	0.199	0.249	0.102	0.068
		(0.427)	(0.000)	(0.000)		(0.558)	(0.001)	(0.341)	(0.000)	(0.000)	(0.002)	(0.001)	(0.006)	(0.029)	(0.148)
5	*ShareClasses*	0.735	0.024	-0.143	-0.055		0.727	0.520	0.691	0.520	0.828	0.369	0.193	0.285	-0.063
		(0.000)	(0.616)	(0.002)	(0.238)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.032)	(0.000)	(0.176)
6	*FrontLoadBreaks*	0.699	0.132	-0.200	0.033	0.621		0.727	0.977	0.505	0.956	0.355	0.353	0.183	0.048
		(0.000)	(0.005)	(0.000)	(0.482)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.301)
7	*RearLoadBreaks*	0.647	0.172	-0.201	-0.072	0.494	0.687		0.652	0.379	0.785	0.142	0.378	0.064	0.012
		(0.000)	(0.000)	(0.000)	(0.122)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.016)	(0.000)	(0.172)	(0.806)
8	*FrontLoad*	0.784	0.079	-0.257	0.049	0.718	0.911	0.641		0.513	0.933	0.361	0.319	0.182	0.047
		(0.000)	(0.093)	(0.000)	(0.300)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.318)
9	*NoLoad_12b1*	0.581	0.196	-0.311	0.074	0.619	0.450	0.359	0.514		0.667	0.290	0.341	0.213	0.007
		(0.000)	(0.000)	(0.000)	(0.116)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.878)
10	*Structural_Complexity*	0.773	0.162	-0.250	-0.018	0.807	0.770	0.683	0.795	0.835		0.367	0.362	0.219	0.015
		(0.000)	(0.001)	(0.000)	(0.695)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.757)
11	*FundsinFiling*	0.347	0.236	-0.097	0.172	0.422	0.433	0.290	0.423	0.331	0.427		0.126	0.658	-0.083
		(0.000)	(0.000)	(0.103)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.164)	(0.000)	(0.160)
12	*Repetition*	0.363	0.259	-0.363	-0.112	0.112	0.277	0.319	0.237	0.329	0.288	0.163		-0.096	0.263
		(0.000)	(0.004)	(0.000)	(0.216)	(0.218)	(0.002)	(0.000)	(0.008)	(0.000)	(0.001)	(0.072)		(0.293)	(0.003)
13	*Length*	0.162	-0.123	0.090	0.125	0.313	0.181	0.163	0.244	0.229	0.276	0.587	-0.073		-0.038
		(0.001)	(0.009)	(0.055)	(0.007)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.424)		(0.413)
14	*WordsPerSentence*	0.006	0.028	-0.106	0.036	-0.043	0.031	-0.097	0.007	-0.059	-0.058	-0.265	0.229	-0.070	
		(0.905)	(0.554)	(0.023)	(0.446)	(0.360)	(0.510)	(0.037)	(0.876)	(0.207)	(0.219)	(0.000)	(0.011)	(0.132)	

Panel B: Pearson-Spearman correlations, Expense Disclosure Sample

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	*Fees*		0.077	-0.404	0.509	0.630	0.855	0.588	0.875	0.685	0.852	0.463	0.446	0.755	0.577	0.739
			(0.398)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2	*TE*	0.402		-0.210	-0.006	-0.035	0.030	-0.005	0.060	0.065	0.025	0.006	0.028	0.160	0.082	0.144
		(0.000)		(0.020)	(0.945)	(0.701)	(0.738)	(0.954)	(0.510)	(0.475)	(0.786)	(0.948)	(0.760)	(0.076)	(0.368)	(0.113)
3	*Size*	-0.581	-0.692		-0.105	-0.107	-0.308	-0.263	-0.318	-0.026	-0.252	-0.194	-0.226	-0.353	-0.136	-0.302
		(0.000)	(0.000)		(0.246)	(0.238)	(0.001)	(0.003)	(0.000)	(0.771)	(0.005)	(0.031)	(0.012)	(0.000)	(0.133)	(0.001)
4	*Turnover*	0.209	0.175	-0.247		-0.044	0.166	0.049	0.243	0.165	0.134	0.293	0.249	0.269	0.197	0.261
		(0.020)	(0.053)	(0.006)		(0.629)	(0.067)	(0.590)	(0.007)	(0.069)	(0.138)	(0.001)	(0.006)	(0.003)	(0.029)	(0.004)
5	*ShareClasses*	0.689	0.142	-0.204	-0.022		0.802	0.582	0.724	0.524	0.863	0.323	0.193	0.519	0.559	0.561
		(0.000)	(0.117)	(0.024)	(0.805)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.032)	(0.000)	(0.000)	(0.000)
6	*FrontLoadBreaks*	0.849	0.263	-0.390	0.074	0.830		0.706	0.974	0.641	0.977	0.320	0.353	0.799	0.637	0.791
		(0.000)	(0.003)	(0.000)	(0.419)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
7	*RearLoadBreaks*	0.759	0.320	-0.403	0.030	0.528	0.758		0.595	0.361	0.759	0.116	0.378	0.446	0.374	0.447
		(0.000)	(0.000)	(0.000)	(0.741)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.201)	(0.000)	(0.000)	(0.000)	(0.000)
8	*FrontLoad*	0.850	0.323	-0.479	0.174	0.792	0.957	0.706		0.656	0.937	0.317	0.319	0.844	0.680	0.838
		(0.000)	(0.000)	(0.000)	(0.054)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
9	*NoLoad_12b1*	0.649	0.355	-0.464	0.044	0.601	0.621	0.398	0.639		0.715	0.360	0.341	0.661	0.522	0.652
		(0.000)	(0.000)	(0.000)	(0.626)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
10	*Structural_Complexity*	0.811	0.316	-0.412	0.021	0.859	0.890	0.698	0.858	0.837		0.334	0.362	0.767	0.653	0.773
		(0.000)	(0.000)	(0.000)	(0.816)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
11	*FundsinFiling*	0.493	0.339	-0.228	0.197	0.480	0.448	0.392	0.428	0.436	0.510		0.126	0.264	0.066	0.214
		(0.000)	(0.000)	(0.011)	(0.029)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.164)	(0.003)	(0.469)	(0.018)
12	*Repetition*	0.363	0.259	-0.363	-0.112	0.112	0.277	0.319	0.237	0.329	0.288	0.163		0.278	0.155	0.253
		(0.000)	(0.004)	(0.000)	(0.216)	(0.218)	(0.002)	(0.000)	(0.008)	(0.000)	(0.001)	(0.072)		(0.002)	(0.088)	(0.005)
13	*Length_ExpDisc*	0.650	0.312	-0.481	0.227	0.624	0.709	0.358	0.787	0.665	0.674	0.313	0.162		0.758	0.977
		(0.000)	(0.000)	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.074)		(0.000)	(0.000)
14	*WordsPerSent_ExpDisc*	0.605	0.229	-0.410	0.182	0.648	0.708	0.470	0.761	0.623	0.689	0.207	0.127	0.866		0.880
		(0.000)	(0.011)	(0.000)	(0.044)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.021)	(0.160)	(0.000)		(0.000)
15	*Wordiness_ExpDisc*	0.643	0.304	-0.466	0.202	0.639	0.712	0.394	0.784	0.664	0.679	0.283	0.146	0.984	0.924	
		(0.000)	(0.001)	(0.000)	(0.025)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.106)	(0.000)	(0.000)	

Table 3: Analysis of Economic Magnitudes of Fee Dispersion

This table presents the total net assets and weighted average fees for funds with fees in the lowest 20% of our sample (columns i and ii, "low-fee funds") and the remaining 80% of our sample (columns iii and iv, "high-fee funds"). Average fees are weighted by net assets at the fund-class level. Column v presents the difference in fees between low-fee and high-fee funds, and column vi presents the estimated total fee savings if the investors in the high-fee funds instead held low-fee funds. Column vii reports the fee savings as a fraction of the total net assets of S&P 500 index funds.

Year	(i) Low-Fee Funds: Net Assets ($M)	(ii) Low-Fee Funds: Weighted Avg. Fees (bps)	(iii) High-Fee Funds: Net Assets ($M)	(iv) High-Fee Funds: Weighted Avg. Fees (bps)	(v) =(iv)-(ii) High-Low Fee Difference (bps)	(vi) =(v)*(iii) High-Low Fee Difference ($M)	(vii) =(vi)/[(i)+(iii)] High-Low Fee Diff. as a Percent of Total Investment (bps)
1996	$43	9	$2,766	26.2	17.2	$4.8	0.169
1997	$1,295	18.8	$5,773	30.2	11.4	$6.6	0.093
1998	$3,222	18	$12,392	34.8	16.8	$20.8	0.133
1999	$97,494	18.9	$16,703	36.8	17.9	$29.9	0.026
2000	$117,976	18.2	$21,629	40.3	22.1	$47.8	0.034
2001	$90,222	18.1	$17,951	53.8	35.7	$64.1	0.059
2002	$77,412	18.2	$12,462	55.3	37.1	$46.2	0.051
2003	$88,515	18.1	$17,209	64.7	46.6	$80.2	0.076
2004	$115,955	18.1	$23,937	69.7	51.6	$123.5	0.088
2005	$122,970	18.1	$35,932	64.1	46.0	$165.3	0.104
2006	$122,833	17.9	$40,392	63.2	45.3	$183.0	0.112
2007	$136,387	17.3	$39,495	59.8	42.5	$167.9	0.095
2008	$103,774	16.9	$31,131	57.4	40.5	$126.1	0.093
2009	$103,050	15.1	$25,590	57.6	42.5	$108.8	0.085
2010	$145,495	15	$28,602	60.7	45.7	$130.7	0.075
2011	$153,033	15.4	$33,320	61.9	46.5	$154.9	0.083
2012	$153,817	15.4	$33,342	56.1	40.7	$135.7	0.073
2013	$204,271	14.7	$44,523	59.6	44.9	$199.9	0.080
2014	$257,416	14.6	$50,693	62.5	47.9	$242.8	0.079
2015	$288,410	14.2	$62,529	63.4	49.2	$307.6	0.088
2016	$314,438	14.6	$65,064	62.9	48.3	$314.3	0.083
2017	$390,746	13.8	$71,757	63.7	49.9	$358.1	0.077

Table 4: Prediction 1 – Narrative Complexity

This table presents results of regressing total ownership cost (*Fees*) on narrative complexity variables, as per model (1). Panel A presents results for the full prospectus, and Panel B presents results for the expense disclosure. OLS or robust regressions are used, as indicated at the bottom of each column. Year fixed effects are included in all models. All variables are defined in Appendix A. *t*-statistics are in parentheses. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Full Prospectus Sample

	Dependent variable: *Fees*					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
FundsinFiling	0.198***				0.178*	0.172**
	(2.70)				(1.79)	(2.09)
Repetition		0.280***			0.231***	0.269***
		(3.15)			(3.03)	(3.71)
Length			0.071		0.158	0.154
			(1.20)		(1.42)	(1.47)
WordsPerSentence				0.063	0.085	0.055
				(1.04)	(1.44)	(1.36)
Estimation	OLS	OLS	OLS	OLS	OLS	Robust Regression
Adjusted R^2	0.090	0.174	0.028	0.024	0.336	0.407
Observations	286	123	458	458	123	123

Panel B: Expense Disclosure Sample

	Dependent variable: *Fees*					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
FundsinFiling	0.282***				0.139***	0.118***
	(2.86)				(3.47)	(3.83)
Repetition		0.280***			0.158***	0.170***
		(3.15)			(3.74)	(4.13)
Length_ExpDisc			0.453***			
			(5.89)			
WordsPerSent_ExpDisc				0.346***		
				(3.57)		
Wordiness_ExpDisc					0.366***	0.374***
					(6.06)	(7.78)
Estimation	OLS	OLS	OLS	OLS	OLS	Robust Regression
Adjusted R^2	0.210	0.222	0.579	0.342	0.680	0.751
Observations	123	123	123	123	123	123

Table 5: Prediction 2 – Structural Complexity

This table presents results of regressing total ownership cost (*Fees*) on structural complexity variables, as per model (1). Panel A presents results for the full prospectus, and Panel B presents results for the expense disclosure. OLS or robust regressions are used, as indicated at the bottom of each column. Year fixed effects are included in all models. All variables are defined in Appendix A. *t*-statistics are in parentheses. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Full Prospectus Sample

	Dependent variable: *Fees*						
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
ShareClasses	0.355***						
	(5.92)						
FrontLoadBreaks		0.437***					
		(10.65)					
RearLoadBreaks			0.340***				
			(6.83)				
FrontLoad				0.448***			
				(10.07)			
NoLoad_12b1					0.316***		
					(4.64)		
Structural_Complexity						0.454***	0.436***
						(10.53)	(14.96)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	Robust
Adjusted R^2	0.438	0.715	0.442	0.735	0.373	0.758	0.849
Observations	458	458	458	458	458	458	458

Panel B: Expense Disclosure Sample

	Dependent variable: *Fees*						
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
ShareClasses	0.377***						
	(6.31)						
FrontLoadBreaks		0.509***					
		(8.85)					
RearLoadBreaks			0.349***				
			(4.24)				
FrontLoad				0.521***			
				(8.66)			
NoLoad_12b1					0.408***		
					(5.26)		
Structural_Complexity						0.507***	0.475***
						(7.81)	(10.95)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	Robust
Adjusted R^2	0.363	0.717	0.310	0.755	0.443	0.711	0.842
Observations	123	123	123	123	123	123	123

Table 6: Narrative Complexity of Specific Disclosures

This table presents results of regressing total ownership cost (*Fees*) on narrative complexity variables, specifically for disclosure items that are unlikely to be affected by structural complexity, marketing, and other unobservable strategies. Panel A presents results for funds' objective disclosures, and Panel B presents results for funds' equity risk disclosures. Appendix C provides examples of equity risk disclosures. OLS or robust regressions are used, as indicated at the bottom of each column. Year fixed effects are included in all models in Panel A. All variables are defined in Appendix A. *t*-statistics are in parentheses. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Analysis of fund objective statement length and words per sentence

	Dependent variable: *Fees*			
	(i)	(ii)	(iii)	(iv)
Length_ObjDisc	0.176***		0.169***	
	(2.75)		(3.14)	
WordsPerSent_ObjDisc		0.051		0.063
		(0.49)		(0.66)
Estimation	OLS	OLS	Robust Regression	Robust Regression
Adjusted R^2	0.038	-0.045	0.042	-0.039
Observations	123	123	123	123

Panel B: Analysis of equity risk disclosure summarization, length, and words per sentence

	Dependent variable: *Fees*							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Summarization	-0.079**				-0.097***			
	(-2.53)				(-3.42)			
Summarization_Zero		0.574***				0.651***		
		(3.03)				(3.87)		
Length_RiskDisc			0.116**				0.121**	
			(2.34)				(2.48)	
WordsPerSent_RiskDisc				0.031***				0.032***
				(3.42)				(3.77)
Estimation	OLS	OLS	OLS	OLS	Robust Regression	Robust Regression	Robust Regression	Robust Regression
Adjusted R^2	0.075	0.097	0.038	0.210	0.215	0.197	0.069	0.231
Observations	25	25	25	25	25	25	25	25

Table 7: Additional Analysis of Narrative Complexity and Discretionary Fees

This table presents results of regressing total ownership cost (*Fees*) on narrative complexity variables, as per model (2). Panel A presents results for the full prospectus, and Panel B presents results for the expense disclosure. Control variables are included in all columns to isolate the discretionary component of *Fees*. OLS or robust regressions are used, as indicated at the bottom of each column. Year fixed effects are included in all models. All variables are defined in Appendix A and winsorized at 1% and 99%. *t*-statistics are in parentheses. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Full Prospectus Sample

	Dependent variable: *Fees*					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
FundsinFiling	0.129**				0.043	0.049
	(2.03)				(0.52)	(0.67)
Repetition		0.125*			0.134**	0.134**
		(1.85)			(2.13)	(2.50)
Length			0.049		0.211**	0.229**
			(1.00)		(2.11)	(2.38)
WordsPerSentence				0.004	0.018	0.005
				(0.09)	(0.46)	(0.16)
Size	-0.231***	-0.255***	-0.225***	-0.223***	-0.236***	-0.253***
	(-5.25)	(-5.32)	(-4.87)	(-4.59)	(-4.70)	(-5.30)
Turnover	0.168***	0.151***	0.175***	0.181***	0.137***	-0.510
	(7.90)	(6.29)	(6.88)	(7.86)	(5.86)	(-1.32)
TE	-0.043	-0.021	-0.041	-0.045*	-0.007	-0.613
	(-1.51)	(-0.79)	(-1.56)	(-1.67)	(-0.31)	(-0.99)
Estimation	OLS	OLS	OLS	OLS	OLS	Robust Regression
Adjusted R^2	0.449	0.524	0.365	0.357	0.601	0.541
Observations	281	123	452	452	123	123

Panel B: Expense Disclosure Sample

	Dependent variable: *Fees*					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
FundsinFiling	0.147*				0.101**	0.080**
	(1.87)				(2.51)	(2.62)
Repetition		0.125*			0.097**	0.105***
		(1.85)			(2.42)	(3.06)
Length_ExpDisc			0.323***			
			(4.55)			
WordsPerSent_ExpDisc				0.226***		
				(4.90)		
Wordiness_ExpDisc					0.290***	0.316***
					(5.59)	(8.66)
Size	-0.296***	-0.285***	-0.167***	-0.260***	-0.141***	-0.143***
	(-6.45)	(-5.32)	(-3.71)	(-6.65)	(-4.06)	(-4.69)
Turnover	0.197***	0.202***	0.173***	0.195***	0.142***	0.132***
	(5.86)	(6.29)	(6.26)	(7.95)	(5.53)	(6.26)
TE	-0.045	-0.022	-0.046*	-0.033	-0.044**	-0.046***
	(-1.46)	(-0.79)	(-1.74)	(-1.19)	(-2.06)	(-4.37)
Estimation	OLS	OLS	OLS	OLS	OLS	Robust Regression
Adjusted R^2	0.576	0.563	0.726	0.653	0.771	0.863
Observations	123	123	123	123	123	123

Table 8: Additional Analysis of Structural Complexity and Discretionary Fees

This table presents results of regressing total ownership cost (*Fees*) on structural complexity variables, as per model (2). Panel A presents results for the full prospectus, and Panel B presents results for the expense disclosure. Control variables are included in all columns to isolate the discretionary component of *Fees*. OLS or robust regressions are used, as indicated at the bottom of each column. Year fixed effects are included in all models. All variables are defined in Appendix A and winsorized at 1% and 99%. *t*-statistics are in parentheses. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Full Prospectus Sample

	Dependent variable: *Fees*						
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
ShareClasses	0.332***						
	(6.15)						
FrontLoadBreaks		0.387***					
		(14.71)					
RearLoadBreaks			0.297***				
			(7.77)				
FrontLoad				0.391***			
				(12.00)			
NoLoad_12b1					0.248***		
					(3.90)		
Structural_Complexity						0.403***	0.406***
						(13.60)	(15.17)
Size	-0.188***	-0.103***	-0.144***	-0.097***	-0.160***	-0.101***	-0.100***
	(-5.49)	(-5.36)	(-4.33)	(-4.76)	(-2.72)	(-5.66)	(-6.49)
Turnover	0.183***	0.148***	0.181***	0.128***	0.154***	0.149***	0.173***
	(7.80)	(16.38)	(13.08)	(10.30)	(5.70)	(12.05)	(14.50)
TE	-0.026	0.003	-0.023	0.001	-0.045*	-0.007	-0.001
	(-1.59)	(0.27)	(-1.23)	(0.09)	(-1.93)	(-0.73)	(-0.28)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	Robust
Adjusted R^2	0.730	0.848	0.666	0.835	0.563	0.889	0.904
Observations	452	452	452	452	452	452	452

Panel B: Expense Disclosure Sample

	Dependent variable: *Fees*						
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
ShareClasses	0.350***						
	(7.75)						
FrontLoadBreaks		0.417***					
		(10.56)					
RearLoadBreaks			0.262***				
			(5.61)				
FrontLoad				0.422***			
				(8.43)			
NoLoad_12b1					0.288***		
					(3.73)		
Structural_Complexity						0.420***	0.414***
						(10.05)	(10.93)
Size	-0.263***	-0.126***	-0.225***	-0.113**	-0.197**	-0.132***	-0.141***
	(-6.81)	(-3.14)	(-5.15)	(-2.38)	(-2.30)	(-4.58)	(-5.69)
Turnover	0.255***	0.205***	0.237***	0.175***	0.208***	0.216***	0.218***
	(20.51)	(16.80)	(14.88)	(10.10)	(8.67)	(21.59)	(24.51)
TE	-0.016	0.008	-0.002	-0.001	-0.015	0.006	0.012*
	(-0.82)	(0.85)	(-0.11)	(-0.10)	(-0.70)	(0.60)	(1.68)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	Robust
Adjusted R^2	0.842	0.900	0.669	0.878	0.689	0.918	0.933
Observations	123	123	123	123	123	123	123

Table 9: Parent-Level Information

This table presents summary statistics of parent-level characteristics. We present the average of each variable, across all funds of the parents, for parents of funds with fees in the lowest 20% of our sample ("low-fee funds") and the remaining 80% of our sample ("high-fee funds"). *t*-statistics are in parentheses. Rows (i) through (vii) are parent-year observations for 2000 – 2017. Row (viii) contains one observation for each parent company in our expense disclosure sample. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

	Variable	N	Average for Parents of Low-Fee S&P 500 Funds	Average for Parents of High-Fee S&P 500 Funds	Difference (Low – High)	t-stat
(i)	Average *Fees*	427	0.540	1.435	-0.895***	(-6.30)
(ii)	Average Fund *Size*	427	6.082	5.268	0.814*	(1.77)
(iii)	Public Company Indicator	427	0.468	0.508	-0.039	(-0.21)
(iv)	Number of mutual funds	427	169.037	80.662	88.375**	(2.35)
(v)	Number of passive funds	427	58.596	12.300	46.295***	(3.38)
(vi)	Number of active funds	427	110.388	68.359	42.029	(0.97)
(vii)	Number of fund categories	427	29.479	20.080	9.399***	(3.79)
(viii)	Number of other services offered	28	3.000	2.136	0.864	(0.98)

Table 10: Is High-Fee & High-Complexity a Parent-Level Strategy?

This table presents results of our analysis of complexity and fees among all funds issued by S&P 500 funds' parent companies. Panel A presents results of regressing parent-level narrative complexity variables on S&P 500 fund narrative complexity variables. Panel B presents results of regressing parent-level structural complexity on S&P 500 fund structural complexity. Panel C presents results of regressing parent-level *Fees* on parent-level narrative complexity variables. Panel D presents results of regressing parent-level *Fees* on parent-level structural complexity. Panels A and B use OLS regressions. Panels C and D use OLS or robust regressions, as indicated at the bottom of each column. Year fixed effects are included in all models. All variables are defined in Appendix A *t*-statistics are in parentheses. Standard errors are clustered by fund. *** indicates significance at 1%; ** at 5%; and * at 10%.

Panel A: Regression of Parent-Level Average Narrative Complexity on S&P 500 Fund Complexity

	Dependent variable: Parent-Level Average Complexity			
	Length	*WordsPerSentence*	*Length_ExpDisc*	*WordsPerSent_ExpDisc*
	(i)	(ii)	(iii)	(iv)
S&P 500 Fund Narrative Complexity	0.759***	0.470***	0.824***	0.782***
	(15.51)	(8.77)	(8.09)	(8.32)
Adjusted R^2	0.645	0.513	0.674	0.606
Observations	424	424	123	123

Panel B: Regression of Parent-Level Average Structural Complexity on S&P 500 Fund Complexity

	Dependent variable: Parent-Level Average Complexity	
	Structural_Complexity (Full Prospectus Sample)	*Structural_Complexity* (Expense Disclosure Sample)
	(i)	(ii)
S&P 500 Fund Structural Complexity	0.626***	0.663***
	(7.32)	(4.91)
Adjusted R^2	0.387	0.432
Observations	424	123

Panel C: Parent-Level Analysis of Narrative Complexity

	Dependent variable: Parent-Level *Fees*							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Length	-0.001		0.011					
	(-0.01)		(0.16)					
WordsPerSentence		0.039		0.030				
		(0.78)		(0.64)				
Length_ExpDisc					0.414***		0.422***	
					(5.84)		(6.47)	
WordsPerSent_ExpDisc						0.411***		0.427***
						(5.17)		(5.28)
Estimation	OLS	OLS	Robust Regression	Robust Regression	OLS	OLS	Robust Regression	Robust Regression
Adjusted R^2	(0.016)	(0.013)	(0.009)	(0.008)	0.409	0.402	0.455	0.445
Observations	424	424	424	424	123	123	123	123

Panel D: Parent-Level Analysis of Structural Complexity

	Dependent variable: Parent-Level *Fees*			
	(i)	(ii)	(iii)	(iv)
Structural_Complexity (Full Prospectus Sample)	0.475***	0.488***		
	(10.16)	(12.78)		
Structural_Complexity (Expense Disclosure Sample)			0.475***	0.508***
			(7.09)	(8.38)
Estimation	OLS	Robust Regression	OLS	Robust Regression
Adjusted R^2	0.596	0.724	0.544	0.684
Observations	424	424	123	123