Mutual Funds and Risk Disclosure: Information Content of Fund Prospectuses*

Jonathan Krakow†, Timo Schäfer‡

January 15, 2020

Abstract

In this paper, we shed light on the informational value of U.S. mutual fund risk disclosures by examining the qualitative content of funds' prospectuses. First, approximating a fund's disclosed riskiness by the amount of risk disclosure, we find that funds inform extensively about their idiosyncratic risks and less about their systematic risks. Second, using methods from textual analysis, we document that around one-third of the variation in the content of funds' risk disclosure is fund-specific, while a substantial part of a fund's risk disclosure is determined at the fund group level. Third, we show that regular updates of the disclosed risks have predictive power concerning funds' alphas, indicating that risk statements provide information in predicting returns in the cross-section. Our results suggest that updates from outperforming funds provide valuable information about future performance, while we find a negative, less pronounced effect for underperforming funds. Finally, using an event study framework combined and relying on matching techniques, we find no significant changes in investors' behavior after the voluntary delivery of a simplified disclosure document, i.e., a summary prospectus, in addition to the full prospectus.

JEL classification: D8, G14, G23

Keywords: Mutual Funds, Risk Disclosure, Textual Analysis, Fund Performance

*We wish to thank Jules van Binsbergen, Charles Calomiris, Marco di Maggio, Miguel Ferreira, Mike Gallmeyer, Michael Hasler, Alexander Hillert, Pedro Matos, Per Östberg, Denitsa Stefanova (discussant), Alexandre Ziegler, and seminar participants at the University of Frankfurt, the University of Melbourne as well as the University of Zurich for very valuable comments. The authors declare no conflicts of interest.

†Department of Banking and Finance, University of Zurich, Plattenstrasse 32, CH-8032 Zurich, Switzerland. e-mail: jonathan.krakow@bf.uzh.ch. Jonathan Krakow acknowledges financial support from the Forschungskredit of the University of Zurich, grant no. [FK-17-015].

‡Department of Business Informatics and Information Economics, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, D-60323 Frankfurt am Main, Germany, e-mail: t.schaefer@wiwi.uni-frankfurt.de

1. Introduction

The Securities and Exchange Commission (SEC) in the U.S. requires managing firms of mutual funds to disclose qualitative information about several fund characteristics in their yearly updated full prospectuses. Besides the provision of quantitative information on a fund's risk, performance, and expenses, the value of qualitative fund disclosure provided in funds' prospectuses has not been investigated in the academic literature in spite of the large mutual funds market in the U.S., i.e., $17.7 trillion of total net assets invested in mutual funds at the end of 2018.[1] According to two studies conducted by the Investment Company Institute (ICI), the information disclosed in a fund's prospectus is a relatively important source of information for investors. For instance, more than 90% of mutual fund–owning U.S. households review the investment objective and risks associated with investing in a particular fund, and 34% of mutual fund shareholders rely on information in the prospectus to make a final investment decision (Investment Company Institute 2006; 2019).

In February 2009, the SEC enacted a rule on disclosing key information in a summary section at the beginning of a fund's full prospectus (SEC, 2009). The aim of introducing such a summary section is to provide particularly retail investors with key information in a user-friendly and concise format, making it easier for them to find, compare, and interpret information related to a fund's risks and performance to make an informed investment decision. The summary section informs qualitatively about, for instance, a fund's principal risks, investment strategy, and primary objective. In this paper, we use methods from textual analysis to scrutinize the qualitative content provided in summary sections in funds' full prospectuses by focusing on funds' risk disclosures. To this end, we analyze a sample of U.S. open-end mutual funds between 2011 and 2018 that report a summary section in their full prospectuses.

First, we investigate whether the amount of disclosed risk in a fund's prospectus corresponds to a fund's exposed risk through its portfolio holdings and what, in general, are determinants of the heterogeneity in the extent of disclosures. We approximate the amount of disclosed risk by the number of words that appear in a fund's risk statement. The SEC recently pointed out that the length of risk statements within one fund category might vary a lot.[2] To shed more light on this, we relate mutual funds' disclosed risk to funds' general risk, and additionally for U.S. equity funds, to their idiosyncratic as well

[1] For a detailed view on the annual statistics, see "2019 Investment Company Fact Book", available at https://www.icifactbook.org/deployedfiles/FactBook/Site%20Properties/pdf/2018/2018_factbook.pdf.

[2] See, for instance, the "Request for Comment on Fund Retail Investor Experience and Disclosure", available at https://www.sec.gov/rules/other/2018/33-10503.pdf

as systematic risk. Similar to the SEC, we document that the amount of disclosed risk varies significantly over time and in the cross-section. We find that, in general, a fund's amount of disclosed risk can be positively related to its idiosyncratic and general risk while controlling for a set of fund characteristics and fund category and group fixed effects. However, we document a small negative, statistically insignificant relationship with a fund's systematic risk when using within fund group and category variation. Moreover, we find that a substantial part of the variation in the amount of disclosed risk is determined at the fund group level, suggesting that funds within a fund group report, on average, risk statements of similar length.

Second, to shed more light on the influence of a fund's group and category regarding funds' risk disclosures, we use methods from textual analysis to compare the actual written content of funds' risk disclosures – instead of relying exclusively on the length. We compute median similarities of funds' disclosures within fund groups, categories, the whole sample, and random pairs. If prospectuses provide valuable information, we would expect that the similarity of prospectuses within a fund category is higher than within the same fund group, as funds within the same category tend to have a similar risk profile. Our results show that within fund group similarity is, on average, four times as high as within fund category similarity. Additionally, we investigate the comparability of funds' qualitative disclosures by computing the median of a fund's similarity with all other funds in a given year. Put differently, we analyze why some funds' disclosures are, on average, similarly different from all other funds in our sample. We decompose the variation in funds median similarities into year, fund group, and category fixed effects to quantify the contribution of the respective fixed effect. Our results show that for all risk, primary objective, and strategy narratives, the fund group effects dominate with 46.4%, 51.1%, and 43.6% of the variation, respectively. Yet, the idiosyncratic contribution (from the residuals) amounts to 31.9% in the risk disclosure analysis, which suggests that at least one-third of the variation in the median similarity is due to fund-specific characteristics.

Third, to understand the value of specific disclosed information in more detail, we analyze the value of disclosed risks over time concerning funds' performance. This analysis is in addition motivated by a guidance update provided by the SEC in 2016 that points to the relevance of timely updates of funds' risk statements with respect to changes in market conditions a fund is exposed to.[3] In particular, we refer to regular updates of the content in risk statements – measured by the relative difference a fund's risk statements between two consecutive years. We find that risk statements provide relevant information in explain-

[3] See the guidance update by the SEC, March 2016, available at https://www.sec.gov/investment/imguidance-2016-02.pdf.

ing returns in the cross-section controlling for funds' past performance. More precisely, funds that experienced overperformance in the prior year provide valuable information in their updates, explaining future performance. Funds, however, that underperformed in the past, use these updates to provide their investors an indicator of activism, and therefore, results suggest that management of those funds gets in particular active regarding content changes in the fund risk statement when the fund continues relatively underperformances.

Finally, while having analyzed the informativeness of risk disclosures in prospectuses so far, we investigate investors' reactions in light of the information disclosure hypothesis, i.e., whether funds are rewarded for voluntarily disclosing additional information in a simplified document. On the contrary, investors might not respond to it as the provided information has already been disseminated to them through other channels. We refer to a fund's summary prospectus as an additional separate filing that contains exactly the same information as the summary section in the full prospectus in a more compact format, but the provision of a summary prospectus is not mandatory (SEC, 2009). We exploit variation due to funds' voluntary introduction of a summary prospectus to provide causal evidence on investors' response to funds that file a summary prospectus, which provides more concise key risk- and performance-related information than the full prospectus. Since the decision to publish a summary prospectus is neither random nor exogenous, we rely on a propensity score matching estimator to rule out differences in fund characteristics in which treatment (publishing a summary prospectus) and control groups differ. Controlling for the different characteristics, we find that the introduction of such a summary prospectus in itself has only a small positive significant effect on fund flows of funds publishing a summary prospectus. This finding holds when we control for total net asset standardized fund flows as well as for rank standardized fund flows. Therefore, our findings confirm the results of Beshears, Choi, Laibson, and Madrian (2010) empirically.

Literature contribution

Our paper builds on several strands of the literature. First, we add to the recent literature on firms' risk disclosures by analyzing qualitative risk information on mutual funds' by their respective managing firms. Previous research that examines qualitative disclosures on risk focuses primarily on corporate risks. Campbell, Chen, Dhaliwal, Lu, and Steele (2014) analyze risk factor disclosures (RFD) in 10-K filings and conclude that the information that companies disclose in their risk factor section reflects the exposed risks. Chiu, Guan, and Kim (2018) confirm this finding and additionally show that RFDs contain information that is particularly relevant to credit investors. Contrary, Kravet and Muslu

(2013) find that RFDs in 10-K filings do not tend to cover firm-level risks and are more likely boilerplate. However, these contradictory findings are partially resolved by Bao and Datta (2014). Using a topic modeling approach for RFDs in 10-K filings, Bao and Datta (2014) show that only two-thirds of the identified risk types are informative to investors, and not necessarily all of them increase investors' risk perceptions. Similarly, Hope, Hu, and Lu (2016) show by introducing a computing algorithm to quantify the specificity of firms' qualitative risks disclosures that market reactions are positively associated with the quality of risk disclosures. In line with the empirical results, Heinle and Smith (2017) develop a theoretical model that considers the price effects of risk disclosure and offers a rationale for the present findings. They show that firms have an incentive to disclose risk-related information truthfully since firms can expect to reduce uncertainty premiums.

Second, our paper relates to the literature that considers the impact of mandatory disclosures in the fund industry. Brown, Goetzmann, Liang, and Schwarz (2008) examine empirically the value of information about operational risk and conflict of interest variables disclosed by hedge funds through the SEC. They find mixed evidence for the effectiveness of risk disclosure. The results suggest that financial institutions and well-informed investors may receive little additional information from mandatory disclosures, while individual investors benefit from it. Kozup, Howlett, and Pagano (2008), as well as Beshears et al. (2010), investigate the impact of introducing a summary prospectus on investors' fund selection experimentally. Results from Kozup et al. (2008) suggest that providing a summary prospectus has an impact on investors' fund evaluations and risk perceptions. In contrast, Beshears et al. (2010) find no evidence that the summary prospectus affects the fund choices of investors. Similarly, Choi, Laibson, and Madrian (2009) conduct a laboratory experiment and find that providing additional information does not affect investors' choices substantially.

Finally, this paper contributes to the literature on the managerial behavior of fund managers and the following reaction of investors. Jain and Wu (2000) and Gallaher, Kaniel, and Starks (2015) show that funds do not experience superior performance after the advertising of a fund, but attract significantly more money in comparison with a group of control funds. Cooper, Gulen, and Rau (2005) examine how investors react to changes in fund names, finding abnormal positive fund flows with no improvement in performance. Hillert, Niessen-Ruenzi, and Ruenzi (2016) provide evidence that investors react to the writing style of shareholder letters, which is another written communication channel of fund managers with investors besides prospectuses. Further, the literature shows that funds change their risk-taking behavior over time (e.g., Brown and Goetzmann, 1997; Huang, Sialm, and Zhang, 2011).

The remainder of the paper is structured as follows. Section 2 provides background information regarding the introduction of a summary section of U.S. mutual funds in their prospectuses and describes the data. Section 3 analyzes determinants of funds' disclosed amounts of risk. Then, Section 4 examines the drivers of the specificity and comparability of funds' risk disclosures. Section 5 provides the analysis of year-to-year updates in funds' risk disclosures regarding future performance. Section 6 investigates the effects of publishing a summary prospectus on investors behavior. Finally, Section 7 concludes.

2. Background and Data

In this section, we first provide background information on the mandatory introduction of a summary section in the U.S. in mutual fund prospectuses. We review the motivation for the enactment of this simplify rule, and provide some details on the structure of the content. Then, we describe how we obtain our data and preprocess them. Finally, we provide summary statistics on the main financial characteristics of mutual funds in our example as well as a description of the textual data from the prospectuses that we link to the financial data from Morningstar.

2.1. Background

The SEC requires that all investment companies registered under the Investment Company Act of 1940 and to offer their share under the Securities Act of 1933 provide a statutory prospectus in which they inform their shareholders about their activities and to protect (potential) investors by disclosing relevant information about the offered securities.[4] These statutory prospectuses are often criticized as long and complicated, using a complex and legalistic language, which results in difficulties for investors to use them efficiently and to compare different funds. Accordingly, in April 2009, the SEC adopted a new disclosure framework that requires mutual funds to provide a summary section in their statutory prospectuses for securities that they offer for sale (SEC, 2009).[5] This summary section is intended to provide investors who are overwhelmed by the choices among thousands of available funds with useful and understandably key information that is el-

[4]The statutory prospectus is defined in the Securities Act of 1933, see http://legcounsel.house.gov/Comps/Securities%20Act%20Of%201933.pdf.

[5]Since June 1998, open-end mutual funds had to provide a risk-return summary. However, this information was disclosed on a fund group level and, therefore, not tailored to a fund's specific investments. With the adaption of the new rule in 2009, investment companies had to inform about the risk of each fund individually.

ementary to an informed investment decision. Moreover, a fund has in order to react to changes over time at least once per year to update the disclosed information.

The information provided in the summary section is at the front of the statutory prospectus, in a standardized order and contains streamlined details on the following dimensions: (1) investment objectives; (2) costs and fees; (3) principal investment strategies, risks, and performance; (4) investment advisers and portfolio managers; (5) brief purchase and sale and tax information; and (6) financial intermediary compensation.

Importantly, if a fund group issues several funds, the investment company has to publish for each fund a separate summary section. Many funds are also offered in multiple share classes. Since all share classes, however, are invested in the same underlying portfolio, this separation of individual summary sections does not apply to multiple share classes. All share classes of a fund are exposed to the same risks and differ primarily in their fees, e.g., management fees and 12b-1 fees. Hence, the summary section describes the characteristics of one fund but can describe several share classes of this fund, e.g., regarding costs and performance.

Finally, to simplify disclosure documents, the SEC provides mutual funds with the option to write a separate *summary prospectus* additionally, which contains the same information as the summary section in the statutory prospectus. While writing a summary section in the statutory prospectus is mandatory, the additional delivery of a summary prospectus is not (about 85% of all mutual funds offered a summary prospectus to investors by the end of 2018). The summary prospectus contains the same information (in the same order) as required in the summary section of the statutory prospectus that requests information on investment objectives, expenses, and fund's exposed risks. In our analysis, we will focus on qualitative information provided in the "Principal Risks" section, i.e., narrative risk disclosure, that contains information on *" risks to which the Fund's portfolio as a whole is subject and the circumstances reasonably likely to affect adversely the Fund's net asset value, yield, and total return"*.[6] See Section A in the Appendix for two examples of funds' risk disclosures.

Since the SEC considers these risk disclosures as eminent for investors, the SEC regularly publishes guidance to improve risk disclosures regarding accuracy and conciseness. For instance, the SEC staff often observes too long, legalistic, and complex risk narratives that do not reflect a concise summary of the principal risks.[7] Besides, the SEC remarks that investment companies shall review risk disclosures when market conditions change.[8]

[6] Item 4(b) of Form N-1A, see https://www.sec.gov/files/formn-1a.pdf

[7] See the guidance update by the SEC, June 2014, available at https://www.sec.gov/investment/im-guidance-2014-08.pdf.

[8] For instance, see the guidance update by the SEC, March 2016, available at

2.2. *Data*

2.2.1. *Data collection*

We obtain data on mutual funds' prospectuses from the SEC "Mutual Fund Prospectus Risk/Return Summary Data Sets" (MFSD).[9] According to the Securities Acts of 1933 and 1940, investment companies registered with the SEC have to disclose relevant investment information about the offered securities. Since January 2010, mutual funds have also to provide a summary section about their investments. This dataset provides quarterly updates of text and numeric information that is extracted from the risk/return summary sections of mutual fund prospectuses. The values of some text variables, however, are truncated since the maximal length of values is set to 2048 bytes. Hence, to circumvent the problem of truncated text data, we download prospectus data separately, i.e., filings 497[10] (initial registration) and 485BPOS[11] (post-effective amendment), that is specified in the MFSD dataset. Note that for some funds, we observe several 485BPOS filings per fund and year. In these cases, in line with Baghai, Becker, and Pitschner (2019), we only keep the longest risk statement for each fund and year.[12]

2.2.2. *Data preprocessing*

Prospectus data is provided in the eXtensible Business Reporting Language (XBRL) format. The extraction of relevant sections in a prospectus can be identified due to the XBRL tree and key structure. For each fund and year, we extract the most recent data on a fund's risk statement, strategy narrative, and primary objective. We clean text data from html code, stem words, remove stopwords, and convert numeric values into <*numeric*>.

2.3. *Summary statistics*

The final sample comprises all matched U.S. open-end mutual funds regulated under the supervision of the SEC from 2009 to 2018, covering a total of 8,194 funds. For our empirical analysis regarding the informativeness of risk disclosures, we merge mutual funds' prospectus data from the EDGAR database from the SEC with the Morningstar survivorship bias free mutual fund database. The Morningstar database includes information on

https://www.sec.gov/investment/im-guidance-2016-02.pdf.

[9] The dataset is available at https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets.

[10] https://www.law.cornell.edu/cfr/text/17/230.497

[11] https://www.law.cornell.edu/cfr/text/17/230.485

[12] Regarding the length, we refer to the number of words after removing the html code from the filings, while Baghai et al. (2019) refer to the number of sentences.

fund returns, various risk measures, as well as several mutual fund characteristics, like total assets under management, costs, and security holdings.

To merge the prospectus data from the SEC with the database in Morningstar, we establish a unique link between a fund's ticker symbol and the Series ID assigned from the SEC to each individual fund. Many funds are offered in multiple share classes. However, all share classes of a fund are exposed to the same risks in the portfolio. Hence, since the Morningstar database provides data by share class, we aggregate a fund's share classes characteristics into a single fund characteristic by taking the market value-weighted average (see e.g., Gallaher et al. (2015)). For the mapping (see Table 10 in the Appendix for details on the matching process), we use the investment companies file of the SEC to identify 31,796 tickers of share classes under their regulation.[13] 13,553 of those share classes have not a ticker symbol in this file. Excluding funds with missing ticker symbols in the Morningstar database reduces our sample to 39,138 share classes, which results in 8,194 funds and their corresponding prospectuses. Furthermore, we drop observations where a fund's total net assets in a given month are smaller than 1 million USD. After matching, we conduct several plausibility tests to ensure that the data in Morningstar correspond to the funds identified in the SEC reports.[14]

Table 1 reports the sample summary statistics on company size, turnover, return, risk, and various holding measures in our sample from 2009 to 2018. We also include a fund's age (in months), its expense ratio, and a fund's size to analyze whether more established funds, more expensive funds, and larger funds differ in their risk disclosure behavior. Next, based on data on total net assets and returns, we compute the monthly fund net flows, which are defined as the growth rate of the total assets under management adjusted for reinvested returns. We follow the methodology of Huang et al. (2011) and calculate it formally as:

$$Flow_{f,t} = \frac{TNA_{f,t} - TNA_{f,t-1}(1 + RF_{f,t})}{TNA_{f,t-1}(1 + RF_{f,t})}, \tag{1}$$

where $TNA_{f,t}$ is fund f's total net assets at time t and $RF_{f,t}$ is the fund's return over the prior month.[15] To account for outliers, we winsorize at both the bottom and the top parts

[13]The basic identification information on investment companies year by year is available at https://www.sec.gov/open/datasets-investment_company.html.

[14]For instance, we test whether there is a unique link between a fund's Series ID and assigned Fund ID in Morningstar. Since both identifiers are unique at the fund level, we exclude cases where Series IDs are assigned to multiple Fund IDs.

[15]This measure ensures that fund flows cannot be below -100%. Most studies define fund's net flows as Flow $= (TNA_{f,t} - TNA_{f,t-1}(1 + RF_{f,t}))/(TNA_{f,t-1})$. However, our approach, following the methodology of Huang et al. (2011), guarantees that fund outflows will not be below -100%. The Pearson correlation coefficient between the two measures is 0.9998.

of the distribution at the 1% level.

[Insert Table 1 about here]

The average fund in our sample has total net assets of USD 1,741 million, is slightly below 14 years old, and has an annual expense ratio of 0.69%. Regarding fund's holdings and the asset allocation, we observe that the average fund invests more than 61% in equity, about 30% in fixed income, and almost 6% in cash. The investor's average return of mutual funds in our sample is 0.36% per month, while the average fund' risk, estimated as the standard deviation of the fund's monthly returns over the prior 36 months, is 1.90.

To investigate the characteristics of funds' prospectuses both in the cross-section and over time, we extract all risk statements from fund prospectuses between 2011 and 2018.[16]

Table 2 presents summary statistics of the length of risk statements over time.

The average length of a fund's risk statement in 2011 is about 790 words. In contrast, the average risk statement comprises about 1,550 words in 2018, meaning that the length of risk statements nearly doubles relative to 2011 in 8 years. Additionally, one observes large heterogeneity in the cross-section concerning the number of words funds use in the respective risk statements. While some of the risk statements include only a few dozens words, others contain more than several thousand words.

[Insert Table 2 about here]

To understand this considerable variation in the cross-section, we examine in a first test the relation of a fund's exposed risk with the amount of disclosed risk. In general, we would expect that mutual funds that are exposed to more substantial risks in their portfolios also report more extensively about it in their risk disclosures.

3. Risk disclosure and fund risk

We start our analysis by investigating the relation of disclosed risk in the summary sections of funds' prospectuses and the corresponding risks funds are exposed to via their portfolios. In particular, we examine whether risks disclosed in textual form are informative, i.e., whether they are positively related to variables of risks in funds' portfolios. To address this research question and to analyze the relationship, we use, as a first proxy for

[16] Other papers that are using information from funds' prospectuses in their analysis are Abis (2017) and Baghai et al. (2019). Abis (2017) applies machine learning methods to funds' prospectuses to classify funds as either quantitative or discretionary. Baghai et al. (2019) apply methods from textual analysis to funds' prospectuses to obtain a proxy for the use of credit ratings in funds' investment mandates.

the informativeness of disclosed risks, the length of the risk statements.[17] The length of a fund's risk statement and the number of disclosed risk factors are positively correlated, meaning that longer risk statements contain more risk factors.

[Insert Figure 1 near here]

Figure 1 shows the distribution of the length of the risk statements across all published risk statements in 2018. In line with the comment from the SEC, we document in our sample a significant heterogeneity in the length of risk statements, i.e., a risk statement's range of 200 vs. more than 7,000 words. Accordingly, we ask whether the exposed risk of a fund's portfolio, assuming that riskier funds provide longer risk disclosures, explains this heterogeneity?

To investigate this hypothesis, we relate in a first regression the amount of disclosed information as a dependent variable to a fund's general exposed portfolio risk. Following the approach suggested by the literature on mutual funds, we measure the general risk of a fund (*Fund Risk*) by calculating the standard deviation of the returns over the prior 36 months as a risk measure, see e.g., Huang et al. (2011). Second, for all U.S. equity funds, we calculate both the *Systematic* and the *Idiosyncratic Risk*. In line with Huang et al. (2011), we use the market model (CAPM) beta as a proxy for a fund's systematic risk since it measures the relationship between the fund's excess returns over Treasury bills and the excess returns of the benchmark index. We compute the beta in the market model based on the prior 36 months fund returns. To quantify a fund's idiosyncratic risk, we calculate the part of the variation in the returns that cannot be explained by the variation in the returns of the market model as $1 - R^2$.[18]

3.1. Fund's general risk

To examine the determinants of a fund's disclosed risk formally, we start by a simple OLS regression where we regress the length of a risk statement – the amount of disclosed risk – as dependent variable on the general risk variable and control for several fund as well as fund group characteristics like fund flows and returns over the prior period, fund

[17] This proxy is in line with previous research (Campbell et al., 2014; Célérier and Vallée, 2017; Ghent, Torous, and Valkanov, 2017). For instance, concerning financial products, Célérier and Vallée (2017) use the length of descriptions of structured products as an approximation for product complexity, and Ghent et al. (2017) take the number of pages of a mortgage-backed security prospectus as an input to estimate product complexity. Regarding corporate disclosure, Campbell et al. (2014) calculate the length of companies' risk factor disclosures in annual reports to test whether this measure correlates with pre-disclosure information on firm risk.

[18] We multiply the idiosyncratic risk measure with 100 so that it ranges from 0 to 100.

size, expense ratio, turnover ratio, age, and a dummy that indicates whether a fund files a summary prospectus (1) or not (0). Hence, we estimate the following panel regression model:

$$\begin{aligned}\text{Length}_{f,t} =&\beta_0 + \beta_1 \text{Fund Risk}_{f,t} + \beta_2 \text{Fund Flow Period}_{f,t} + \beta_3 \text{Return Period}_{f,t} \\ &+ \beta_4 \log(\text{Fund Size}_{f,t}) + \beta_5 \text{Expense Ratio}_{f,t} + \beta_6 \text{Turnover Ratio}_{f,t} \\ &+ \beta_7 \log(\text{Age}_{f,t}) + \beta_8 \text{SP Dummy}_{f,t} + \nu_{category} + \nu_{fundgroup} + \nu_t + \epsilon_{f,t},\end{aligned} \tag{2}$$

where $\text{Length}_{f,t}$ is the number of words at time t in fund's f risk statement. $\text{Fund Risk}_{f,t}$ is the respective risk variable. Further, we control for the fund flow and the return in the year, the fund size, the expense ratio, the turnover, the age, and a summary prospectus dummy (SP Dummy). $\nu_{category}$, $\nu_{fundgroup}$, and ν_t are fund category, fund group, and year fixed effects, respectively, and we cluster standard errors at all the fund group, and year level. The coefficient of interest β_1 captures the marginal effect of the fund's respective risk on the amount of risk disclosure. It allows us to test whether the disclosed risk is positively aligned with the fund's exposed risk dimension.

The inclusion of fund category fixed effects allows us to estimate β_1 by exploiting variation of disclosed risk of funds within a specific category. The amount of disclosure might vary substantially across different fund categories, e.g., small-cap stock funds vs. municipal bond funds, that can be distinguished with regards to their risk-taking behavior by construction. Moreover, we add fund group fixed effects, taking into account that different investment companies – independent of a fund's exposed risk – provide varying amounts of (risk-related) information. The inclusion of a fund group fixed effect incorporates the hypothesis that the decision of the amount of disclosed information in a risk statement is made at the investment company level.

Table 3 shows the results from estimating Equation 2 for the fund's general risk ($\text{Fund Risk}_{f,t}$) with year fixed effects. Results of the first specification (1) in Table 3 provide evidence that the exposed risk of a fund is negatively related to the amount of disclosed risk. A one-unit higher risk measure results, on average, in a "56" words shorter risk statement. Hence, the higher the exposed risk of a fund, the shorter is the risk statement, controlling for various fund characteristics. This result stands in contrast to our hypothesis that riskier funds inform more about their exposed risk in their risk statements.

In a second (2) and third (3) specification, we add category fixed effects and alternatively fund group fixed effects to our control variables. In both specifications, the R-squared increases from 0.140 in the specification (1) to 0.348 and 0.637, respectively,

pointing to the relevance of considering funds' categories and unobserved characteristics at the fund company level to explain variation in the amount of risk disclosure. In particular, the increase in the R-squared when including group fixed effects illustrates that the length of the disclosed risk statements is considerably determined at the fund group level and not only at the fund level. Thus, the length of a fund's risk section is, to a large extent, determined on the fund group dimension independent of a fund's risk.Furthermore, we confirm in both specifications a negative relationship between the exposed general risk of a fund and the amount of disclosed risk. The magnitude of the coefficients is smaller but still highly significant. Finally, specification (4) controls for all fixed effects. In this specification, we observe a positive relationship between a fund's exposed risk and amount of disclosed risk, implying that riskier funds within the same fund category and within the same fund group write more about risks in their risk statements when they are subject to higher risks. Thus, at this stage, our results suggest that a comparison of funds disclosed risk across categories and fund groups is rather difficult since we only document a positive relationship between exposed and disclosed risk when we control for those two dimensions.

Besides a fund's general risk exposure, we also included several other fund characteristics to the regressions as controls. We find for all four specifications that the coefficient of the age variable is negative and highly significant, and the coefficient of the summary prospectus dummy is considerably large and highly significant. These findings indicate that older funds write relatively shorter risk statements compared to younger funds and that the decision to provide a summary prospectus results in general in longer risk statements. Both findings are robust, independent of the inclusion of any fixed effects.

[Insert Table 3 about here]

As a robustness check of our findings, due to the skewness of the risk disclosure distribution in Figure 1, we run the same analysis with *log of words* as the dependent variable. Detailed results can be found in Table 12 in the Appendix. Similarly to our previous findings, we document a statistically significant negative relationship between exposed and disclosed risk as long as we do not control for category and fund group fixed effects for the same time. Finally, Table 13 in the Appendix documents a similar relationship for exposed and disclosed risk for equity funds only.

3.2. *Fund's idiosyncratic and systematic risk*

In a second analysis, we distinguish a fund exposed risk in an idiosyncratic and systematic risk part. More precisely, we test whether the idiosyncratic or the systematic risk of a fund can partially explain a fund's amount of disclosed risks. To obtain measures for these two risk dimensions, we use a subsample with US. equity funds only and estimate the systematic risk (*market beta*) as well as the idiosyncratic risk ($1 - R^2$) based on the estimations of a 1-factor CAPM.[19] We estimate the corresponding idiosyncratic and systematic risk measures based on the past 36 months returns. Table 4 presents the results from Estimating Equation 2 with idiosyncratic and systematic risk as independent variables.

Table 4 shows that the amount of disclosed risks is positively related to a fund's idiosyncratic risk. This finding amplifies – also in terms of statistical significance – when we control for fund group variation (specification (2)). Furthermore, this result is robust when recovering the marginal effect of a fund's idiosyncratic risk, β_1, using both within fund group as well as category variation in specification (3). Hence, the higher the idiosyncratic risk of an equity fund in our sample, the more information is disclosed in the fund's risk statement. Note, however, the main driver of the variation in the amount of disclosed risks are unobserved characteristics at the fund group level, which is revealed by the substantial increase in the R-squared from specification (1) to (2).

Conducting the same analysis with a fund's systematic risk as an independent variable, we find the opposite effect. Specifications (4), (5), and (6) in Table 4 provide evidence for a negative relation between systematic and disclosed risk. The results indicate that a higher systematic risk of a fund does not result in a longer but rather in a shorter risk statement. The coefficient β_1, however, decreases substantially in magnitude when controlling for a fund group fixed effect (specification (5)), and we do not find statistical significance anymore when we control for both fund group and category fixed effects (specification (6)). This finding, however, is not surprising since the categorization of a fund, by definition, provides already information on the systematic risk of a fund. Concerning the explanatory power of our model specifications, we document again a substantial increase in magnitude when controlling for a fund group fixed effect as the R-squared in Table 4 indicate.

[Insert Table 4 about here]

Overall, we find that a fund's risk statement, approximated by the length of these risk

[19] Table 11 in the Appendix shows the summary statistics for equity funds only.

statements, is reasonably informative about a fund's exposed risk, meaning that each of the three risk measures has a significant impact on the amount of disclosure. In particular, we document that funds disclosed risk is positively driven by the general and idiosyncratic risk of a fund and negatively with the fund's systematic risk. The effect of systematic risk on the amount of disclosure becomes insignificant when estimating the relationship using within category and fund group variation. However, the results in Tables 3 and 4 also show that much of the variation in the disclosure amount is due to unobserved characteristics at the fund group and category level.

4. Similarity in prospectuses

The preceding section has provided evidence on the informativeness of the amount of disclosed risks concerning the general and idiosyncratic risk of a fund. For this matter, we use the length of the risk statements as a proxy for disclosed riskiness of a fund. However, our results indicate that a large part of the variation in the disclosure amount can be traced back to unobserved characteristics at the fund group and category level. To shed more light on this finding, we examine the specificity of the actual *written* content in funds' prospectuses and fund characteristics that drive variations in the disclosure specificity. For instance, in a recent report, the SEC points out that "several fund groups [that] tailor their risk disclosure for each fund in the fund group rather than rely on generic, standardized, risk disclosures across funds."[20]

We refer to standardized textual disclosure that is highly similar across funds as boilerplate. To analyze content-based differences in funds' textual disclosures, we compute the cosine similarity between two funds' f and f' disclosures to quantify the distance between the two text documents without considering the order words in a document[21]:

$$\text{sim}_{f,f'} = \frac{\vec{V}(\mathbf{d}_f)\vec{V}(\mathbf{d}_{f'})}{||\vec{V}(\mathbf{d}_f)||_2||\vec{V}(\mathbf{d}_{f'})||_2} \quad \in [0,1], \tag{3}$$

where $\vec{V}(\mathbf{d}_{f'})$ is the word vector of fund f in vector space notation using term frequency - inverse document frequency (tf-idf) weights and $||\cdot||_2$ is the L_2 norm for normalization. Intuitively speaking, the cosine similarity tells us how different two text documents are

[20] See https://www.sec.gov/investment/accounting-and-disclosure-information/principal-risks/adi-2019-08-improving-principal-risks-disclosure.

[21] For instance, Brown and Tucker (2011) measure year-by-year cosine similarities in a company's Management Discussion and Analysis (MD&A) disclosures in 10-K filings to examine the informativeness of the MD&A section.

by calculating the angle between the two word vectors.[22] The cosine similarity takes a value of 0 (1) when two documents are entirely different (identical). In contrast to a uniform weighting of words using the vector space operator $\vec{V}(\cdot)$, the tf-idf schema takes into account that some words appear in a large (low) number of prospectuses and get a lower (higher) weight. We do not explicitly risk factors a fund describes in the risk statement. Our approach, however, has the advantage that we can directly compare various risk statements with each other, independent of the fact that there do not exist universal definitions for each risk category to which investment companies refer to in their disclosures.

First, we investigate whether, on average, the textual similarity of funds within a fund group or within a fund category is more extensive. In line with Brown and Tucker (2011) and Lang and Stice-Lawrence (2015), we compute the similarity within a group by calculating the similarity of a fund and any other fund in the respective group – i.e., sample, category, fund group (CIK) – and take the mean of all pair-wise similarities. As the similarity of two funds from the same managing company might be upward biased due to similarity in a company's writing style, we use only stemmed (proper) nouns in funds' prospectuses and hence, do not consider any stylistic words.[23] If prospectuses are informative on average, one would expect that the average similarity of prospectuses within a category is relatively high and that the similarity of funds within the same fund group is relatively low. In other words, funds in the same category are similar regarding the risk dimensions they are exposed to because of comparable investment strategies and risks in their portfolios, and thus, we expect within category variation to be relatively higher than within fund group variation.[24] Figure 2 illustrates the results of the median similarity within fund groups, categories, all funds, and a random pair of funds for each year in our sample.

[Insert Figure 2 near here]

The median similarity between a random pair of funds (black line) is the lowest across

[22] For instance, the vector space operator $\vec{V}(\cdot)$ transforms the two sentences $\mathbf{d}_1$="this is a risk"=("this", "is", "a", "risk")$'$ and $\mathbf{d}_2$="this is no problem"=("this", "is", "no", "problem")$'$ into word vectors in the same vector space, i.e., $\vec{V}(\mathbf{d}_1) = (1,1,1,1,0,0)'$ and $\vec{V}(\mathbf{d}_2) = (1,1,0,0,1,1)'$. Then, the first element of the vector $\vec{V}(\mathbf{d}_1)$ is the count of the word "this" in document $\mathbf{d}_1$, the second element refers to the word "is", and the sixth element to the word "problem". This transformation using $\vec{V}(\cdot)$ allows us to compare the two documents in the same vector space and quantify their similarity using the definition of cosine similarity in Equation 3.

[23] We identify (proper) nouns in funds' prospectuses via parts-of-speech (POS) tagging.

[24] As an assumption of this, we implicitly stipulate that a fund group consists of funds from different categories.

the four groups we consider, while the median similarity between all funds (green line) in a given year is similarly higher. The red line depicts the results for the median similarity within a fund category. Results imply that within-category similarity is, on average, higher than in the two groups with a random pair of funds and all funds. This finding supports somewhat the hypothesis that funds' risk disclosures within a category group capture more specific information than when supposing no or a random group structure. Finally, when considering the median similarity within a fund group (blue line), we find that the similarity is, on average, three times as high as within a fund category. This finding lends strong support to the hypothesis that similarities in funds' risk statements are highest within fund groups and, in particular, that they are significantly higher compared to the median similarity within a fund category.

Second, we conceive a fund's average deviation from other funds' prospectuses in a given year as disclosure specificity (comparability)) and analyze drivers thereof. As in Lang and Stice-Lawrence (2015), we compute the median of a fund's similarity with all other funds in a given year (instead of defining specific peer funds). In particular, we suppose that the more dissimilar a fund's disclosure is, on average, from other funds' disclosures, the more specific and informative a fund's disclosure is. Using this measure, we run a linear regression of a fund's median similarity on year, fund group, and fund category fixed effects to investigate how much of the variation in the median similarity of a fund is due to each of the fixed effects. We use fund group and category fixed effects to scrutinize the contribution of the fund group and fund category concerning the content similarity in prospectuses. Hence, we can estimate the contribution of the fund group and fund category concerning the content similarity in prospectuses, which allows us to decompose the variation in the median similarity into several fixed effects. Moreover, this regression exercise is sort of like a clustering problem, in which we aim to understand why some funds' disclosures are similarly different from all other funds and, hence, are comparable. Besides the risk disclosure section, to generalize our findings to other parts of a fund's prospectus, we also consider the primary objective and strategy statement subsections in the summary section of a fund's prospectus. Table 5 presents the results.

[Insert Table 5 about here]

We find that the most substantial part of the variation in fund's risk disclosures median similarity plays out at the fund group level (46.4%), while the contribution of a fund's category is only about a quarter of the size compared to the group level effect (11.9%). Nevertheless, the relative contribution of the residual is still comparatively high, which suggests that 31.9% of the variation can be traced back to fund's idiosyncratic effects,

i.e., variation within a category, fund group, and year. We report similar findings for a fund's primary objective and strategy narrative median similarity. The interpretation of our findings can be biased due to measurement errors, i.e., we might not completely rule out a fund group's writing style by only using stemmed (proper) nouns to compute pair-wise similarities.

5. Risk disclosure and fund performance

Our analysis so far indicates that content-based variation in qualitative risks funds disclose in summary sections of their prospectuses mainly plays out at the fund group level. Still, the amount of disclosed risks can be related to a fund's general and idiosyncratic risk exploiting within a fund group and category variation. Besides scrutinizing primarily cross-sectional effects of risk disclosure, recent academic research (see, e.g., Brown and Tucker, 2011) and regulators have pointed to the importance of timely updates of firms' disclosures in general. In particular, regarding the case of funds' prospectuses, the SEC states in a guidance update in 2016 that investment companies should review their risk disclosures on an ongoing basis and update their risk statements in light of current market conditions.[25] Results from Table 2 show that there is a monotonous increase in the average amount of risk disclosure from one year to the next and Table 14 in the Appendix shows summary statistics on year-to-year changes in the length of funds' risk statements.

Yet, it is not clear whether these updates are informative to investors or whether this average increase is mainly driven by pure actionism of the respective fund. To elaborate on the informativeness of updates, we compute difference scores of funds' risk disclosures on a year-to-year basis to determine the magnitude of updates and relate these changes to the future performance of a fund. More precisely, we calculate the similarity score for each prospectus of a fund in year t compared to the prospectus of a fund in $t-1$ using Equation 3. Then, in line with Brown and Tucker (2011), the difference score is 0 (1), when the similarity scores of the prospectuses in two successive years is 1 (0).

Next, to analyze the relationship between performance and disclosure updates over time, we predict yearly performance measures for all U.S. equity funds between 2012 and 2018.[26] In particular, we relate various performance measures of funds to differences in the disclosed information. The findings of this analysis will allow us to assess whether

[25] See the guidance update by the SEC, March 2016, available at https://www.sec.gov/investment/im-guidance-2016-02.pdf.

[26] Since we calculate the difference between two risk statements over time, we do not have an independent variable for 2011 as our sample of prospectuses starts in 2011.

funds update their risk statements (concerning the content of disclosure) in line with their future performance.

Table 14 in the Appendix presents summary statistics regarding the similarity scores in the risk statements over time. The results in Table 14 indicate considerable variation concerning updates in risk statements across both the fund and time dimension.

For the performance measures, we compute a fund's alphas, i.e., the CAPM alpha, the Fama and French (1993) 3-factor alpha, and the Carhart (1997) 4-factor alpha, for our set of U.S. equity funds. To estimate funds' alphas, we use funds' as well as stock market factors' returns on a daily basis over the prior twelve months.[27] Then, for our panel regressions, we use the estimated alphas as dependent variables that we regress on the text-based measure Diff Score$_{f,t}$, and a set of fund characteristics as control variables. Finally, we include a year as well as fund fixed effects to control for the impact of time and cross-sectional effects that are not filtered out. Hence, we estimate the following panel data model:

$$\begin{aligned}\hat{\alpha}_{f,t} =& \beta_0 + \beta_1 \text{Diff Score}_{f,t-1} + \beta_2 \log(\text{Age}_{f,t}) + \beta_3 \log(\text{Fund Size}_{f,t}) + \\ & \beta_4 \log(\text{Company Size}_{f,t}) + \beta_5 \text{Expense Ratio}_{f,t} + \\ & \beta_6 \text{Fund Flow Period}_{f,t} + \nu_t + \nu_f + \epsilon_{f,t}, \end{aligned} \tag{4}$$

for each of the performance measures. Diff Score$_{f,t}$ is the difference between two years at time t in fund's f risk statements. Further, we control for the fund age, the fund size, the company size, the expense ratio, and fund flow of that year. ν_t and ν_f are year and fund fixed effects, respectively, and we cluster standard errors at the year and fund level. In order to control for outperforming ($\alpha \geq 0$) and underperforming ($\alpha < 0$) funds, we extend the model in Equation 4 in a second specification by an interaction term of a dummy, that is 1 if $\alpha \geq 0$ and 0 otherwise, with the difference score, i.e., Diff Score$_{f,t-1} \times \mathbb{1}_{\hat{\alpha}_{f,t-1} \geq 0}$. The results of these regressions formulated in Equation 4 are presented in Tables 6 and 7, respectively.

[Insert Table 6 about here]

Table 6 reports no statistical and economic significant relationship between differences in disclosed information in risk statements in two consecutive years of a fund for all specifications. Moreover, we find that fund size has a positive effect on fund performance,

[27] We thank Kenneth French for providing data on stock market factors, available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

while fund flows in that period have a significantly positive effect on fund performance. Thus, at this stage, our results indicate that changes in risk statements have no explanatory power for fund performance.

To elaborate on this finding in more detail, we then differentiate between outperforming and underperforming funds. The idea behind this differentiation is that there might be two different motives for updates that cancel each out other when not controlling for the out- and underperformance in the year before. Non-performing funds might update because they want to show some action to their investors while performing funds update in order to adapt to changes in the market conditions. Hence, in Table 7, we present results that differentiate between those two types.

[Insert Table 7 about here]

Results in Table 7 show a much more differentiated picture regarding the explanatory power of updates in the risk statements and future fund performance. We find a positive statistically significant relationship between updates and outperforming funds and a slightly negative – and statistically significant – relationship between underperforming funds and their corresponding updates in the risk statements. Our results suggest that funds that experienced outperformance in the year before an update provide information in their risk statements that explain predictive performance, while funds that underperform use updates in their risk statements to show investors activism.

6. Simplified disclosure and fund flows

So far, in this paper, we analyzed the informativeness of risk disclosures in the summary section of funds' prospectuses. Yet, we have not considered investor reactions concerning the information in prospectuses' summary sections to examine whether the disclosed information provides investors with additional material information on their investment decision, expressed in increasing fund flows. As an alternative hypothesis, one might argue that the information provided has already been disseminated through other channels, and therefore, investors do not respond to the additional availability of key information in the summary section.

Hence, to emphasize the importance of the information in this summary section, the SEC has not only introduced a new disclosure framework requiring all mutual funds to adapt to it but also introduced the option for mutual funds to publish a separate summary prospectus, i.e., filing 497K, see SEC (2009). Note that a summary prospectus contains

exactly the same information as the summary section but is a separate document from the full prospectus.

In line with the academic literature (see e.g., Brown et al., 2008), we elaborate on the information disclosure hypothesis, i.e., whether funds are rewarded for voluntarily disclosing additional information. By additional information, we refer in our context to the additional provision of key information in a separate filing, i.e., the summary prospectus. The literature finds mixed evidence as Kozup et al. (2008) conclude that a simplified disclosure framework affects the investor's decision process while Beshears et al. (2010) do not find such an effect.

We measure the effect of introducing a summary prospectus on investors' behavior by examining changes in fund flows after the release of a summary prospectus. More precisely, to identify the causal effect of this reform, we exploit the fact that the release of a summary prospectus is not mandatory for a fund, and therefore, we can compare treated vs. non-treated (control) funds using an event study framework. Hence, we classify a fund to the treatment group when it starts to provide a summary prospectus in 2009 or 2010, and to the control group otherwise. Figure 3 plots the number of funds that decide to publish for the first time a summary prospectus in 2009 and 2010.

[Insert Figure 3 near here]

This setting, however, does not represent a random, exogenous assignment of funds to the treatment and control group, respectively. Instead, investment companies decide endogenously whether their funds belong to the treatment group or not, i.e., we observe self-selection of funds in a non-experimental setting. Accordingly, to compute *abnormal flows* of funds that introduce a summary prospectus in our sample, we use propensity score matching to make treated funds with control funds comparable across multiple characteristics that are relevant for investors' behavior and hence, for fund flows.[28] Our selection of covariates is motivated by Cooper et al. (2005) and Hartzmark and Sussman (2019), and comprises the 1-month lagged log of fund size, the cumulative returns as well as flows to the fund in the six months before the introduction of a summary prospectus, the log of the age of the fund in months, and the standard deviation of returns in the six months before.

Former literature provides evidence that fund flows are strongly related to past performance (Warther, 1995; Chevalier and Ellison, 1997). Moreover, this relationship is non-linear in the sense that investors react to positive returns in the recent past but do not sell poorly performing funds (Gruber, 1996; Carhart, 1997; Goetzmann and Peles, 1997; Sirri and Tufano, 1998; Lynch and Musto, 2003). In addition, we control for fund size,

[28] We use the propensity score matching procedure based on Sekhon (2011).

assuming that larger funds have more capacities to publish such a summary prospectus. Finally, we include the age of a fund, since results in Section 3 indicate that older funds provide less qualitative information compared to younger funds.

To estimate a propensity score for each fund, i.e., the probability to publish a summary prospectus, we estimate a logistic regression for each month in the period of interest by assigning the dependent variable for each treated fund a dummy of 1, and otherwise a 0. Panel A in Table 8 shows that treated funds significantly differ from funds in the control group at the 1% level for all covariates.

[Insert Table 8 about here]

Then, for each treated fund, a matching fund that does not publish a summary prospectus is identified as the fund with the closest propensity score to the treated fund. We use the Mahalanobis distance to compute the weighting of each covariate, and we allow for ties. To control for unobservable characteristics across time, we match treated funds in a specific month only with control funds in the same month. After running this exercise, Panel B in Table 8 shows the differences between each covariate's mean of the treatment and the matched control group. Results indicate that for all covariates there is, on average, no difference between the treatment and control group which implies appropriate covariate balance.[29] Abnormal fund flows are computed for each treated fund in our sample compared to their matching fund.

[Insert Figure 4 near here]

Figure 4 illustrates the monthly fund flows before and after the publication of a summary prospectus controlling for fund characteristics according to the propensity score matching estimator. Before the delivery of a summary prospectus, the funds in the two groups were receiving similar levels of flows. After the release, the treated funds experienced slightly higher fund inflows in the months of treatment, and these flows are slightly higher than the flows of the control group in most months after the treatment. To test this finding more formally, we then calculate the cumulative abnormal fund (CAF) flows over various event windows in the month(s) before and after the treatment. We then test the hypothesis whether these CAF are different from 0. Table 9 reports CAFs to all treated funds and provides the corresponding t-values.

[Insert Table 9 about here]

[29] We conducted several matching strategies to transform the matching problem into subgroups using variables such as the prospectus' benchmark or fund's category. However, since these variables are usually very granular, many subgroups are left with only very few observations resulting in a poor matching balance.

As reported in Panel A of Table 9, in the month(s) prior to the release of the summary prospectus, treated funds do not experience statistically significant abnormal flows. In the 6 months after the release of a summary prospectus, CAFs in excess to control funds are, on average, 1.20%, which is slightly statistically significant (t-value = 2.01). Over the complete year after the introduction, the CAF are, on average, 1.76% (t-value = 1.76). Thus, we document a small positive effect on CAFs for treated funds after the introduction of the summary prospectus. However, t-values indicate only some statistical significance for this finding. Therefore, we test whether this finding is robust to normalized fund flows. We normalize fund flows according to two suggestions from the literature (i) by dividing flows by TNA_{t-1}, see e.g., Cooper et al. (2005), and (ii) by allocating funds into deciles according to fund size and then ranking flows in percentiles within these deciles, see Hartzmark and Sussman (2019).

Panel B and Panel C in Table 9 show the corresponding results with respect to the cumulative abnormal normalized flows. We document again a statistically significant, but a rather small effect for size normalized fund flows. Moreover, we do not find any statistical evidence at all for CAFs different from 0 for the rank normalized fund flows. Therefore, in line with Beshears et al. (2010), our results imply that the additional provision of information disclosed via summary prospectuses does not (or only slightly) affect the investment decisions of investors.

7. Conclusion

In this paper, we examine the informational value and the drivers thereof in risk disclosures of U.S. mutual fund prospectuses. To this end, we use methods from textual analysis to analyze a sample of U.S. open-end mutual funds between 2011 and 2018 that report their principal risks to which they are exposed to in the summary section of their full prospectuses.

The starting point is to scrutinize whether funds' risk disclosures reflect the exposed risks in their portfolio holdings. Therefore, we approximate the amount of disclosed risks by the number of words that appear in a fund's risk statement. We provide evidence that a fund's disclosed risks increase with a fund's idiosyncratic and general risks, but decreases – though not significantly – with its systematic risk while controlling for a set of fund characteristics as well as fund category and fund group fixed effects. Moreover, we find that a large part of the variation in length plays out at the fund group level.

Second, to analyze in more detail the impact of a fund's group and category on funds' risk disclosures, we compute similarities of the written content of funds' risk disclosures.

We document that around one-third of the variation in the content of funds' risk disclosure is fund-specific, while at the same time a substantial part of a fund's risk disclosure is determined at the fund group level.

Third, we show that regular updates of funds' disclosed risks have predictive power concerning funds' alphas, indicating that risk statements provide information in predicting returns in the cross-section. Our results suggest that outperforming funds use these updates to provide information about future performance, while updates of underperforming funds signal activism to investors.

Finally, in line with Beshears et al. (2010), our results imply that the additional provision of a summary prospectus does not (or only slightly) affect investors' behavior expressed by abnormal fund flows. To identify the causal effect of the SEC reform in 2009, we use an event study framework and exploit the fact that the release of a summary prospectus is not mandatory for a fund. By matching treated and non-treated funds along several fund characteristics, we can not reject the hypothesis of no effect on investors' behavior.

Overall, our findings suggest that funds inform, on average, reasonably accurate about their risks. However, there is considerable heterogeneity in the cross-section, which impedes a simple comparison of different funds for investors, and that is determined to a large extent at the fund group level. To shed more light on the underlying mechanisms and implications for investors, this paper offers a starting point for further research examining the informational value of soft information in prospectuses of financial products in general.

References

Abis, S., 2017. Man vs. machine: Quantitative and discretionary equity management. Unpublished working paper. Columbia Business School .

Baghai, R., Becker, B., Pitschner, S., 2019. The private use of credit ratings: Evidence from investment mandates. Swedish House of Finance Research Paper .

Bao, Y., Datta, A., 2014. Simultaneously discovering and quantifying risk types from textual risk disclosures. Management Science 60, 1371–1391.

Beshears, J., Choi, J., Laibson, D., Madrian, B. C., 2010. How does simplified disclosure affect individuals' mutual fund choices? University of Chicago Press pp. 75–96.

Brown, S., Goetzmann, W., Liang, B., Schwarz, C., 2008. Mandatory disclosure and operational risk: Evidence from hedge fund registration. The Journal of Finance 63, 2785–2815.

Brown, S. J., Goetzmann, W. N., 1997. Mutual fund styles. Journal of Financial Economics 43, 373–399.

Brown, S. V., Tucker, J. W., 2011. Large-sample evidence on firms' year-over-year md&a modifications. Journal of Accounting Research 49, 309–346.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H.-M., Steele, L. B., 2014. The information content of mandatory risk factor disclosures in corporate filings. Review of Accounting Studies 19, 396–455.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of Finance 52, 57–82.

Célérier, C., Vallée, B., 2017. Catering to investors through security design: Headline rate and complexity. The Quarterly Journal of Economics 132, 1469–1508.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. Journal of Political Economy 105, 1167–1200.

Chiu, T.-T., Guan, Y., Kim, J.-B., 2018. The effect of risk factor disclosures on the pricing of credit default swaps. Contemporary Accounting Research .

Choi, J. J., Laibson, D., Madrian, B. C., 2009. Why does the law of one price fail? an experiment on index mutual funds. The Review of Financial Studies 23, 1405–1432.

Cooper, M. J., Gulen, H., Rau, P. R., 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. The Journal of Finance 60, 2825–2858.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Gallaher, S. T., Kaniel, R., Starks, L. T., 2015. Advertising and mutual funds: From families to individual funds. CEPR Discussion Paper No. DP10329 .

Ghent, A. C., Torous, W. N., Valkanov, R., 2017. Complexity in structured finance. The Review of Economic Studies pp. 1–29.

Goetzmann, W. N., Peles, N., 1997. Cognitive dissonance and mutual fund investors. Journal of Financial Research 20, 145–158.

Gruber, M. J., 1996. Another puzzle: The growth in actively managed mutual funds. The Journal of Finance 51, 783–810.

Hartzmark, S. M., Sussman, A. B., 2019. Do investors value sustainability? a natural experiment examining ranking and fund flows. Journal of Finance .

Heinle, M. S., Smith, K. C., 2017. A theory of risk disclosure. Review of Accounting Studies 22, 1459–1491.

Hillert, A., Niessen-Ruenzi, A., Ruenzi, S., 2016. Mutual fund shareholder letters: flows, performance, and managerial behavior. Working paper .

Hope, O.-K., Hu, D., Lu, H., 2016. The benefits of specific risk-factor disclosures. Review of Accounting Studies 21, 1005–1045.

Huang, J., Sialm, C., Zhang, H., 2011. Risk shifting and mutual fund performance. The Review of Financial Studies 24, 2575–2616.

Investment Company Institute, 2006. Understanding investor preferences for mutual fund information. ICI Release, Available at https://ici.org/pdf/rpt_06_inv_prefs_summary.pdf .

Investment Company Institute, 2019. What us households consider when they select mutual funds, 2018. ICI Release, Available at https://ici.org/pdf/per25-03.pdf .

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.

Kozup, J., Howlett, E., Pagano, M., 2008. The effects of summary information on consumer perceptions of mutual fund characteristics. Journal of Consumer Affairs 42, 37–59.

Kravet, T., Muslu, V., 2013. Textual risk disclosures and investors' risk perceptions. Review of Accounting Studies 18, 1088–1122.

Lang, M., Stice-Lawrence, L., 2015. Textual analysis and international financial reporting: Large sample evidence. Journal of Accounting and Economics 60, 110–135.

Lynch, A. W., Musto, D. K., 2003. How investors interpret past fund returns. The Journal of Finance 58, 2033–2058.

SEC, 2009. Enhanced disclosure and new prospectus delivery option for registered open-end management investment companies. SEC Release, Available at https://www.sec.gov/rules/final/2009/33-8998.pdf .

Sekhon, J., 2011. Multivariate and propensity score matching software with automated balance optimization: The matching package for r. Journal of Statistical Software, Articles 42, 1–52.

Sirri, E. R., Tufano, P., 1998. Costly search and mutual fund flows. The Journal of Finance 53, 1589–1622.

Warther, V. A., 1995. Aggregate mutual fund flows and security returns. Journal of Financial Economics 39, 209–235.

Figures and Tables

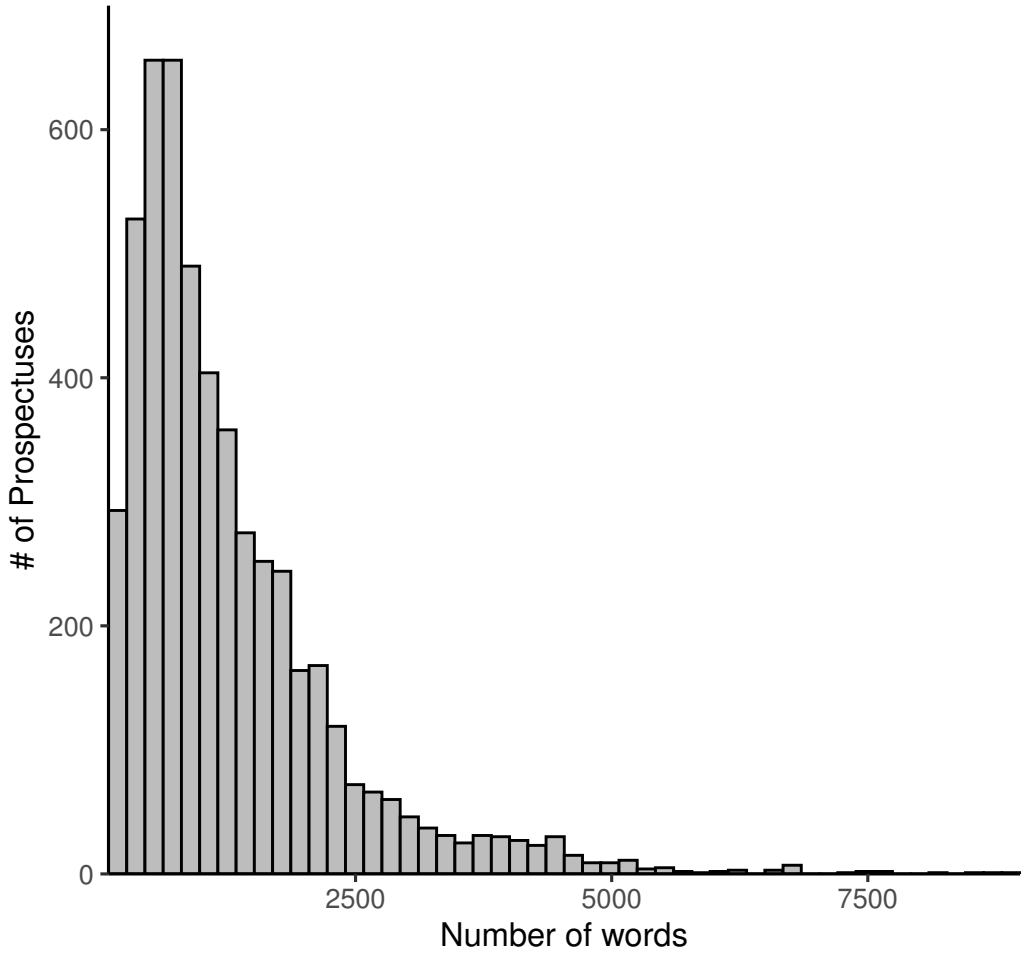

Figure 1. **Distribution of number of words in funds' risk statements in 2018**

Figure 1 displays the distribution of number of words in funds' risk statements. The number of words is calculated after removing the html code from the filings. In total the sample consists of 5,164 risk statements in 2018. See also Figure 5 in the Appendix for a comparison across funds in 2011.

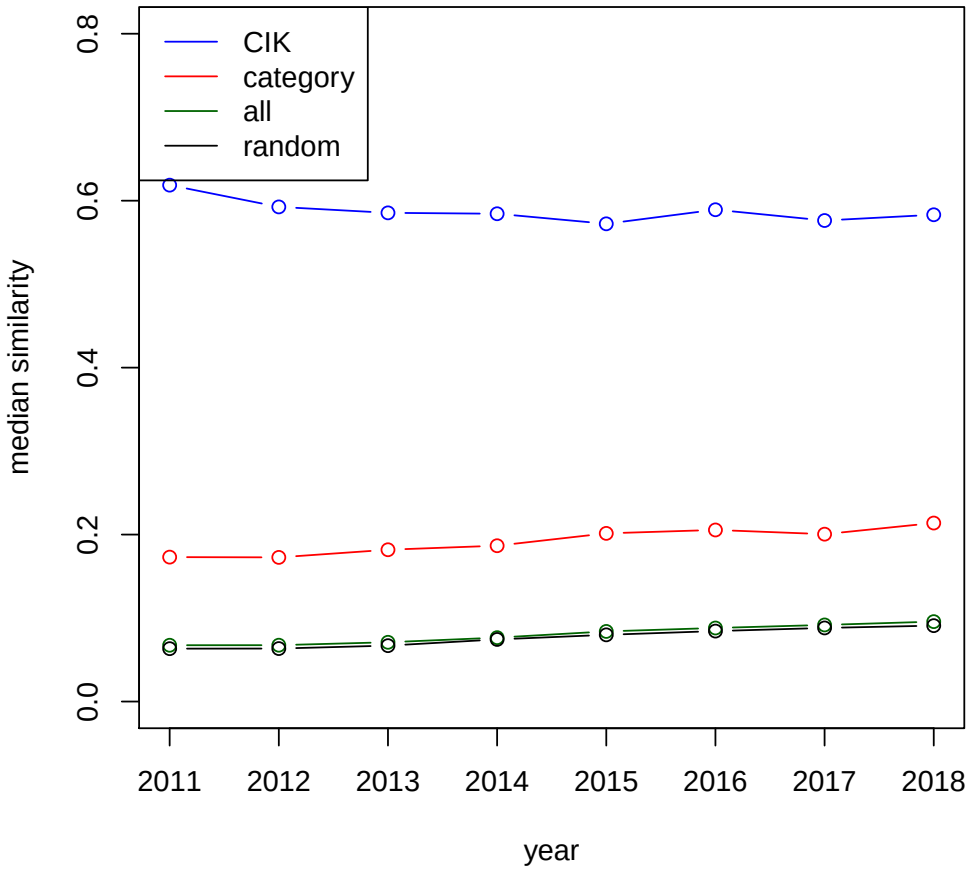

Figure 2. **Within group median similarity over time**

Figure 2 shows the yearly median similarity of a fund (matched with Morningstar and with information on the fund category) with any other fund in its managing company, i.e., CIK, (blue line), fund category (red line), and in the complete fund universe (green line). For the CIK and category effect, we exclude groups with one only fund as the similarity is 1 by definition and thus, would bias the results in Figure 2.

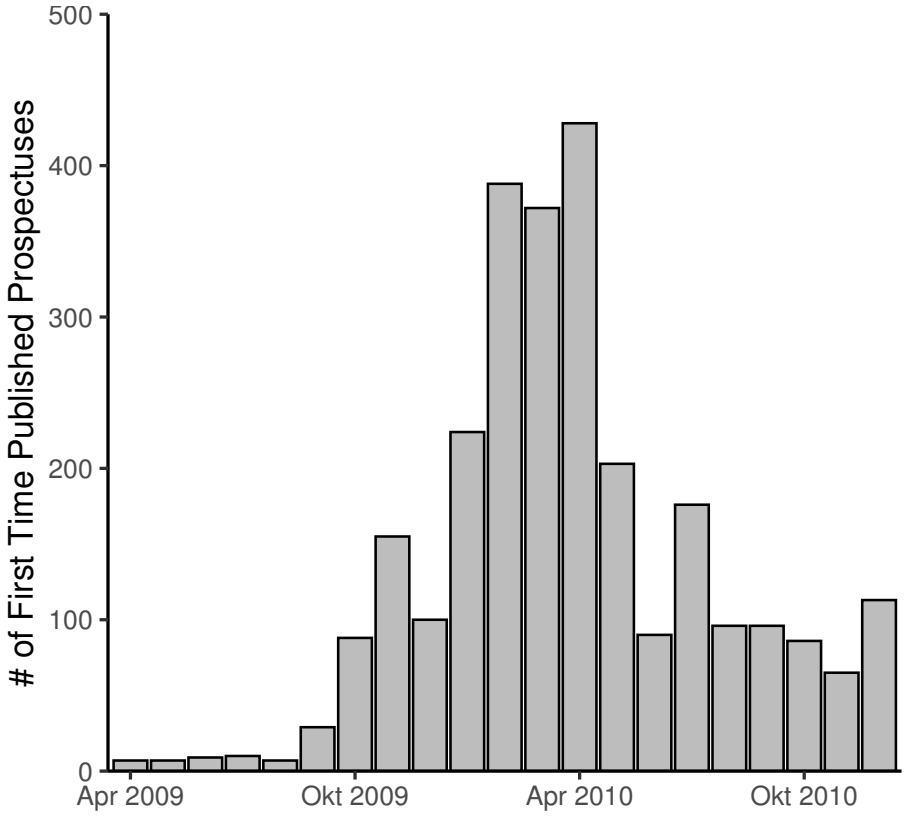

Figure 3. **Release of summary prospectuses over time**

Figure 3 shows the number of funds that decide in a month between April 2009 and December 2012 to publish for the first time a summary prospectus. In total 2,749 funds are in the treatment group.

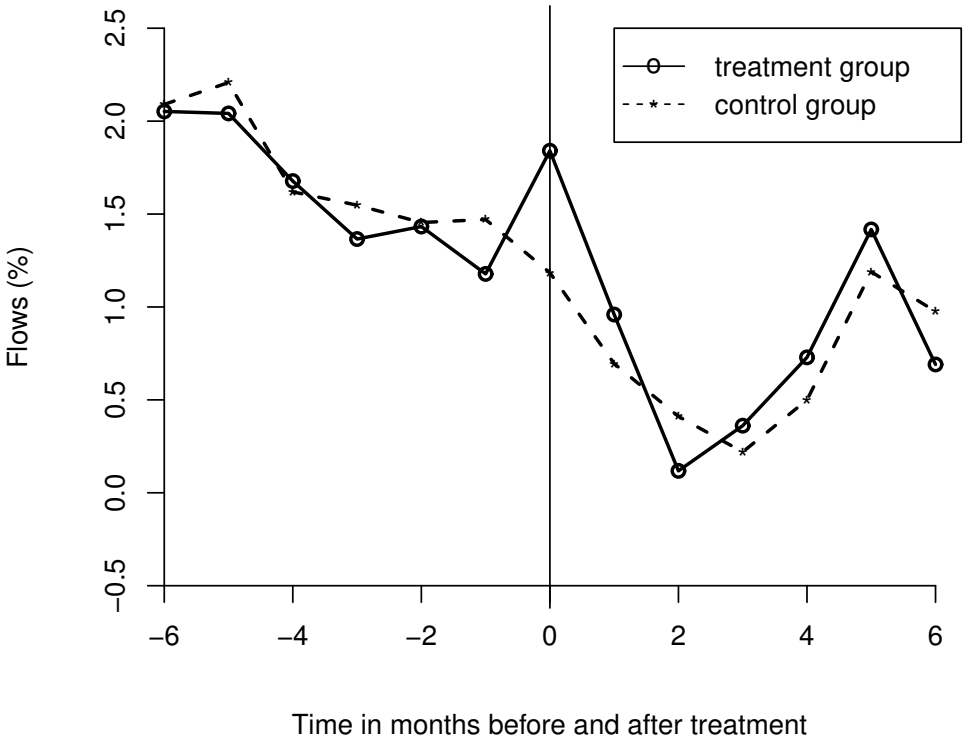

Figure 4. **Fund flows before and after treatment**

Figure 4 shows the monthly net flows in percent for both treated and non-treated funds 6 months before and 6 months after the introduction of a summary prospectus. The vertical line indicates the first time release of a summary prospectus.

Table 1: **Summary statistics**

Table 1 shows summary statistics (mean, standard deviation, median, first (p1) and last (p99) percentile, and number of observations of the fund characteristics. All fund characteristics are defined in detail in Table 15. The variable *Fund Flow* is winsorized at the bottom and the top percentile. The variable *Age* is calculated based on the IPO of the oldest share class. The sample includes all U.S. mutual funds in the SEC database that are linked to a fund in the Morningstar database. The sample consists of 8,194 funds between 2009 and 2018.

Statistic	Mean	Std	p1	Median	p99	Count
Fund Flow (in %)	1.20	7.82	-20.11	0.23	45.91	738224
Fund Size (in mn.)	1741.05	8508.68	2.04	264.68	24614.43	745038
log(Fund Size (in mn.))	19.30	2.07	14.53	19.39	23.93	745038
log(Company Size)	22.66	2.04	16.98	22.82	26.60	539581
Age (in Months)	166.15	142.80	3.00	141.00	723.00	745038
log(Age (in Months))	4.68	1.10	1.10	4.95	6.58	745038
Expense Ratio (in %)	0.69	4.99	0.00	0.38	4.28	721230
Turnover Ratio (in %)	92.68	433.63	0.44	46.00	779.00	724522
Return	0.36	2.42	-7.44	0.19	8.25	740950
Risk	1.90	1.93	0.07	1.19	8.46	622262
Sizescore	221.79	82.90	-3.56	251.32	342.91	361788
Cash Proportion (in %)	5.91	33.54	-13.39	2.55	85.82	479698
Bond Proportion (in %)	30.80	44.78	0.00	0.08	104.53	475741
Equity Proportion (in %)	61.51	49.08	0.00	89.48	100.04	480141
Holdings	287.62	786.99	4.00	96.00	2992.30	483371

Table 2: **Summary statistics length**

Table 2 shows summary statistics (mean, min, first quartile (q1), median, third quartile (q3), max, and number of observations) of the risk statements, i.e., the number of words from 485BPOS filings. The sample includes all U.S. open funds with SEC data. The sample period is from 2011 to 2018.

Year	Mean	Min	q1	Median	q3	Max	Count
2011	790.55	29.0	409.0	623.0	966.0	6,179.0	10,453
2012	886.6	86.0	440.0	696.0	1,094.0	6,732.0	11,573
2013	975.82	86.0	472.0	770.0	1,220.0	6,249.0	11,685
2014	1,103.37	86.0	521.0	877.0	1,392.0	6,994.0	12,054
2015	1,225.41	93.0	574.0	974.0	1,576.0	8,492.0	12,527
2016	1,371.3	91.0	610.0	1,081.0	1,802.0	8,489.0	12,689
2017	1,474.41	91.0	661.0	1,181.0	1,947.0	8,753.0	12,688
2018	1,550.66	91.0	694.0	1,238.0	2,046.0	9,209.0	12,576

Table 3: **Determinants of risk statements' length**

Table 3 shows the coefficients of OLS regressions with number of words as the dependent variable on various fund characteristics. In the first specification, we regress disclosed risk on fund risk only controlling for a year fixed effect. The control variables are defined in detail in Table 15. In column (2), we add category fixed effects (Category FE) to the regression. In column (3), we add investment company fixed effects (CIK FE) to the regression. Finally, in column (4), we control for category and investment company fixed effects simultaneously. Standard errors are reported in parentheses and are double-clustered at the category and fund group level. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent variable	Num. of Words (1)	Num. of Words (2)	Num. of Words (3)	Num. of Words (4)
Fund Risk	-56.201^{***}	-24.870^{***}	-41.817^{***}	28.689^{***}
	(2.958)	(3.057)	(2.690)	(2.661)
Flow Period	0.487^{***}	-0.063	-0.517^{***}	-0.261^{***}
	(0.131)	(0.116)	(0.092)	(0.078)
Return Period	-7.296^{***}	-2.160^{***}	-1.620^{***}	1.524^{***}
	(0.664)	(0.597)	(0.459)	(0.394)
log(Fund Size)	-37.156^{***}	-20.364^{***}	3.885^{*}	10.192^{***}
	(2.476)	(2.228)	(2.263)	(1.958)
Expense Ratio	-26.979^{***}	-28.899^{***}	-2.594	-9.153^{*}
	(7.706)	(6.940)	(6.109)	(5.203)
Turnover Ratio	0.210^{***}	0.026^{*}	0.021^{*}	-0.008
	(0.016)	(0.015)	(0.012)	(0.010)
log (Age)	-178.648^{***}	-123.123^{***}	-182.109^{***}	-105.377^{***}
	(6.396)	(6.130)	(5.719)	(5.031)
Summary Prospectus	57.503^{***}	45.524^{***}	42.653^{***}	59.527^{***}
	(11.696)	(10.354)	(14.052)	(11.855)
Year FE	Y	Y	Y	Y
Category FE	N	Y	N	Y
CIK FE	N	N	Y	Y
Observations	30,964	30,964	30,964	30,964
R^2	0.140	0.348	0.637	0.745

Table 4: **Determinants of length of risk statements - idiosyncratic and systematic risk**

Table 4 shows the coefficients of OLS regressions with number of words as the dependent variable. In the specifications (1)-(3), we regress disclosed risk on the idiosyncratic risk. The control variables are defined in detail in Table 15. In the specifications (4)-(6), we regress disclosed risk on the systematic risk. Standard errors are reported in parentheses and are double-clustered at the category and issuer level. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent variable	Number of Words			Number of Words		
	(1)	(2)	(3)	(4)	(5)	(6)
Idiosyncratic Risk	0.593*	3.892***	0.695***			
	(0.327)	(0.155)	(0.237)			
Systematic Risk				−205.843***	−87.695***	−5.938
				(12.325)	(13.223)	(13.173)
Flow Period (%)	−0.051	−0.372***	−0.332***	−0.231**	−0.394***	−0.334***
	(0.110)	(0.078)	(0.075)	(0.109)	(0.079)	(0.075)
Return Period	−2.691***	0.041	1.382***	0.238	−0.687*	1.378***
	(0.521)	(0.367)	(0.356)	(0.546)	(0.375)	(0.360)
log(Fund Size)	−15.963***	4.778**	3.669*	−13.774***	2.577	3.677*
	(2.232)	(2.017)	(1.963)	(2.217)	(2.063)	(1.972)
Expense Ratio	−17.393***	4.398	2.134	20.278***	16.362***	2.938
	(5.436)	(4.268)	(4.088)	(5.809)	(4.449)	(4.206)
Turnover Ratio	0.278***	−0.004	−0.011	0.263***	0.018	−0.007
	(0.024)	(0.019)	(0.018)	(0.024)	(0.019)	(0.018)
log(Age)	−86.770***	−102.304***	−81.276***	−93.859***	−119.328***	−81.857***
	(6.145)	(5.136)	(5.054)	(6.106)	(5.186)	(5.055)
Summary Prospectus	−1.283	62.159***	61.542***	−29.753***	60.813***	61.371***
	(10.150)	(12.188)	(11.649)	(10.204)	(12.501)	(11.751)
Year FE	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Category FE	*Y*	*N*	*Y*	*Y*	*N*	*Y*
CIK FE	*N*	*Y*	*Y*	*N*	*Y*	*Y*
Observations	16,412	16,412	16,412	16,412	16,412	16,412
R^2	0.197	0.663	0.693	0.210	0.650	0.693

Table 5: **Variance decomposition using fixed effects**

Table 5 shows the results from regressing a fund's median similarity with all other funds in a given year on year, fund category, and fund group fixed effects. Our set of funds comprises all funds matched with Morningstar and with information on the fund category.

	risk statement	strategy statement	primary objective
	Fraction accounted by (%)		
Year FE	9.8	1.3	0.3
Category FE	11.9	13.0	5.6
CIK FE	46.4	43.6	51.5
Residual	31.9	42.2	42.6
Total	100.0	100.0	100.0
Observations	35,683	35,683	35,683

Table 6: **Changes in risk statements and fund performance**

Table 6 shows the coefficients of OLS regressions with fund performance on differences in the disclosed risk information and various fund characteristics. The dependent variable is the CAPM alpha, the Fama and French (1993) 3-factor alpha, the Carhart (1997) 4-factor alpha, and the fund's return. Performance is measured on a yearly basis in percent. Alphas are computed using beta coefficients obtained from a regression using monthly fund returns over 12 months. The difference score variable is defined between 0 and 1. All control variables are defined in detail in Table 15 in the Appendix. Period of interest is 2012–2018. Standard errors are reported in parentheses and are double-clustered at the fund and year level. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

	Dependent variable:			
	$\hat{\alpha}$ (CAPM)	$\hat{\alpha}$ (3FF)	$\hat{\alpha}$ (4FF)	yearly returns
	(1)	(2)	(3)	(4)
Diff Score	0.071	−0.026	0.016	0.434
	(0.122)	(0.127)	(0.126)	(0.999)
log(Age)	−0.321***	−0.385***	−0.344***	−3.338***
	(0.089)	(0.092)	(0.091)	(0.726)
log(Fund Size)	0.090***	0.084***	0.083***	−0.302
	(0.026)	(0.027)	(0.027)	(0.211)
log(Company Size)	0.146***	0.143***	0.136***	−0.673**
	(0.037)	(0.038)	(0.038)	(0.300)
Expense Ratio	−0.022	−0.041*	−0.039	0.411**
	(0.024)	(0.024)	(0.024)	(0.193)
Fund Flow Period	0.103***	0.107***	0.111***	−0.490
	(0.037)	(0.039)	(0.039)	(0.306)
Year FE	*Y*	*Y*	*Y*	*Y*
Fund FE	*Y*	*Y*	*Y*	*Y*
Observations	7,318	7,318	7,318	7,318
R^2	0.444	0.442	0.417	0.661

Table 7: **Changes in risk statements and fund performance**

Table 7 shows the coefficients of OLS regressions with fund performance on differences in the disclosed risk information and various fund characteristics. The dependent variable is the CAPM alpha, the Fama and French (1993) 3-factor alpha, the Carhart (1997) 4-factor alpha, and the fund's return. Performance is measured on a yearly basis in percent. Alphas are computed using beta coefficients obtained from a regression using monthly fund returns over 12 months. The difference score variable is defined between 0 and 1. All control variables are defined in detail in Table 15 in the Appendix. Period of interest is 2012–2018. Standard errors are reported in parentheses and are double-clustered at the fund and year level. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

	Dependent variable:		
	$\hat{\alpha}$ (CAPM)	$\hat{\alpha}$ (3FF)	$\hat{\alpha}$ (4FF)
	(1)	(2)	(3)
Diff Score	-0.226^{*}	-0.346^{***}	-0.284^{**}
	(0.127)	(0.130)	(0.127)
log(Age)	-0.350^{***}	-0.398^{***}	-0.376^{***}
	(0.088)	(0.091)	(0.090)
log(Fund Size)	0.094^{***}	0.089^{***}	0.091^{***}
	(0.026)	(0.027)	(0.026)
log(Company Size)	0.147^{***}	0.142^{***}	0.133^{***}
	(0.037)	(0.038)	(0.037)
Expense Ratio	-0.021	-0.039	-0.038
	(0.023)	(0.024)	(0.024)
Fund Flow Period	0.068^{*}	0.072^{*}	0.085^{**}
	(0.037)	(0.039)	(0.038)
Diff Score$_{f,t} \times \mathbb{1}_{\hat{\alpha} \geq 0}$	2.752^{***}		
	(0.333)		
Diff Score$_{f,t} \times \mathbb{1}_{\hat{\alpha} \geq 0}$		4.131^{***}	
		(0.399)	
Diff Score$_{f,t} \times \mathbb{1}_{\hat{\alpha} \geq 0}$			5.463^{***}
			(0.440)
Year FE	Y	Y	Y
Fund FE	Y	Y	Y
Observations	7,318	7,318	7,318
R^2	0.451	0.452	0.433

Table 8: **Pretreatment characteristics with unmatched and matched samples**
Table 8 shows the mean for each of the five covariates for the treatment and control group respectively and the standardized differences between the groups. Panel A presents the results for the unmatched sample, while Panel B shows the results for the matched sample using propensity score matching. We use the Mahalanobis distance to compute the weighting of each covariate and allow ties, i.e., a treated fund matches more than one fund from the control group. Finally, matched funds are weighted such that the sum of weighted funds is equal to the original number of observations (Sekhon, 2011).

	N	Flows Period	log(Size)	Return Period	log(Age)	Std(Returns)
Panel A: **Unmatched sample**						
mean Treatment	2749	9.79	19.43	3.99	4.88	1.76
mean Control	90751	12.24	18.92	6.40	4.66	2.57
std. Difference		-10.67	26.74	-42.21	26.49	-40.66
KS Bootstrap p-value		0.00	0.00	0.00	0.00	0.00
Panel B: **Propensity score matching**						
mean Treatment	2749	9.79	19.43	3.99	4.88	1.76
mean Control	2751	9.70	19.44	3.95	4.88	1.71
std. Difference		0.40	-0.51	0.73	0.15	2.41
KS Bootstrap p-value		0.42	0.76	0.83	0.39	0.40

Table 9: **Cumulative abnormal returns around introduction**

Table 9 shows average cumulative abnormal flows earned by funds around the first publishing of a summary prospectus in 2009 and 2010. Fund flows are defined as $[TNA_{f,t} - TNA_{f,t-1}(1 + RF_t)]/TNA_{f,t-1}(1 + RF_t)$. Abnormal fund flows are computed for each fund that introduces a summary prospectus in our sample with respect to the closest mutual fund that does not introduce such a summary prospectus on the basis of its propensity score (see Table 8). For the propensity scores we estimate for each event date a logistic regression by assigning the dependent variable a dummy of 1 if the fund publishes a summary prospectus, and all other funds 0. From the date that a fund publishes the first time a summary prospectus, it is removed from the control group as a potential match. For the propensity scores we use the following independent variables: the 1-month lagged log of total net assets, the accumulative past 6 months return as well as 6-months flows to the fund before the introduction of a summary prospectus, the log of fund age in months, and the standard deviation of the returns over the 6 months before the summary prospectus. Panel A reports cumulative abnormal flows. Panel B and C report normalized abnormal cumulative fund flows. In Panel B fund flows are normalized by diving flows by total net assets (see e.g., Cooper et al. (2005)). In Panel C fund flows are normalized according to a method from Hartzmark and Sussman (2019), who propose to calculate the rank of fund flows. We report t-statistics in parenthesis for the null hypothesis that the event window abnormal flows are 0. The sample includes all U.S. mutual funds in the SEC database that publish the first time a summary prospectus in 2009 or 2010.

	N	Months						
		-6 to 0	-3 to 0	-1	1	0 to 3	0 to 6	0 to 12
Panel A: Cumulative abnormal fund flows								
CAF (%)	2,749	0.05%	0.34%	-0.29%	0.26%	0.64%	1.20%	1.76%
		(0.08)	(0.98)	(-1.67)	(1.38)	(1.78)	(2.01)	(1.76)
Panel B: Cumulative abnormal normalized fund flows (flows/tna$_{-t}$)								
CAF Normal. (%)	2,749	0.17	0.07	0.01	0.04	0.12	0.23	0.38
		(0.66)	(1.20)	(0.25)	(1.42)	(2.16)	(2.31)	(2.20)
Panel C: Cumulative abnormal normalized fund flows (rank based)								
CAF Rank Normal. (%)	2,749	0.14	0.06	-0.00	-0.00	-0.03	-0.22	-0.80
		(0.35)	(0.28)	(-0.05)	(-0.13)	(-0.14)	(-0.57)	(-1.10)

Appendix A. Example risk statements

A.1. Hancock Horizon Burkenroad Small Cap Fund[30]

Principal Risks

As with all mutual funds, there is no guarantee that the Fund will achieve its investment objective. You could lose money by investing in the Fund. **A Fund share is not a bank deposit and is not insured or guaranteed by the FDIC, or any government agency.** The principal risks affecting shareholders' investments in the Fund are set forth below.

Equity Risk – Since it purchases equity securities, the Fund is subject to the risk that stock prices will fall over short or extended periods of time. Historically, the equity market has moved in cycles, and the value of the Fund's securities may fluctuate from day to day. 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. These factors contribute to price volatility, which is the principal risk of investing in the Fund.

Small-Capitalization Company Risk – The smaller-capitalization companies that the Fund invests in may be more vulnerable to adverse business or economic events than larger, more established companies. In particular, investments in these small companies may pose additional risks, including liquidity risk, because these companies tend to have limited product lines, markets and financial resources, and may depend upon a relatively small management group. Therefore, small-capitalization stocks may be more volatile than those of larger companies. These securities may be traded over-the-counter or listed on an exchange. The Fund is also subject to the risk that the Adviser's particular investment style, which focuses on small-capitalization stocks, may underperform other segments of the equity market or the equity market as a whole.

REITs Risk – REITs are pooled investment vehicles that own, and usually operate, income-producing real estate. REITs are susceptible to the risks associated with direct ownership of real estate, such as the following: (i) declines in property values; (ii) increases in property taxes, operating expenses, interest rates or competition; (iii) overbuilding; (iv) zoning changes; and (v) losses from casualty or condemnation. REITs typically incur fees that are separate from those of the Fund. Accordingly, the Fund's investments in REITs will result in the layering of expenses such that shareholders will indirectly bear a proportionate share of the REITs' operating expenses, in addition to paying Fund expenses."

[30] This risk statement from the fund's summary prospectus is available at https://www.sec.gov/Archives/edgar/data/890540/000139834419010287/fp0042397_497k.htm

A.2.Franklin Mutual Shares Fund[31]

Principal Risks

You could lose money by investing in the Fund. Mutual fund shares are not deposits or obligations of, or guaranteed or endorsed by, any bank, and are not insured by the Federal Deposit Insurance Corporation, the Federal Reserve Board, or any other agency of the U.S. government.

Market The market values of securities or other investments owned by the Fund will go up or down, sometimes rapidly or unpredictably. The market value of a security or other investment may be reduced by market activity or other results of supply and demand unrelated to the issuer. This is a basic risk associated with all investments. When there are more sellers than buyers, prices tend to fall. Likewise, when there are more buyers than sellers, prices tend to rise.

Stock prices tend to go up and down more dramatically than those of debt securities. A slower-growth or recessionary economic environment could have an adverse effect on the prices of the various stocks held by the Fund.

Smaller and Midsize Companies Securities issued by smaller and midsize companies may be more volatile in price than those of larger companies, involve substantial risks and should be considered speculative. Such risks may include greater sensitivity to economic conditions, less certain growth prospects, lack of depth of management and funds for growth and development, and limited or less developed product lines and markets. In addition, smaller and midsize companies may be particularly affected by interest rate increases, as they may find it more difficult to borrow money to continue or expand operations, or may have difficulty in repaying any loans.

Value Style Investing A value stock may not increase in price as anticipated by the investment manager if other investors fail to recognize the company's value and bid up the price, the markets favor faster-growing companies, or the factors that the investment manager believes will increase the price of the security do not occur or do not have the anticipated effect.

Foreign Securities (non-U.S.) Investing in foreign securities typically involves more risks than investing in U.S. securities, including risks related to currency exchange rates and policies, country or government specific issues, less favorable trading practices or regulation and greater price volatility. Certain of these risks also may apply to securities of U.S. companies with significant foreign operations.

Regional Focus Because the Fund may invest at least a significant portion of its assets in companies in a specific region, including Europe, the Fund is subject to greater risks of adverse developments in that region and/or the surrounding regions than a fund that is more broadly diversified geographically. Political, social or economic disruptions in the region, even in countries in which the Fund is not invested, may adversely affect the value of investments held by the Fund. Current political uncertainty surrounding the European Union (EU) and its membership, including the 2016 referendum in which the United Kingdom voted to exit the EU, may increase market volatility. The financial instability of some countries in the EU together with the risk of such instability impacting other more stable countries may increase the economic risk of investing in companies in Europe.

Derivative Instruments The performance of derivative instruments depends largely on the performance of an underlying instrument, such as a currency, security, interest rate or index, and such instruments

[31]This risk statement from the fund's summary prospectus is available at https://www.sec.gov/Archives/edgar/data/825063/000137949119002311/filing186636985.htm.

often have risks similar to the underlying instrument, in addition to other risks. Derivatives involve costs and can create economic leverage in the Fund's portfolio, which may result in significant volatility and cause the Fund to participate in losses (as well as gains) in an amount that significantly exceeds the Fund's initial investment. Certain derivatives have the potential for unlimited loss, regardless of the size of the initial investment. Other risks include illiquidity, mispricing or improper valuation of the derivative instrument, and imperfect correlation between the value of the derivative and the underlying instrument so that the Fund may not realize the intended benefits. The successful use of derivatives will usually depend on the investment manager's ability to accurately forecast movements in the market relating to the underlying instrument. Should a market or markets, or prices of particular classes of investments move in an unexpected manner, especially in unusual or extreme market conditions, the Fund may not achieve the anticipated benefits of the transaction, and it may realize losses, which could be significant. If the investment manager is not successful in using such derivative instruments, the Fund's performance may be worse than if the investment manager did not use such derivative instruments at all. When a derivative is used for hedging, the change in value of the derivative may also not correlate specifically with the currency, security, interest rate, index or other risk being hedged. Derivatives also may present the risk that the other party to the transaction will fail to perform. There is also the risk, especially under extreme market conditions, that an instrument, which usually would operate as a hedge, provides no hedging benefits at all.

Merger Arbitrage Securities and Distressed Companies A merger or other restructuring, or a tender or exchange offer, proposed or pending at the time the Fund invests in merger arbitrage securities may not be completed on the terms or within the time frame contemplated, which may result in losses to the Fund. Debt obligations of distressed companies typically are unrated, lower-rated, in default or close to default and are generally more likely to become worthless than the securities of more financially stable companies.

Liquidity From time to time, the trading market for a particular security or type of security or other investments in which the Fund invests may become less liquid or even illiquid. Reduced liquidity will have an adverse impact on the Fund's ability to sell such securities or other investments when necessary to meet the Fund's liquidity needs, which may arise or increase in response to a specific economic event or because the investment manager wishes to purchase particular investments or believes that a higher level of liquidity would be advantageous. Reduced liquidity will also generally lower the value of such securities or other investments. Market prices for such securities or other investments may be relatively volatile.

Management The Fund is subject to management risk because it is an actively managed investment portfolio. The Fund's investment manager applies investment techniques and risk analyses in making investment decisions for the Fund, but there can be no guarantee that these decisions will produce the desired results."

Appendix B. Additional Figures

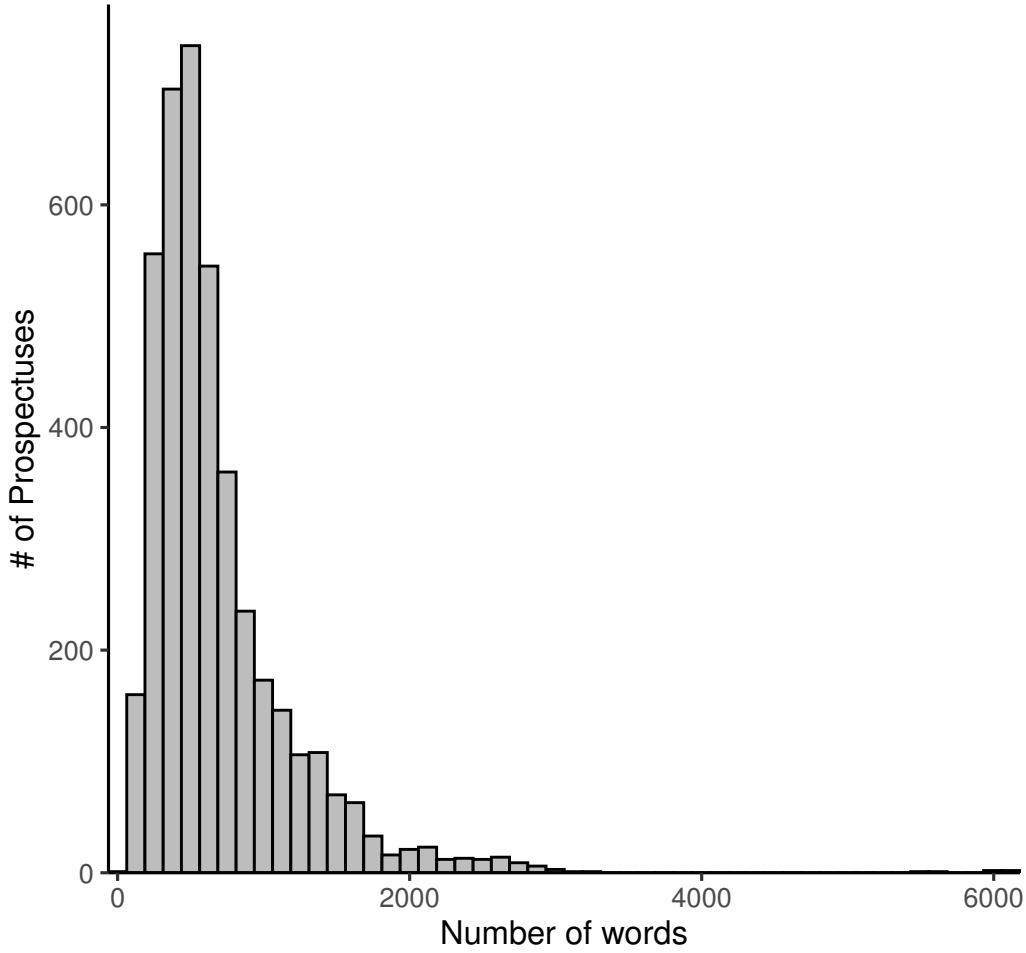

Figure 5. **Distribution of number of words in funds' risk statements in 2018**

Figure 1 displays the distribution of number of words in funds' risk statements. The number of words is calculated after removing the html code from the filings. In total the sample consists of 4,136 risk statements in 2011.

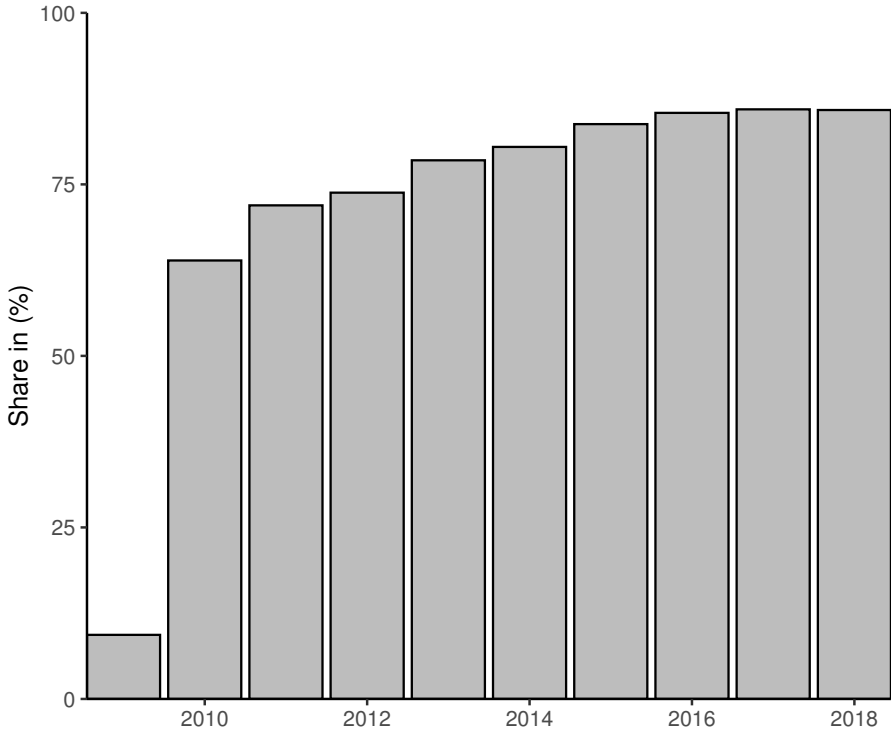

Figure 6. **Share of funds that publish a summary prospectus**

Figure 6 illustrates for each year between 2009 and 2018 the share of funds that publish a summary prospectus.

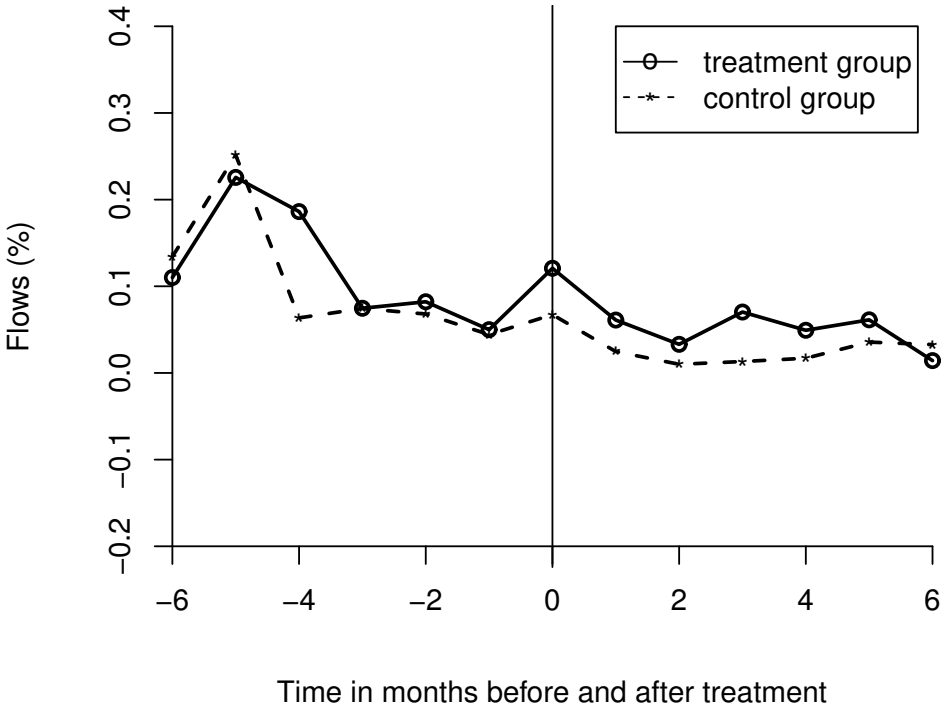

Figure 7. **Normalized fund flows before and after treatment**

Figure 7 shows the monthly normalized net flows in percent for both treated and non-treated funds 6 months before and 6 months after the introduction of a summary prospectus. The vertical line indicates the first time release of a summary prospectus. Fund flows are normalized by dividing flows by lagged total net assets.

Appendix C. Additional Tables

Table 10: **Data Merge Process**

Table 10 describes the process of merging and shows the final number of distinct funds and share classes in our sample. 497(K) and 485BPOS filings are available in the SEC database. Starting point is the investment companies file of the SEC with 45,349 share classes which correspond to 13,165 prospectuses. 31,796 of those share classes have a ticker symbol and 26,591 of the ticker symbol can we match with ticker symbols in the Morningstar database, which are 9,150 prospectuses, respectively. Merging this data with fund (not share class) specific variables, we have data for 26,540 share classes, respectively 9,133 prospectuses. Finally, we merge the data with the Series ID data from the SEC. Series ID is used as an identifier of fund level at the SEC. Our final sample has observations for 26,138 share classes and 8,194 prospectuses.

Level	SEC	MS ticker data	Merge FundID	Merge Series_ID
Share class	45,349	26,591	26,540	26,138
	(31,796 with ticker)			
Fund level	15,118	9,139	9,136	8,952
Prospectus	13,165	9,150	9,133	8,938

Table 11: **Summary statistics equity funds**

Table 11 shows summary statistics (mean, standard deviation, median, first (p1) and last (p99) percentile, and number of observations of all equity fund characteristics in our sample. All fund characteristics are defined in detail in Table 15. The variable *Fund Flow* is winsorized at the bottom and the top percentile. The variable *Age* is calculated based on the IPO of the oldest share class. The sample includes all U.S. equity mutual funds in the SEC database that are linked to a fund in the Morningstar database. The sample consists of 4,385 funds between 2009 and 2018.

Statistic	Mean	Std	p1	Median	p99	Count
Fund Flow (in %)	0.96	7.51	-18.61	0.14	39.09	390122
Fund Size (in mn.)	1835.05	9829.13	2.05	275.13	25123.92	393806
log(Fund Size (in mn.))	19.32	2.09	14.53	19.43	23.95	393806
log(Company Size)	22.59	2.03	16.58	22.83	26.33	286047
Age (in Months)	169.96	145.05	3.00	148.00	774.00	393806
log(Age (in Months))	4.73	1.07	1.10	5.00	6.65	393806
Expense Ratio (in %)	0.80	3.57	0.04	0.48	5.07	382021
Turnover Ratio (in %)	77.69	183.30	2.00	48.00	581.00	385616
Return	0.49	3.01	-8.82	0.37	9.87	391458
Risk	2.64	2.09	0.36	1.91	9.22	334525
Sizescore	214.82	88.81	-0.13	246.96	344.76	252227
Cash Proportion (in %)	3.41	11.09	-0.74	1.81	43.09	251998
Bond Proportion (in %)	0.64	6.52	0.00	0.00	15.63	248804
Equity Proportion (in %)	95.08	12.05	41.53	97.53	100.40	252498
Holdings	232.25	539.72	7.00	88.00	2803.00	253611

Table 12: **Determinants of risk statements' length (logwords)**

Table 12 shows the coefficients of OLS regressions with number of words as the dependent variable on various fund characteristics. In the first specification, we regress disclosed risk on fund risk only controlling for a year fixed effect. The control variables are defined in detail in Table 15. In column (2), we add category fixed effects (Category FE) to the regression. In column (3), we add investment company fixed effects (CIK FE) to the regression. Finally, in column (4), we control for category and investment company fixed effects simultaneously. Standard errors are reported in parentheses and are double-clustered at the category and fund group level. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent variable	log(Words) (1)	log(Words) (2)	log(Words) (3)	log(Words) (4)
Fund Risk	-0.079^{***}	-0.039^{***}	-0.054^{***}	0.013^{***}
	(0.003)	(0.003)	(0.002)	(0.002)
Flow Period	0.0004^{***}	0.00001	-0.0002^{***}	-0.0001^{*}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Return Period	-0.009^{***}	-0.003^{***}	-0.003^{***}	0.0004
	(0.001)	(0.001)	(0.0004)	(0.0003)
log(Fund Size)	-0.040^{***}	-0.025^{***}	-0.003	0.009^{***}
	(0.002)	(0.002)	(0.002)	(0.001)
Expense Ratio	-0.013^{*}	-0.014^{**}	-0.002	-0.003
	(0.007)	(0.006)	(0.005)	(0.004)
Turnover Ratio	0.0002^{***}	0.0001^{***}	0.00003^{***}	0.00001
	(0.00001)	(0.00001)	(0.00001)	(0.00001)
log(Age)	-0.179^{***}	-0.126^{***}	-0.168^{***}	-0.100^{***}
	(0.006)	(0.006)	(0.005)	(0.004)
Summary Prospectus	-0.013	-0.012	-0.007	0.023^{**}
	(0.011)	(0.010)	(0.011)	(0.009)
Year FE	Y	Y	Y	Y
Category FE	N	Y	N	Y
CIK FE	N	N	Y	Y
Observations	30,964	30,964	30,964	30,964
R^2	0.172	0.359	0.732	0.829

Table 13: **Determinants of risk statements' length - equity funds**

Table 13 shows the coefficients of OLS regressions with number of words as the dependent variable on various fund characteristics. In the first specification, we regress disclosed risk on fund risk only controlling for a year fixed effect. The control variables are defined in detail in Table 15. In column (2), we add category fixed effects (Category FE) to the regression. In column (3), we add investment company fixed effects (CIK FE) to the regression. Finally, in column (4), we control for category and investment company fixed effects simultaneously. Standard errors are reported in parentheses and are double-clustered at the category and fund group level. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent variable	Num. of Words (1)	Num. of Words (2)	Num. of Words (3)	Num. of Words (4)
Fund Risk	-22.342^{***}	-34.242^{***}	13.729^{***}	6.269^{**}
	(2.539)	(2.565)	(2.571)	(2.558)
Flow Period	-0.160	-0.154	-0.330^{***}	-0.321^{***}
	(0.113)	(0.109)	(0.079)	(0.074)
Return Period	-5.010^{***}	-0.733	-1.199^{***}	1.259^{***}
	(0.538)	(0.537)	(0.371)	(0.357)
log (Fund Size)	-11.276^{***}	-13.767^{***}	5.366^{***}	4.593^{**}
	(2.258)	(2.206)	(2.053)	(1.958)
Expense Ratio	14.621^{**}	13.508^{**}	5.227	0.630
	(5.835)	(5.648)	(4.428)	(4.174)
Turnover Ratio	0.327^{***}	0.274^{***}	0.027	-0.007
	(0.024)	(0.024)	(0.019)	(0.018)
log (Age)	-129.081^{***}	-91.173^{***}	-118.484^{***}	-81.474^{***}
	(6.005)	(6.095)	(5.185)	(5.044)
Summary Prospectus	-2.299	-24.903^{**}	78.966^{***}	66.621^{***}
	(10.500)	(10.167)	(12.417)	(11.665)
Year FE	Y	Y	Y	Y
Category FE	N	Y	N	Y
CIK FE	N	N	Y	Y
Observations	16,476	16,476	16,476	16,476
R^2	0.132	0.206	0.649	0.693

Table 14: **Determinants of risk statements' length - equity funds**

Table 14 presents summary statistics (mean, min, first quartile (q1), median, third quartile (q3), max, and number of observations) on text-based variable on fund-year level, i.e., the absolute change in number of words (changes) and the cosine similarity (similarities) from year $t-1$ to t.

				changes			
	mean	min	q25	median	q75	max	count
year							
2012	72.79	−3407.0	0.0	15.0	108.0	2166.0	9785.0
2013	71.08	−2425.0	0.0	19.0	107.0	3904.0	10785.0
2014	94.58	−4276.0	0.0	31.0	145.0	2501.0	11044.0
2015	93.35	−2031.0	0.0	30.0	138.0	4964.0	11453.0
2016	133.35	−4002.0	0.0	56.0	215.0	6053.0	11798.0
2017	96.28	−2644.0	0.0	33.0	143.0	3240.0	11807.0
2018	67.28	−2829.0	0.0	13.0	100.0	2981.0	11787.0
				similarities			
	mean	min	q25	median	q75	max	count
year							
2012	0.94	0.11	0.93	0.97	1.0	1.0	9785.0
2013	0.95	0.16	0.94	0.98	1.0	1.0	10785.0
2014	0.95	0.21	0.94	0.98	0.99	1.0	11044.0
2015	0.96	0.26	0.94	0.98	1.0	1.0	11453.0
2016	0.95	0.23	0.93	0.97	0.99	1.0	11798.0
2017	0.96	0.21	0.95	0.98	1.0	1.0	11807.0
2018	0.97	0.03	0.97	0.99	1.0	1.0	11787.0

Table 15: **Variable definitions**

Table 15 provides definitions of all variables used in this paper. MS indicates Morningstar, C refers to own calculation and SEC indicates data from the Securities and Exchange Comission.

Variable name	Description	Source
Fund Flow (%)	Computed as $(TNA_{f,t} - TNA_{f,t-1}(1 + RF_{f,t}))/(TNA_{f,t-1}(1 + RF_{f,t}))$, where $TNA_{f,t}$ corresponds to fund f's total net assets (TNA) in month t and $RT_{f,t}$ denotes fund f's return in month t. The variable is winsorized at the 1st and 99th percentile.	MS, C
Investor Return (%)	Percentage return calculated as the change in monthly net asset value minus management fees and other regular costs.	MS
$\hat{\alpha}$ (CAPM) (%)	Performance alpha from the market model. Market returns are from the Kenneth French library.	MS, C
$\hat{\alpha}$ (3FF) (%)	Performance alpha from a model including the Fama and French factor returns for the market. Market returns are from the Kenneth French library.	MS, C
$\hat{\alpha}$ (4FF) (%)	Performance alpha from a model including the Fama and French factor returns for the market as well as the Carhart momentum factor. Market returns are from the Kenneth French library.	MS, C
Fund Size (in mn.)	Fund's total net assets in million USD, aggregated at the share class level.	MS
log (Fund Size)	Logarithm of fund's total net assets in million USD, aggregated at the share class level.	MS, C
log (Company Size)	Logarithm of company's total net assets in million USD, aggregated at the fund level.	MS, C
Expense Ratio (%)	A fund's annual expense ratio expressed in percent.	MS
Age (in months)	Fund's age computed as the difference from end 2018 to the inception date of the oldest share class.	MS, C
log(Age)	Logarithm of fund's age in months computed as the difference from end 2018 to the inception date of the oldest share class.	MS, C
Turnover Ratio (%)	Lesser of purchases or sales (excluding all securities with maturities of less than one year) and dividing by average monthly net assets.	MS
Fund Risk	Standard deviation of monthly returns. Calculated for the past 36 months.	MS, C

Variable definitions, Table 15 continued

Variable name	Description	Source
Systematic Risk	A fund's beta measured as the slope of the regression of the excess return on the fund as the dependent variable and the excess return on the benchmark as the independent variable.	MS, C
Idiosyncratic Risk	A fund's risk measured as residual part from the slope of the regression of the excess return on the fund as the dependent variable and the excess return on the benchmark as the independent variable.	MS, C
Size Score	Companies are scored according to their relative size. Each stock is given a Size Score that ranges from -100 (very micro) to 400 (very giant).	MS
Summary Prospectus Dummy	Dummy if the fund issues a summary prospectus.	SEC, C
Category	Morningstar assigns each fund to one of 70 categories.	MS
Fund group	Investment company that issues the fund. An investment company can issue several funds.	SEC
Cash	The percentage of the fund's net assets in cash.	MS
Fixed Income	The percentage of the fund's net assets in fixed income.	MS
Equity	The percentage of the fund's net assets in equity.	MS
Foreign Investment	The percentage of the fund's net assets in foreign assets, calculated as the sum of investments in equity and in fixed income.	MS, C
Passive	Dummy that indicates if a fund is an index fund or not.	MS
Number of Holdings	Number of holdings that are in the portfolio of a fund.	MS
Number of Words	Number of words in the risk statements.	SEC,C
log(Number of Words)	Logarithm of number of words in the risk statements.	SEC,C
Sim	Sim is the similarity score between a risk statement compared to other risk statements.	C
Diff Score	Diff Score is calculated as the difference of similarity between risk statements of a fund in two consecutive years.	C
Return Period	Return over the 12 months before the calendar year end.	MS,C
Fund Flow Period	Flow over the 12 months before the calendar year end.	MS,C