Understanding Risk: Lessons from Mutual Fund Disclosures

Nikola Kiprijanovski*

September 25, 2026

Abstract

Do mutual fund disclosures reveal risks that matter for asset prices? I use risk disclosures to show how the corporate finance of asset management shapes asset pricing. Disclosed market, size, and value/growth risks align with return-based factor exposures in sign and magnitude, establishing that the disclosures are credible and informative. I then construct disclosure-sorted return spreads to measure intermediary-level risks, including redemption pressure, concentration, mandates, management, and operations. These risks are not spanned by standard factors and are priced in fund returns. Redemption and concentration also price individual stocks, showing that organizational frictions in asset management have asset-pricing consequences.

Keywords: *Mutual Funds; Risk Disclosures; Intermediary Risk; Asset Pricing; Corporate Finance of Asset Management; Organizational Frictions.*

*INSEAD (nikola.kiprijanovski@insead.edu). I am very grateful to Sergei Glebkin, Frederico Belo, and Semyon Malamud for their continued guidance and support. I thank Winston Dou and Wenqian Huang for discussing the paper. For valuable feedback, I thank Joel Peress, Massimo Massa, Wei Wu, Ben Charoenwong, Naveen Gondhi, Bernard Dumas, Boris Vallee, Xinyu Liu, Pierre Mabille, Olivier Dessaint, Ahmed Gueciouer, Alejandro Lopez-Lira, Junyuan Zou, Lin Shen, Jason Kang, Yiran Wang, Xinwei Li, and participants at the Young Scholars Finance Consortium (2026), Liechtenstein Workshop AI in Finance, Wharton-INSEAD Doctoral Consortium, INSEAD Brownbags, INSEAD PhD Seminar.

1 Introduction

Mutual funds are not passive containers of asset-level risk. They are financial intermediaries that collect capital from investors, allocate it across firms, and operate under contracts, mandates, liquidity promises, and flow-performance incentives. These institutional features shape the risks fund managers care about. A fund manager does not only face the risk that the underlying securities perform poorly. The manager also cares about investors redeeming capital, mandate constraints limiting portfolio adjustment, the tradeoff between attracting new capital after successful portfolio bets and facing career risk after unsuccessful ones, and operational complexity that makes the fund harder to manage, monitor, or market. These are risks of the intermediary, not simply risks of the assets.

This distinction matters because the mutual fund sector is large enough for its internal frictions to affect market outcomes. If many funds face similar constraints, their portfolio choices may become correlated for reasons that originate inside delegated asset management. The aforementioned redemption concerns, benchmark and mandate restrictions, concentration incentives, and operational constraints can therefore influence not only funds' operating margins and returns, but also the prices of the underlying assets. In this sense, mutual funds are more than passive vehicles where pre-existing risks are held and traded. They may also be an origination point of systematic risk.

Recent work from Dou, Kogan, and Wu (2020) shows that agency frictions in active mutual funds can generate priced risk. In their model, fund managers care about assets under management because fee revenue and compensation depend on fund size. Common fund-flow shocks therefore affect managers' objective functions, leading them to hedge flow risk through their portfolio choices. This hedging demand, in turn, makes common fund flows a priced risk factor in equilibrium. Their mechanism illustrates the broader idea motivating this paper: risks and constraints originating

inside delegated asset management can become systematic when many funds face similar incentives.

This paper takes that insight to the data from a different direction. Rather than beginning with a specific friction, such as common fund flows, I use mutual fund prospectus disclosures to ask which fund-level considerations managers themselves identify as material. These include redemption pressure, liquidity constraints, concentration, mandate rigidity, management risk, turnover, tracking error, and operational complexity among others. I then ask whether these disclosures are credible, whether they are associated with systematic differences in fund returns, and whether the same disclosed considerations are reflected in the broader cross-section of stock returns.

The paper relates to intermediary and institutional asset pricing, where constraints and incentives of financial institutions shape equilibrium prices.[1] It also relates to recent work that uses portfolio holdings to infer investor demand, preferences, and portfolio-relevant characteristics.[2] Holdings data reveal what intermediaries do, but they require the researcher to infer the risks and constraints that motivate those choices. Prospectuses provide a complementary source of information: they reveal, in the fund's own language, the risks and constraints managers are concerned with.

The empirical object of the paper is the "Principal Risks" section of mutual fund prospectuses. Mutual funds are required to describe, in plain language, the main risks of their investment strategy and portfolio. These disclosures are useful for two reasons. First, they are written at the fund level rather than the firm level, which makes them naturally suited for studying systematic risks, including those that could potentially arise from portfolio management and delegation. Second, they are forward-looking statements meant to inform investors of the risks they will be exposed to if

[1] See, among others, He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), Basak and Pavlova (2013), (Basak & Buffa, 2019), Pavlova and Sikorskaya (2022), Buffa, Vayanos, and Woolley (2022), and Cuoco and Kaniel (2011). Haddad and Muir (2025) provide a recent review.

[2] For example, the demand-based asset pricing literature uses holdings to estimate demand systems and demand elasticities (e.g., Koijen & Yogo, 2019), while recent work on asset embeddings uses holdings to recover firm-level characteristics relevant for portfolio choice (Gabaix, Koijen, Richmond, & Yogo, 2025).

they were to invest. Holdings and returns reveal what funds did, while disclosures go one level deeper and reveal what risks and constraints drive their behaviour.

The central question of the paper is therefore not whether text can be used to generate additional asset-pricing factors. Rather, it is whether fund managers, through their mandated disclosures, convey unique information about risks, in particular ones that stem from delegated asset management itself, that are created or amplified by agency conflicts, mandate constraints, flow-performance incentives, and the organizational structure of funds. Put more succinctly, the paper asks: which risks do mutual fund managers publicly identify as material, what can we learn from them, and do these intermediary-level risks matter for asset prices?

There are valid reasons to be skeptical that prospectuses can answer these questions, but the direction of disclosure incentives is ambiguous ex ante. Prospectus language may be boilerplate, shaped by legal caution, or designed to market the fund to investors. Managers may over-disclose risks to reduce litigation exposure, underdisclose risks to avoid deterring risk-averse investors, or use broad language that does not correspond cleanly to economic exposures. Yet these same institutional forces can also support accurate disclosure. Managers have incentives to describe material risks truthfully in order to maintain investor trust, set expectations about the fund's strategy, reduce reputational costs when risks materialize, and avoid inconsistencies between the fund's public description and its actual portfolio. Whether these statements contain credible information about the risks funds face is therefore an empirical question.

I begin by testing this credibility directly. Using the risk section in the prospectuses, I extract individual risk items, embed them in a semantic vector space, and cluster them into interpretable risk topics. I then compare these disclosed risks to return-based factor exposures. The validation focuses on familiar dimensions where the mapping is clear: market, size, and value/growth risk.[3] The results show that dis-

[3]The following are examples of "small-cap" and "value investing" risk disclosures:

closures align with exposures in both magnitude and direction. Funds disclosing market risk have higher market betas, while disclosures related to size and value correctly map onto the signed structure of the SMB and HML factors: small-cap and value risk disclosures predict higher corresponding betas, whereas large-cap and growth predict lower ones, respectively. These patterns hold at both the extensive and intensive margins, rejecting the view that the prospectuses are merely boilerplate and establishing that the text contains meaningful information about fund exposures.

Having established credibility on the basis of the smaller set of canonical risks, I shift focus to the full set of risks that include ones not well captured by standard return-based models. Many of the most important disclosed risks are not traditional asset-level ones. They concern the organization and operation of the fund itself: redemption pressure, portfolio concentration, management risk, turnover, tracking error, mandate restrictions, and operational complexity. These disclosures point to an organizational layer of risk that standard asset-pricing models do not directly measure. To quantify this layer, I construct disclosure-sorted return spreads by comparing funds that disclose a given risk with funds that do not. Because disclosure choices are only weakly correlated across risk topics, this sorting provides a valid identification strategy for isolating the return component associated with each disclosed risk.

The disclosure-sorted spreads are not spanned by existing factor models and are priced in the cross-section of mutual fund returns. The strongest and most consistent pricing patterns are associated with organizational risks, especially redemption risk, concentration risk, and risks related to management or operational complexity. This evidence supports the idea that the contribution of the disclosures is not simply that they produce additional factors. Rather, the disclosures emphasize intermediary-level frictions that are of economic significance for the funds themselves.

Small-Cap Securities Risk. Securities of smaller companies are usually less stable in price and less liquid than those of larger, more established companies. Therefore, they generally involve greater risk. (Aberdeen Global Equity Impact Fund, 2022).
Value Investing Risk. The determination that a security is undervalued is subjective. The market may not agree with the Adviser's determination and the security's price may not rise to what the Adviser believes is its full fair value. (Fuller & Thaler Behavioral Small-Cap Equity Fund, 2022)

The paper then asks whether these intermediary risks have effects beyond the mutual fund universe. If the disclosure-sorted spreads only captured fund-specific outcomes, they would not be expected to price individual stocks. However, I find that several disclosure-based risks, particularly those related to redemptions and concentration, are also priced in the broader cross-section of stock returns. This result is central to the paper's economic interpretation. It suggests that risks originating within delegated asset management can spill over into the pricing of the underlying assets.

To better understand the mechanism, I study the two most prominent organizational risks in more detail. Redemption risk disclosures identify funds that are smaller, face more volatile flows, hold larger cash buffers, own less liquid portfolios, and display weaker flow-performance sensitivity. These patterns are consistent with the idea that redemption concerns impose a real constraint on portfolio choice. Funds exposed to redemption pressure manage their portfolios differently, tilting toward securities that help them absorb or hedge investor-flow risk. When many funds respond to these constraints in similar ways, their demand generates systematic price pressure and lower subsequent returns for the affected securities. This behavior is related to the flow-hedging mechanism emphasized by Dou et al. (2020), but the evidence also shows that redemption disclosures capture more than a simple replication of common flow beta. The stocks most exposed to redemption-disclosing funds do not line up perfectly with the standard flow-hedging asset-selection channel, suggesting that disclosures capture a broader set of liquidity, clientele, and portfolio-management constraints.

Concentration risk points to a different agency channel. Funds disclosing concentration risk are mechanically more concentrated, have more volatile returns, and face more volatile flows. They also benefit asymmetrically from strong performance: concentration-disclosing funds receive higher baseline flows and additional flow rewards when alpha is high, but do not face a symmetric penalty when alpha is low. This pattern is consistent with convex flow-performance incentives, in the spirit of

Chevalier and Ellison (1997). These incentives direct capital toward securities that increase the variance of fund performance. In doing so, the demand of concentration-disclosing funds affects the pricing of the very assets they hold.

Taken together, the evidence supports the view that mutual fund disclosures reveal risks and constraints that emerge from within the asset management sector. These risks are not limited to the cash-flow risk of the underlying firms or to standard covariance with macroeconomic factors. They arise from the institutional structure of intermediation itself observed in the way funds are organized, evaluated, constrained, and compensated. The paper therefore contributes to asset pricing by showing that the set of relevant state variables includes the constraints and incentives of the intermediaries that allocate capital.

The paper makes three contributions. First, it contributes to the intermediary and institutional asset-pricing literature by providing empirical evidence that organizational frictions within mutual funds are reflected in returns. Existing work shows that delegated asset management can affect prices when fund managers' incentives differ from those of their investors. Dou et al. (2020) show that common fund-flow shocks become priced because active fund managers hedge the effect of flows on assets under management, while Cen, Dou, Kogan, and Wu (2023) provide direct evidence that fund flows are pertinent for managers' income and career outcomes. This paper complements these studies by using prospectus disclosures to recover a broader set of fund-level risks and constraints that managers identify as material. The results suggest that redemption risk is an important part of the story, but not the only one: concentration, mandate-related, management, and operational risks also appear to play significant roles for fund returns and for the pricing of the underlying stocks.

Second, the paper contributes to the growing literature applying textual analysis to mutual fund disclosures. Sheng, Xu, and Zheng (2023) and Krakow and Schäfer (2020) develop fund-level measures of disclosure quality: Sheng et al. (2023) define informativeness as the ability of disclosed risks to explain future idiosyncratic returns,

while Krakow and Schäfer (2020) quantify textual uniqueness relative to fund-family peers. In contrast, my goal is to use disclosures to better understand the risks and constraints faced by the intermediary itself. The paper is also related in spirit to Lopez-Lira (2023), who extracts priced factors from corporate 10-K risk disclosures using Latent Dirichlet Allocation. I differ both in domain and method: by focusing on fund-level prospectus disclosures rather than firm-level 10-K filings, I can study risks arising from delegated portfolio management, and by using embedding-based clustering I can represent the semantic content of disclosures in a more context-aware way.[4]

Third, the paper contributes to the mutual fund literature by showing that agency problems and organizational constraints are visible in mandated risk disclosures. Prior work studies whether active managers add value (Berk & Green, 2004), how skill and compensation are measured (Ibert, Kaniel, Van Nieuwerburgh, & Vestman, 2017; Kaniel, Lin, Pelger, & Van Nieuwerburgh, 2023), and how funds compete, communicate, or differentiate their products using prospectus language (Abis, 2020; Abis & Lines, 2021, 2022; Kostovetsky & Warner, 2020). The paper is also related to work on risk-taking incentives in delegated asset management. Chevalier and Ellison (1997) provide early evidence that mutual fund managers adjust portfolio risk in response to convex flow-performance incentives, while Rajan (2006) emphasizes that investors' inability to observe risk can create agency problems and excessive risk taking. I complement this literature by showing that fund managers, through their mandated disclosures, convey informative signals about the risks and constraints that shape their portfolio choices, and that these signals reflect forces affecting both fund returns and the pricing of underlying stocks.[5]

More broadly, the paper also connects to work showing that textual data can pre-

[4]The embedding-based approach relies on transformer-based models to represent each risk disclosure as a high-dimensional semantic vector. This allows the clustering procedure to group disclosures by meaning rather than by word overlap alone, which is particularly useful when funds describe similar risks using different language.

[5]While not specifically about mutual funds, Barth, Monin, Siriwardane, and Sunderam (2024) also study "hidden" risks in asset management. Their setting differs from ours in two key ways: they focus on hedge funds rather than mutual funds, and they analyze quantitative reports of market risk exposures, whereas this paper studies qualitative, narrative risk disclosures.

dict financial outcomes. Studies such as Bellstam, Bhagat, and Cookson (2021), Ke, Kelly, and Xiu (2019), Bybee, Kelly, and Su (2023), and Hwang and Kim (2017) show that natural language contains information relevant for innovation, returns, and other financial outcomes. My paper complements this work by showing that even structured regulatory disclosures from mutual funds, often viewed as boilerplate, contain economically meaningful information when analyzed with tools that preserve their semantic content.

The remainder of the paper proceeds as follows. Section 2 describes the prospectus data, the extraction of risk disclosures, and the embedding-based topic model. Section 3 validates the credibility of the disclosures by comparing disclosed risks to standard factor exposures. Section 4 constructs disclosure-sorted return spreads and tests whether disclosed risks are priced in the cross-section of mutual fund returns. Section 5 examines whether these risks are priced in individual stock returns. Section 6 studies the mechanisms behind redemption and concentration risk. Section 7 concludes.

2 Data

I use multiple sources of data for my analysis: the Mutual Fund Prospectus Risk/Return Summary Data Sets provided by the SEC, CRSP, and multi-factor model time series returns.

2.1 CRSP and Factor Models

This study uses the CRSP Survivorship-Bias-Free Mutual Fund Database. For the current scope of the project, we limit the sample to diversified, open-end U.S. equity mutual funds, although the same research procedure can be used for any type of fund, and these extensions are a work in progress. U.S. domestic equity funds are identified using the CRSP Style codes, filtering on the Level 1 character being an "E" (for equity).

For each fund, the dataset provides after-fee raw returns, total net assets (TNA), expense ratios, turnover ratios, and other characteristics. A single fund, referred to as a series by the SEC, can have multiple share classes available for it, each with different characteristics. When multiple share classes exist, fund-level variables like returns, expense ratio, and management fees are computed by TNA-weighted aggregation across share classes. Daily return time series of the five factors of Fama and French (2015) and the momentum factor come from the French data library on Ken French's website.

2.2 Mutual Fund Prospectus - Risk Section

A mutual fund prospectus is a document outlining the fund's investment strategy, risks, fees, past performance, and other key attributes. The Securities Act of 1933 mandates that all publicly offered investments, including mutual funds, must be registered with the SEC and accompanied by a prospectus that provides prospective investors with key information about the investment. In addition, mutual funds are also subject to the Investment Company Act of 1940, which imposes further requirements on the structure, governance, and disclosures of investment companies, including the content and periodic updating of prospectuses. The SEC began mandating electronic submissions of prospectuses by mutual funds in 1996 with the introduction of the EDGAR (Electronic Data Gathering, Analysis, and Retrieval) system.

In June 2009, the SEC adopted a rule (Release IC-28617) requiring mutual funds to submit risk/return summary information in the eXtensible Business Reporting Language (XBRL) format as an exhibit to their registration filings on Form N-1A. The requirement applied to filings starting with initial registration statements and post-effective amendments that became effective on or after January 1, 2011. The XBRL requirement did not apply to the full prospectus, but only to key sections of the risk/return summary, such as: Investment objectives and strategies, Risks, Fees and expenses, and Performance data. What this requirement meant, was that

funds were responsible for providing explicitly tagged and structured data on these sections, allowing for easier consumption of the data for analysis. All submissions following this rule are archived and available for download from the SEC website as Mutual Fund Prospectus Risk/Return Summary Data Sets. The submissions date from December 2010 to the current date, being updated quarterly. For this paper we use data from January 2011 to December 2023 [6]. Multiple types of submissions are related to the prospectus, including 485APOS, 485BPOS, and 497. The first two, called form N-1A, deal with statutory i.e. comprehensive submissions, the initial one, and ongoing ones, respectively. The third one deals with changes or amendments to the statutory one. In this paper we only use the statutory ones, as they are the only ones mandated by the SEC, as opposed to the voluntary nature of the other type. All of the textual data found in these datasets has to satisfy the 1998 Plain English Disclosure — SEC Release No. 33-7497, also known as the Plain English Rule, which ensures that disclosures—particularly those intended for retail investors—are clear, concise, and understandable.

In this paper, the main object of interest is the risk section of the prospectus, officially titled "Principal Risks of Investing in the Fund". The goal of this section is to provide investors with a clear, concise, and balanced description of the principal risks of investing in the fund. As per the SEC's Form N-1A Instructions, Funds have to "summarize the principal risks of investing in the Fund, including the 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.". In practice, the risk section typically begins with a brief introductory paragraph explaining its purpose and underscoring the importance of understanding the risks involved before investing. What follows is usually a list of specific risks, each presented under a distinct heading, often named after the risk itself, such as "Stock Market Risk." or "Operational Risk", or omitting the word risk entirely "Concentration". In the PDF and HTML ren-

[6]Extending the dataset to include all submissions going back to 1996 before the XBRL tagging is an ongoing project.

derings, these items are usually formatted to stand out through bolding, underlining, or other typographical means. After a heading, a description of the risk is usually in order. Here is where the risk is explained and its relevance, as well as its potential impact on the fund's performance, is shown.

2.2.1 Risk Extraction

This paper's analysis begins by converting qualitative risk narratives into a structured, quantitative format. The goal is a granular panel dataset that details, for each fund and year, every specific risk it discloses.

To create this dataset, we isolate the textual pair of a risk heading and its associated description. This approach allows me to analyze a fund's disclosures on both the extensive margin (whether a risk is mentioned) and the intensive margin (the share of text it occupies). This dataset underpins the paper's main analyses, including the validation of disclosures against factor exposures and the construction of new risk factors. The extraction process itself is a significant hurdle, as funds are not required to follow a standard format for these documents.

In fact, great variability is observed in the cross-section of the risk disclosures in all dimensions. Funds can refer to the same risk, while describing it differently, or vice versa. What makes the extraction feat particularly difficult is the non-conformity in the way how funds delimit and format the text for different risks. Beginning in 2023, the SEC's Risk/Return Summary Taxonomy, which outlines the rules for proper digital data structuring and tagging, was changed to include a method for tagging all principal risks disclosed, essentially requiring the funds to self-identify all heading-description pairs. While this is exactly the format we desire, only one year of data is affected by the change, leaving the majority of the data still in need of processing. Examples of risk sections are given in the Appendix.

To extract the risk heading-description pairs, we utilize the help of a commercial

Large Language Model (LLM), specifically Google's lightweight model, Gemini 2.0 Flash. We construct a prompt with instructing the model to transform raw mutual fund prospectus text into a clean, structured format where each risk is clearly identified and separated under a labeled heading. Risk headings and descriptions are extracted as are, with the allowance of slight error correction, usually to fix digital artifacts which appear when converting HTML to plain text or similar. Introductory and concluding paragraphs, which do not talk about explicit risks, are removed. Connecting over the Google Gemini API, we feed every text with the same instructions. We end up with over 150,000 cleaned prospectuses, 83,000 of which are for equity mutual funds. Across all of the prospectuses, we find over 260,000 unique risk heading-description pairs. We are thus able to proceed with my analysis which is at the level of fund-year-risk, as we know every risk for every fund for every year.

The reason we choose to employ this new methodology is to avoid the limitations of the alternative, which in this case means using the entire non-cleaned texts. While there is still a wealth of information within them, we would be limited to using a narrower set of tools for analyzing them, such as the bag of words approaches, word frequency or term counts, dictionary or lexicon based methods etc., which while useful in their own right, are currently surpassed by newer, more powerful, and more comprehensive tools and techniques like topic modeling, embeddings, and (un-)supervised learning, all of which we use in this paper.

To verify the credibility of this procedure we employ several steps. First, for the 2023 batch of prospectuses, there exists a subset of funds that provides both the "old" format, that is the plain text risk section that is processed as described above, as well as in the "new" format, i.e. with explicitly tagged risk headings and deliniated descriptions. We cross-validate this several hundered fund sized subset, and find perfect alignment between both. Next, we refer the results of a pre-LLM manual procedure we did to extract the risks, involving semi-automated cleaning through the extensive use of regex (regular expression) code. The LLM powered extraction was aligned with

the parts of my dataset that we had verified to be correct, and were beating my attempt in the parts where my procedure had given poor results. Additionally, I have manually validated the extracted data for more than a thousand prospectuses.

The resulting panel dataset is extensive. As summarized in Table 1, the sample covers over 83,000 fund-year observations between 2011 and 2023. The table documents a clear trend toward more extensive disclosure over time: the average number of distinct risks disclosed per fund nearly doubled, from 8.65 in 2011 to 16.07 in 2023, while the average length of each description grew by over 55%. This growth in both the breadth and depth of disclosure underscores the richness of the dataset and motivates the need for a scalable analytical approach to distill its informational content.

2.2.2 Text Embeddings and Clustering

Having extracted each risk heading-description pair, and established the fund-year-risk level, we now turn my attention to the question of what is actually being said in the disclosures. The answer to this question is of crucial importance, as it will allow me to link disclosed risks to existing ones, find new risks, as well as differentiate between disclosures that might seem related on a superficial level. While the problem of assigning and using the semantic meaning as an analytical measure is difficult by itself, here, it is compounded by the fact that disclosed risks are very numerous and diverse. Thus, to gain any semblance of statistical power, we need to reduce the high-dimensionality of disclosed risks to a manageable level. In order to do both, we use a state-of-the-art text analysis procedure of embedding-based topic modeling.

Word embeddings represent each risk description as a dense, high-dimensional (in this case, n=1024) numerical vector that encodes the semantic meaning of the text. Unlike traditional approaches such as bag-of-words or TF-IDF, which represent text as sparse vectors based on word frequencies and ignore word order and context, embeddings capture complex relationships between words by placing them in a contin-

uous vector space where distance and direction have interpretable semantic meaning. Texts that are conceptually similar—even if they use different terminology, phrasing, or structure—are located close to one another in this embedding space. This makes embeddings especially well-suited for aggregating heterogeneous textual data like mutual fund risk disclosures, which vary in language but often convey similar underlying ideas.

To generate these embeddings, we employ Stella 1.5B, a large-scale transformer-based language model that has been benchmarked as one of the top-performing embedding engines, with particularly strong performance on clustering tasks. Stella 1.5B, like many other models, is trained on a large corpus of technical, financial, and general text, and is particularly adept at preserving fine-grained distinctions between similar but non-identical texts. Its training follows a two-stage process: first, the model learns to predict missing words in a sentence (a task called masked language modeling), which forces it to understand grammar, context, and the relationships between words. Then, in a second phase called contrastive learning, it is shown pairs of sentences that are either similar or different in meaning, and is trained to place similar ones close together in vector space while pushing dissimilar ones apart. This shapes the embedding space so that in our case passages expressing the same risk, like "exposure to equity market swings" and "broad stock volatility" end up geometrically close, even if they share no words.

Each risk description is passed through the model to obtain a 1,024-dimensional vector. This embedding step forms a crucial foundation for subsequent clustering, as it abstracts from the raw text and maps disclosures into a semantic space where conceptually similar risks are geometrically aligned. This transformation ensures that when disclosures are grouped into topics using clustering algorithms, the resulting clusters reflect genuine thematic coherence rather than superficial lexical similarity. It also avoids the issue of mislabeled headings mentioned above, since the embeddings are agnostic to the heading. This allows the underlying structure of the data to

emerge organically, without relying on predefined taxonomies or manually imposed categories. By limiting the role of subjective judgment in categorizing risks, this approach improves both the scalability and reproducibility of the analysis. The resulting clusters capture how funds themselves articulate risk, helping to mitigate concerns that the identified themes simply mirror conventional expectations of what fund risks are assumed to be. This data-driven organization of risk topics underpins all subsequent empirical analyses in the paper.

Once embedded, we apply k-means clustering to group these vectors into distinct topics. This unsupervised algorithm partitions the semantic space into k clusters, each of which represents a broader "risk topic" composed of conceptually similar disclosures. We experiment with various values of k$\in \{5, 10, 15, 20, 25\}$, which allows me to balance thematic granularity against statistical power. Traditional methods for selecting the optimal number of clusters (e.g., elbow plots or silhouette scores) perform poorly in this context, likely due to the unusually high number of observations and the high-dimensionality of the embedding space. As the results, and thus the quality of fit, of the k-means clustering procedure depend on the initial starting points, we employ both k-means++ initialization and run the algorithms 100 times each, taking as final the run that produced the lowest total within-cluster sum of square (WSS), i.e. lowest within-cluster variance.

To interpret the resulting clusters as risk topics, we return to the original risk headings associated with each description. For each cluster, we compute the most frequently occurring headings, which provide a natural summary of the dominant themes. For example, in the k=5 case, the first cluster includes high frequencies of headings such as management risk (23,694 mentions), redemption risk (16,984), non-diversification risk (13,908), market risk (10,119), and portfolio turnover risk (10,013), along with related terms like concentration, sector, valuation, and tracking error risks. This distribution suggests the cluster reflects risks tied to fund structure, operational complexity, and portfolio composition. Another cluster in the k=5 case is instead dom-

inated by classic equity market terminology. The most frequently occurring headings include market risk (31,873 mentions), equity securities risk (24,168), equity risk (9,242), stock market risk (5,628), and value investing risk (4,350). The cluster also features related terms such as growth stock risk, investment style risk, equity market risk, and stock market volatility. Taken together, these headings suggest that the cluster captures risks associated with equity market exposure and style-based investing. A particularly instructive feature of the highlighted results is that the heading "market risk" appears prominently in both clusters, despite the clusters reflecting distinct themes. This occurs because the descriptions associated with that heading vary across filings: some are more similar in language and content to disclosures in the first cluster, while others align more closely with the themes of the second. The clustering algorithm assigns them accordingly, based on the semantic similarity of the full descriptions rather than the heading alone. That these similarly titled disclosures are separated into different clusters reflects a key strength of the embedding-based approach: it groups risks based on the meaning conveyed in the full text, not just on superficial labels. As a result, the method naturally accommodates variation in language and avoids misclassification that would arise from simpler keyword- or dictionary-based techniques, which treat identical headings as identical concepts regardless of context. As a second step for risk topic validation, we read through sampled descriptions from each of the clusters to validate that they are broadly semantically close within cluster, and different across clusters.

In general, lower values of k yield broader groupings (e.g., "Management Risk" that combines management, redemption, and derivatives risks), while higher values produce more fine-grained distinctions. Throughout the paper we take no stance as to what the correct k is, and we report the results for all of them. The result is a structured dataset where each fund-year observation contains the full list of disclosed risk topics and the proportion of disclosure text devoted to each one. This enables a host of downstream analyses, from measuring alignment with traditional factor exposures to constructing novel risk-mimicking portfolios. Summarized descriptions of the risk

topics for each level of k are provided in Appendix C.

3 Disclosures and Exposures

Establishing a credible baseline for the reliability of mutual fund risk disclosures is a necessary prerequisite for any subsequent analysis. It is not obvious ex ante that these disclosures credibly reflect the risks funds face, given the conflicting incentives of managers when writing prospectuses.

Funds report risks in the context of delegated portfolio management, where disclosure serves not only to inform but also to market the product. As such, managers may be biased toward "vague and boilerplate risk factor disclosures that attempt to catalog every conceivable problem that might surface" to avoid deterring potential investors (Spindler, 2006). Additionally, disclosures can serve as a "litigation shield," creating an incentive to disclose risks indiscriminately, regardless of the fund's actual exposure. These forces could lead to an uninformative, stock-language filled text that bears little relation to a fund's true risk profile.

On the other hand, maintaining trust with investors is paramount, and providing accurate, credible information may be rewarded with more stable inflows. The literature on corporate 10-K filings provides some support for this view. For example, Campbell, Chen, Dhaliwal, Lu, and Steele (2014) find that "managers provide risk factor disclosures that meaningfully reflect the risks they face," while Gaulin (2017) shows that managers time the addition and removal of risk factors to align with expected future outcomes, consistent with providing useful information to investors rather than just boilerplate.

The closest formal treatment of this issue is done by Lopez-Lira (2021), who theoretically examines the problem of risk disclosure choice for firms. Managers optimally decide how much space to allocate to each risk by weighing the legal costs of underdisclosure against the effort costs of writing. The model predicts that, under reasonable conditions, firms disclose their risks truthfully and write longer sections for the risks to which they are more exposed, making disclosure length a monotonic function

of true exposure. This yields a testable prediction: disclosure intensity should be a monotonically increasing in the true risk exposure. While this framework is developed for corporate 10-K filings, its core economic logic provides a powerful and, until now, untested hypothesis for the mutual fund industry. This motivates the central empirical question of this section: Are mutual fund risk disclosures credible?

The theoretical foundation transforms this general inquiry on credibility into a sharply-defined empirical test. If similar economic trade-offs hold in the delegated asset management context, we should observe a direct correspondence between quantitative measures of risk and the narrative emphasis placed on them in prospectuses. We operationalize this test by using estimated Fama–French factor loadings as an objective, market-based proxy for a fund's true risk exposure and our textual measures, both mentions (the extensive margin) and the text proportion (intensive margin), as proxies for disclosure intensity. Finding a significant, positive relationship between a fund's factor beta and its disclosure on the corresponding risk topic would provide the first empirical validation of this disclosure theory for mutual funds and, in doing so, would robustly establish the credibility of our textual data.

To conduct this test, our first step is to define the overlapping set where both quantitative exposures and qualitative risk topics are well-defined. We begin with the most widely accepted and empirically validated risk factors: the five-factor model of Fama and French (2015), augmented with the momentum factor to yield a six-factor specification. On the disclosure side, we verify that risk topics corresponding to "Market" and "Capitalization" appear consistently across all clustering levels, while a "Value" topic emerges starting at k=10. The other factors of profitability, investment, and momentum do not have clear textual counterparts. As such, our analysis focuses on the relationship between a fund's estimated exposures to the MKT, SMB, and HML factors, and its disclosure of the corresponding "Market," "Capitalization," and "Value" risk topics, respectively.

Before describing the empirical tests, one additional data adjustment is required.

The SMB and HML factors, as constructed by Fama and French (1993), are defined as return spreads between two portfolios—small minus big, and high minus low book-to-market, respectively. This structure implies that any meaningful comparison with disclosures requires observing both "legs" of each factor in the disclosure risk topics. However, the unsupervised k-means procedure used to construct disclosure-based topics does not always make this distinction. For instance, the "Capitalization" risk topic includes both small-cap and large-cap risk descriptions, and the "Value" topic contains observations that refer to both value and growth stock investing. Using such mixed-topic disclosures in the regression analysis would yield misspecified estimates, as they do not cleanly correspond to either side of the underlying factor, nor do they mimic their "long-short" nature. If left unadjusted, the SMB estimates computed for funds that disclose "Capitalization" risk, a grouping that combines both small-cap and large-cap disclosures, will effectively be pooled across funds that lie on opposite ends of the SMB dimension, and thus attenuating the relationship in the aggregate.

To address this, we apply a second round of clustering within the relevant topics. Specifically, we isolate the "Capitalization" topic and re-cluster the associated descriptions using k=2, with the same procedure as before (k-means++ initialization and 100 iterations). The result is two clearly distinguishable sub-topics: one focused on small-cap risk and the other on large-cap risk. We apply the same approach to the "Value" topic, successfully separating it into distinct "Value Stock" and "Growth Stock" risk clusters. This additional refinement ensures conceptual alignment between the empirical factors and the disclosed risk topics. Moreover, the internal consistency of these re-clustered results provides further validation of the embedding-based topic modeling framework: even when applied in a narrow, local setting, the method produces intuitive and economically meaningful distinctions. In fact, this adjustment becomes unnecessary for large enough k, as for $k = 25$, the "Value" and "Growth" topics come out organically.

With this conceptual alignment established between the textual topics and the fac-

tors, we can now formally specify the empirical test of our credibility hypothesis. The testing procedure is designed to quantify the relationship between a fund's quantitative risk exposures, proxied by its factor betas, and its narrative disclosures. The procedure begins by estimating each fund's yearly factor loadings. For each fund i and year t, we estimate the six-factor model using daily returns from the preceding year. This choice naturally follows from the requirement that funds update their prospectus annually. Formally, we estimate:

$$\begin{aligned} R_{it} = \alpha_i + \beta_{i,MKT} MKT_t + \beta_{i,SMB} SMB_t + \beta_{i,HML} HML_t + \beta_{i,RMW} RMW_t \\ + \beta_{i,CMA} CMA_t + \beta_{i,MOM} MOM_t + \epsilon_{it} \end{aligned} \tag{1}$$

Using these estimated betas ($\hat{\beta}^f_{i,t}$) as the dependent variables, we then run the following cross-sectional OLS specification for each factor $f \in \{\text{MKT, SMB, HML}\}$:

$$\hat{\beta}^f_{i,t} = \gamma^f_0 + \sum_{l=1}^{k} \gamma^f_l \, \text{risk_topic_measure}_{l,i,t} + \varepsilon^f_{i,t} \tag{2}$$

where $\text{risk_topic_measure}_{l,i,t}$ is one of two proxies for disclosure intensity: (1) a dummy variable for whether fund i mentions topic l in its prospectus for year t (the extensive margin), or (2) the proportion of the risk section text that fund i devotes to topic l (the intensive margin). The regression includes all k risk topics for a given clustering level, ensuring that the coefficients capture the incremental information in a specific disclosure.

The overarching null hypothesis in this section (H_0) is that prospectus language is uninformative boilerplate, implying that the coefficients on all textual measures in all specifications are zero ($\gamma^f_l = 0$ for all l, f). Sharpening this notion further, for each factor, we test a directional null hypothesis against the alternative that disclosures are informative:

- **Market Factor Beta ($\hat{\beta}^{MKT}$):** We test the null $H_0 : \gamma^{\text{MKT}}_{\text{Market Risk}} \leq 0$. Rejection in fa-

vor of a significantly positive coefficient would indicate that measures of market risk disclosure align with realized market betas.

- **Size Factor Beta ($\hat{\beta}^{SMB}$):** We test the joint null hypothesis that $H_0 : \gamma^{\text{SMB}}_{\text{Small-Cap}} \leq 0$ and $H_0 : \gamma^{\text{SMB}}_{\text{Large-Cap}} \geq 0$. Rejecting this null in favor of the predicted opposite signs is a stringent test that requires disclosures to correctly identify a fund's position along the long-short SMB dimension.

- **Value Factor Beta ($\hat{\beta}^{HML}$):** Similarly, we test the joint null $H_0 : \gamma^{\text{HML}}_{\text{Value}} \leq 0$ and $H_0 : \gamma^{\text{HML}}_{\text{Growth}} \geq 0$.

Furthermore, a proper test of credibility requires not only statistical significance and the correct sign, but also economic significance in terms of relative magnitude. If the textual measures are capturing a clean, specific channel of information, then the coefficients on the focal topics (e.g., the $\gamma^{\text{SMB}}_{\text{Small-Cap}}$ in the $\hat{\beta}^{SMB}$ regression) should be economically large and dominate the coefficients on other, theoretically unrelated topics (e.g., $\gamma^{\text{SMB}}_{\text{Derivatives Risk}}$), which we expect to be statistically and economically close to zero. Such a pattern would provide strong evidence for the discriminant validity of our measures and mitigate concerns that the observed relationships are driven by spurious correlations. Thus, a convincing rejection of the uninformative boilerplate hypothesis requires a pattern of large, correctly-signed coefficients on the focal topics for their respective factor betas, and a corresponding set of small, insignificant coefficients on the non-focal topics.

The results, presented in Tables 2 through 16 in Appendix A.1, offer a comprehensive and resounding rejection of the uninformative boilerplate hypothesis. The alignment between narrative content and quantitative exposures holds across all three factors, at both the extensive and intensive margins, and is remarkably consistent across all levels of topic granularity ($k \in \{5, 10, 15, 20, 25\}$).

3.1 Market Factor Beta ($\hat{\beta}^{MKT}$)

Beginning with the market factor, the evidence shows that funds' disclosures of market sensitivity are strongly aligned with their realized MKT betas. Across all specifications (Tables 2, 5, 8, 11, and 14), topics related to general market or equity risk are strong positive predictors of a fund's market beta. This result is notable because, given that all funds in the sample are equity funds, one might expect little variation in either their exposures or their disclosures (as is assumed by Sheng et al. (2023)).

Across all specifications topics related to general market or equity risk are strong positive predictors of a fund's MKT beta. For instance, the $k = 15$ specification (Table 8), the extensive measure yields the largest coefficient for the disclosure mention of "Equity Risk," at 0.113. Funds that mention "Equity Risk" therefore have, on average, higher market betas than those that do not. At the intensive margin, the coefficient on "Equity Risk" intensity is 0.611, again among the largest in magnitude. Interpreted economically, a fund that devotes 75% of its risk disclosure to the "Equity Risk" topic is predicted to have a market beta roughly 0.305 higher than a fund that only devotes 25% to it.

Interestingly, the topic labeled "Market Risk" performs less strongly, with an intensity coefficient of 0.398 and a dummy coefficient of 0.03. This seemingly puzzling result is, in fact, a validation of the nuance captured by our methodology. An examination of the semantic content of these topics (see Appendix C.3) reveals that the "Equity Risk" topic is more focused on risks specific to the stock market and individual securities. In contrast, the "Market Risk" topic encompasses broader economic risks. Given that the Fama-French MKT factor represents the excess return of a broad portfolio of equities, not a direct proxy for macroeconomic fundamentals, it is economically consistent that the "Equity Risk" topic demonstrates a stronger alignment. This distinction becomes even sharper as topic granularity increases. In the $k = 25$ specification (Table 14), "Equity Risk" has the largest coefficient (0.581), while the coefficient

on "Market Risk" falls (0.286) and a newly distinct "Macro Risk" topic appears with an even smaller, negative coefficient (-0.117). This demonstrates the power of an objective, data-driven topic modeling approach. A manual classification would likely have grouped these risks together under a single "market" label, obscuring the economically meaningful distinctions that the model uncovers.

Across all specifications, other equity-related topics—such as those referring to Value, Size, or Sector risks—also load positively on MKT betas, whereas fixed-income–related topics (Credit, Interest Rate, High Yield) tend to have negative coefficients, and unrelated topics cluster around zero. Overall, these patterns provide strong evidence of discriminant validity for our textual measures. We therefore reject the null hypothesis $H_0 : \gamma^{\text{MKT}}_{\text{Market Risk}} \leq 0$ and conclude that fund disclosures credibly reflect market exposure.

3.2 Size Factor Beta ($\hat{\beta}^{SMB}$)

The size factor provides a stronger test of disclosure credibility, as it requires that disclosures align with the directional nature of the SMB factor. The evidence strongly supports the credibility assertion and is remarkably consistent across all specifications (Tables 3, 6, 9, 12, and 15).

The regressions reveal a stark and symmetric opposition between small-cap and large-cap disclosures. At the intensive margin, disclosures of "Small-Cap Risk" are associated with large, positive, and highly significant coefficients. Again, taking the $k = 15$ specification as an example (Table 9), the intensity coefficient is 1.200. This magnitude is economically substantial: a fund dedicating 20% of its risk section to small-cap risks is predicted to have an SMB beta that is 0.12 points higher than a fund that only dedicates 10%. In stark contrast, disclosures of "Large-Cap Risk" carry a large and highly significant negative coefficient of -0.608. This finding is robust at the extensive margin as well. Simply mentioning "Small-Cap Risk" is associated with an

SMB beta that is 0.205 higher, while mentioning "Large-Cap Risk" corresponds to a beta that is 0.097 lower. This pattern is not an artifact of a specific clustering choice; it is present in every specification from $k = 5$ to $k = 25$. This symmetric, opposing effect across both margins is precisely what we predicted and constitutes powerful evidence against the boilerplate hypothesis. Managers are not merely vaguely mentioning "size"; they are accurately communicating their portfolio's specific tilt along the small-versus-large dimension.

The model again reveals other economically intuitive correlations: "Redemption Risk" for instance, consistently loads positively on SMB beta. This is consistent with the well-known fact that small-cap stocks are, on average, less liquid. Similarly, topics like "Foreign Risk" and "Emerging Market Risk" load negatively, which aligns with the fact that many prominent international and emerging market firms are large-capitalization entities. Meanwhile, theoretically unrelated topics, such as "Management Risk" have coefficients that are statistically and economically indistinguishable from zero. This pattern of large and correctly signed coefficients on focal topics, sensible coefficients on related topics, and zero coefficients on unrelated ones is strong evidence against the null of non-informativeness, and so we can decisively reject the joint null hypothesis that $H_0 : \gamma^{\text{SMB}}_{\text{Small-Cap}} \leq 0$ and $H_0 : \gamma^{\text{SMB}}_{\text{Large-Cap}} \geq 0$.

3.3 Value Factor Beta ($\hat{\beta}^{HML}$)

The analysis of the HML factor completes our validation exercise by testing whether managers' qualitative descriptions of their investment style (value or growth) are credibly reflected in their quantitative portfolio tilts. The results, detailed in Tables 4, 7, 10, 13, and 16, are once again perfectly aligned with our directional hypotheses.

The regressions once again document a dichotomy between disclosures of value and growth investing. Disclosures of "Value Risk" consistently predict the large positive HML betas. In the k=15 specification (Table 10), the intensity coefficient is 0.541

and is highly significant. Conversely, disclosures of "Growth Risk" predict large negative HML betas, with a coefficient of -1.010. This strong, symmetric relationship also holds at the extensive margin: simply mentioning "Value Risk" is associated with a 0.119 higher HML beta, while mentioning "Growth Risk" predicts a beta that is 0.128 lower. This confirms that the language managers use to describe their investment style is not cheap talk; it translates into their portfolio's quantitative factor exposures.

The validity of our measures is again supported by the behavior of non-focal topics. Most risk topics show no consistent, significant relationship with HML betas. The clear pattern of large, correctly signed coefficients on the style-related topics, combined with the near-zero coefficients on unrelated ones, provides a robust basis for rejecting the joint null hypothesis $H_0 : \gamma^{\text{HML}}_{\text{Value}} \leq 0$ and $H_0 : \gamma^{\text{HML}}_{\text{Growth}} \geq 0$. These findings confirm that mutual fund disclosures credibly capture systematic exposure along the value–growth dimension as well.

3.4 Summary and Implications for Disclosure Credibility

The preceding analyses provide a compelling body of evidence that systematically rejects the uninformative boilerplate hypothesis, and in turn provides ample support for the credibility claim of risk disclosures. The validation is based on three distinct, but complementary pillars of evidence. First, at a basic level, we find a direct correspondence: disclosures concerning market, size, and value risks are significantly associated with their respective factor betas. Second, and more strikingly, the analysis confirms the precise directional predictions for the long-short SMB and HML factors, demonstrating that managers' disclosures accurately reflect not just the dimension of risk but also the portfolio's specific orientation along that dimension. Third, the results demonstrate strong specificity of the textual measures: coefficients on focal topics are economically large and correctly signed, whereas those on unrelated topics are close to zero.

The findings from the intensive margin are particularly revealing. The finding that a greater proportion of text devoted to a specific risk corresponds to a larger exposure suggests that these disclosures are not a simple qualitative checklist. Instead, they represent a nuanced communication channel through which managers convey quantitative information about the magnitude of their exposures. The findings align with the model of costly and legally incentivized disclosure, but do not rule out alternative explanations, as similar behavior could also reflect reputational or relationship-based incentives to communicate risk credibly.

This robust validation is the critical foundation upon which our subsequent analyses are built. Having established that narrative disclosures are not noise but a credible signal of risk, we can now proceed to our central objective: leveraging this novel data source to identify and price dimensions of risk that are invisible to conventional, return-based models.

4 New Priced Risk Factors from Disclosures

The preceding analysis provides a crucial validation: mutual fund risk disclosures are not boilerplate filler but credible reflections of funds' economic exposures. This finding validates the reliability of the textual data, but this alone does not exhaust their informational value. The breadth of the disclosures hints at risks that extend well beyond the canonical factors. They encompass a wide array of risk concepts, many of which fall outside traditional asset-pricing frameworks, including risks tied to the securities held by the fund (e.g., derivatives, sector concentration, foreign stocks), as well as to the fund's internal organization and operations (e.g. management, redemptions, cybersecurity). With the goal of improving our understanding of risk, this observation raises a broader and more fundamental question: do these disclosures contain new and unique information that reflects genuine risk, above and beyond that captured by established factors?

In asset pricing, risk itself is not directly observable. Its existence and significance are inferred from its effect on asset returns. A risk is considered systematic and economically meaningful if it is priced, that is, if investors demand compensation in the form of a risk premium for bearing it. This principle transforms and disciplines our conceptual inquiry into a testable empirical hypothesis: if the risks detailed in prospectuses are indeed what they are claimed to be, they should manifest as priced factors in the cross-section of returns. Testing for a significant risk premium thus becomes the method by which we validate the economic substance of the disclosures.

This provides a natural starting point for our investigation. The most logical first test is to examine whether these disclosure-based risks are priced within the universe of the assets from which they are derived: mutual funds. If the risks described by fund managers do not explain variation in their own returns, their broader relevance would be highly questionable. Therefore, this section investigates whether risk factors, constructed based on topic disclosure earn a significant risk premium.

To do this, we first construct textual risk factors for each disclosure topic by forming portfolios of funds based on their disclosures. We then test the novelty of these factors by employing the Gibbons, Ross, and Shanken (GRS) test to determine if their returns are simply redundant representations of existing canonical factors. Finally, having established their novelty, we use two-stage Fama-MacBeth regressions to formally test whether these textual factors are priced in the cross-section of mutual fund returns and to quantify how much they improve the explanatory power of standard models.

4.1 Factor Construction

To translate the qualitative risk topics into quantitative test assets, we construct risk-mimicking portfolios in the spirit of Fama and French (1993). For each risk topic j within a given set of k topics, we construct a "Have-minus-Not" (HMN) textual risk factor. This factor is the return on a zero-investment portfolio that goes long an equal-weighted portfolio of all funds that disclose risk topic j in their most recent prospectus (the "Have" portfolio) and short an equal-weighted portfolio of all funds that do not (the "Not" portfolio). This procedure yields a daily return time series for each textual factor, representing the return differential associated with the disclosure of a given risk topic.

The validity of this approach as a test for priced risk hinges on a crucial empirical feature of the data: disclosure patterns across topics, whether measured by mentions or intensity, are largely uncorrelated. This is evidenced in Figures 1-5, which show the correlation matrices of the risk topic mentions across all the specifications. The low correlations that we observe are the cornerstone of my identification strategy, as it implies that funds are not systematically sorted on multiple disclosures simultaneously. As a result, the "Have" and "Not" portfolios for a given topic are, on average, balanced along other, unrelated disclosed risk dimensions. This balancing ensures that the resulting HMN return spread is plausibly attributable to the specific risk be-

ing disclosed, rather than to a confounding, correlated topic. Therefore, if a disclosed topic captures a priced dimension of risk, its corresponding HMN factor should earn a significant, non-zero premium.[7] This construction provides a direct test of whether the narrative language funds use to describe their risks corresponds to differential expected returns in the cross-section.

4.2 Are Disclosed Risks Spanned by Existing Factors?

Before testing whether the HMN factors are priced, it is essential to determine if they offer genuinely new information beyond what is captured by standard asset pricing models. Although these disclosure-based factors are derived from novel textual data, their return profiles could still be linear combinations of the traditional Fama-French six factors. In such a case, their explanatory power would be redundant, and any observed pricing would merely reflect repackaged exposures to known risks. To formally evaluate whether the HMN factors are spanned by the existing factor set, we employ the Gibbons, Ross, and Shanken (GRS) test (Gibbons, Ross, & Shanken, 1989), which provides a joint test of whether the intercepts (alphas) from a set of time-series regressions are all equal to zero. In this context, the test assesses whether the standard FF6 model can fully explain the returns of the HMN portfolios.

The test is implemented by regressing each HMN textual factor's return series on the six Fama-French factors. The null hypothesis is that the alphas from these regressions are jointly zero. A failure to reject this null would imply that our textual factors are redundant. Conversely, a rejection of the null would provide formal evidence that the FF6 model cannot fully explain the returns of the HMN portfolios, meaning our textual factors are not spanned by the canonical model and represent new, potentially orthogonal, dimensions of risk. The GRS test is implemented separately for each set of

[7] The correlations of the HMN factors across themselves and FF6 are provided in Figures 6-10. While some of the factors do have nominally large correlations with the standard Market factor, this is likely a side-effect of the long-short construction where the short leg contains a majority of the assets, which would bring its market exposure close to one, while the long leg is likely not perfectly offsetting it. Regardless, all subsequent empirical tests include the FF6 factors, thus controlling for these correlations.

HMN portfolios derived from a given clustering level k, allowing me to assess whether increasing topic granularity leads to richer, less redundant risk factors.

The results of the GRS test, presented in Table 17, provide a resounding rejection of the null hypothesis across all levels of k. For $k = 5$, the GRS statistic is 2.39 with a p-value of 0.0264. As the number of topics increases, the evidence against the null becomes even stronger. At the extreme, for $k = 25$, the GRS statistic rises to 3.72, with an extremely small p-value. This pattern strongly indicates that the textual risk factors are not redundant. They capture systematic variations in fund returns that are orthogonal to the risks captured by the standard model, justifying their use as new potential pricing factors in the subsequent analysis.

4.3 The Price of Disclosed Risk

Having established that the textual risk factors are not spanned by canonical asset pricing models, we now address the central question of their economic significance: are these disclosure-based risks priced in the cross-section of mutual fund returns? To test this, we implement a three-stage procedure following Fama and MacBeth (1973). In the first stage, we estimate time-series betas for each fund against both the standard $FF6$ factors and our newly constructed Have-minus-Not (HMN) textual factors, re-estimating these betas annually in line with the yearly frequency of prospectus updates. In the second stage, we run daily cross-sectional regressions of fund returns on the estimated betas to obtain a daily time series of slope coefficients for each factor. In the third and final stage, we compute the time-series average of these daily coefficients to arrive at the final risk premium estimate. The Fama-MacBeth slope coefficients, λ, shown in Tables 19 through 23 are these third-stage estimates, presented with Newey-West adjusted t-statistics.

We estimate three primary specifications to test for incremental pricing power: a benchmark "FF6 only" model, a "Textual only" model, and a "Full" model including

both sets of factors. The results reveal that disclosures contain novel information about the structure of expected returns, identifying a new and economically distinct class of systematic risk.

4.3.1 A New Class of Organizational Risk

The first-order result of this analysis is the significant and robust negative pricing of a broad class of what we call "organizational" risk factors, i.e. those arising from the fund's nature as a delegated asset manager, its structure, operations, and constraints, rather than the underlying assets. Funds disclosing these risks systematically underperform those that do not, even after controlling for standard factor exposures.

We identify a consistent set of these negatively priced organizational risks that survive the inclusion of FF6 controls. A standout among these is the "Redemption Risk" factor, which carries a consistently negative and highly significant estimate ranging from $\lambda = -0.018$ to $\lambda = -0.013$. Funds typically describe this risk as stemming from both the iliquidity of their investments and, uniquely to funds, unpredictable redemption pressures (see Appendix C). This suggests that funds which are acutely aware of, and exposed to, redemption risk, behave in a systematically different way than those that are not. The resulting investment strategy then yields lower expected returns. Similarly, the "Concentration Risk" factor is associated with a strong negative premium (e.g., $\lambda = -0.015$, $k = 10$, Table 20). Here, again, funds that are choosing or mandated to invest in only a small number of assets, or are concentrating in assets that are very similar in some regard, are seeing lower expected returns. This negative pricing extends to risks dealing with other aspects of the organization, such as "Management Risk", "Turnover Risk", or "Tracking Risk".

This consistent negative pricing of organizational risks is plausibly connected to the agency frictions inherent in delegated asset management, as suggested by Dou et al. (2020). In their framework, managerial compensation depends primarily on assets

under management rather than on pure performance, creating incentives to hedge against potential outflows even when such behavior lowers expected returns. The negative premium on the Redemption factor is consistent with this logic: funds that explicitly acknowledge redemption and liquidity constraints appear to adopt more defensive strategies that trade off some returns. However, this mechanism is unlikely to be the only explanation. The observed negative premia on organizational risks such as those related to management, concentration, or operations likely reflect a broader set of organizational motives and constraints within delegated portfolio management. These could include mandate rigidity, which confines managers to a prescribed investment universe and forces them to forgo higher-return opportunities; internal governance and risk controls, such as imposing ESG criteria that can constrain allocation decisions; and career incentives, which compel managers to avoid strategies with high long-term potential if they entail significant short-term volatility, since a sharp drawdown could cost them their jobs long before the returns ever materialize.

While organizational frictions are the primary discovery, the textual factors also provide insights into asset-based risks. Certain factors carry positive risk premia consistent with standard theory, serving as an important validation of our methodology. The "Growth" factor, for example, is consistently priced with a positive and significant premium (e.g., $\lambda = 0.020$, $t = 2.30$ in Table 20), indicating that investors demand compensation for exposure to the risks associated with growth-oriented strategies. More subtly, the survival of significant lambdas on textual factors that parallel canonical risks, such as "Market Risk" in the full model of Table 19, is also noteworthy. A significant lambda on a textual "Market Risk" or "Equity Risk" factor, controlling for the quantitative market beta, suggests that the manager's qualitative conception of risk is not perfectly correlated with its return-based counterpart. The results are able to provide context for non-equity assets as well, such as for "Derivatives Risk", the use of which has been studied by Kaniel and Wang (2025). Our findings of a consistent negative risk premium on this risk are in line with their observation that funds using derivatives tend to underperform. This textual representation appears to capture an

additional, incrementally priced dimension of risk, perhaps reflecting forward-looking concerns that are not fully spanned by historical return covariances.

4.3.2 Gains in Explanatory Power

The aggregate importance of these newly identified priced factors is confirmed by their substantial impact on model fit. While the economic story lies in the lambdas, the gains in explanatory power demonstrate that these factors describe a major, previously omitted dimension of the cross-section of fund returns.

As shown in Tables 19 through 23, the "Textual only" models consistently outperform the benchmark $FF6$ model. In the $k = 10$ specification (Table 20), for instance, the textual factors alone achieve an adjusted R^2 of 70%, far exceeding the 54% explained by the six canonical factors. This indicates that the practitioner's view of risk, as captured in disclosures, provides a more comprehensive picture of return variation.

The "Full" model, which integrates both perspectives, yields the most complete description. The adjusted R^2 rises steadily with topic granularity, from 70% at $k = 5$ up to 83% at $k = 25$ (Table 23). The improvement in explanatory power levels off as k increases, indicating that further topic refinement adds little new information and that we are close to the effective limit of distinct, informative risk dimensions. In sum, the analysis of risk premia identifies the specific economic content of disclosures, while the gains in explanatory power confirm their collective, first-order importance in understanding the cross-section of mutual fund returns.

5 Pricing the Cross-section of Stocks

The evidence thus far establishes that textual risk factors derived from mutual fund disclosures are priced within the universe of funds and provide significant explanatory power beyond canonical models. However, a crucial question remains: are these factors truly systematic? If the risks articulated by fund managers capture fundamental forces, they should not only explain variation in fund returns but should also be priced in the broader cross-section of individual stock returns. Testing the factors in this out-of-sample setting provides a powerful validation of their systematic nature.

To investigate this, we extend the analysis from the fund level to the individual stock level. The methodology mirrors the previous section, employing a two-stage Fama-MacBeth regression framework to ensure methodological consistency. In the first stage, for each stock in the CRSP universe, we estimate its time-series exposure (beta) to each of the textual HMN factors over rolling windows. In the second stage, we run daily cross-sectional regressions of individual stock returns on their estimated betas to obtain a time series of risk premia (λ) for each textual factor. The time-series average of these daily premia is then tested for statistical significance. A significant average premium indicates that the market systematically compensates investors for bearing the risk captured by that textual factor, confirming its status as a priced, systematic risk.

The results are presented in Tables 24 through 28. While the slope coefficients are more noisy than in the case of fund returns, an expected outcome given the higher idiosyncratic volatility of individual stocks relative to diversified fund portfolios, they again broadly support the hypothesis that these textual risk factors are priced and, indeed, systematic. The most noteworthy finding is that the "organizational" risks, which were a key discovery in the fund-level analysis, survive at the stock level as well.

Specifically, "Concentration" and "Redemption" risk emerge as consistently priced

factors. Among these two, "Concentration" risk stands out, carrying a significant negative risk premium across all specifications in which it appears ($k = 10$ through $k = 25$). The estimated premium ranges from -0.012 to -0.016, corresponding to an economically meaningful annualized risk premium of approximately -1.5%. The "Redemption" factor also demonstrates a consistent negative premium, with significant lambdas of -0.008 and -0.009 in the $k = 10$ and $k = 25$ specifications, respectively. While not statistically significant for $k = 15$ and $k = 20$, the t-statistics in those cases remain economically suggestive.

That "organizational" risks, are priced in individual stocks is a powerful and non-intuitive result. These are fundamentally fund-based risks, rooted in the operational realities and strategic limitations of delegated portfolio management. Their significance at the stock level suggests a powerful spillover mechanism through which fund-level frictions become systematic risk factors priced by the entire market.

To unpack the exact economic channels driving these risk premia, the following subsections dissect the anatomy of the two most prominent organizational frictions: Redemption Risk and Concentration Risk. By mapping these textual disclosures to the physical construction of the portfolios, their flow-performance sensitivities, and the institutional characteristics of the stocks they hold, we can definitively parse how managerial agency frictions transform into systematic asset pricing factors.

Beyond these organizational factors, other textual factors also show evidence of systematic pricing. The most consistent among these are "Foreign" risk, which appears with a statistically significant, negative premium, and "Growth" risk which comes out significantly positive across all specifications.

Looking at the final column, the analysis confirms that the disclosure-based factors capture novel information about the cross-section of stock returns. While the canonical Fama-French 6-factor model naturally explains a significant portion of return variation, our textual factors are not redundant. The key finding is that a model incorpo-

rating both sets of factors consistently outperforms the benchmark alone, and the risk premia on several textual factors remain significant. This provides strong evidence that our disclosure-based factors are not spanned by the FF6 factors and represent unique dimensions of systematic risk.

In summary, this analysis provides strong out-of-sample evidence for the validity of the disclosure-based risk factors. They are not merely priced within the mutual fund universe but also in the broader cross-section of stock returns. This demonstrates that the qualitative risk assessments of fund managers capture real, systematic economic risks, and it provides a novel, text-based perspective on how fund-level frictions can transmit to the wider market. The findings also open up fundamental questions about the modern asset pricing paradigm. First, they challenge the conventional, purely asset-based view of risk, suggesting that the institutional structure of the market and the constraints of its dominant investors may themselves constitute priced, systematic factors. Second, they invite inquiry into the mechanisms through which such institutional frictions become embedded in prices. Do these patterns reveal underlying asset characteristics that would have been priced regardless of who holds them, or do they instead arise precisely because of the structure of the market and the nature of its participants?

6 The Economic Mechanisms of Priced Organizational Risks

6.1 The Anatomy of Redemption Risk

To understand the economic drivers behind the negative risk premium associated with the text-based Redemption Risk factor, we examine the characteristics of the funds that explicitly disclose this risk. Panel A of Table 29 shows that, compared to their non-

disclosing peers, funds flagging redemption risk tend to be smaller ($1.5B vs $3.0B TNA) and experience both higher overall flows and significantly higher flow volatility (66.28% vs 11.42%). Operationally, these funds maintain larger cash buffers (5.44% vs 2.14%), hold slightly less in equities, and exhibit higher portfolio turnover (0.86 vs 0.47).

To verify whether this disclosure captures a tangible portfolio constraint rather than boilerplate language, we estimate a panel regression of the portfolio liquidity measure (L) developed by Ľuboš Pástor, Stambaugh, and Taylor (2020) on the disclosure dummy (Panel B, Model 1). The coefficient is negative and statistically significant ($\beta = -0.056$) indicating that disclosing funds hold less liquid portfolios, which could expose them to higher costs during periods of elevated redemptions.

As we have established before, bearing this risk also carries a performance cost. As shown in Model 2 of Panel B, funds disclosing redemption risk systematically underperform, generating lower returns even after controlling for standard Fama-French six-factor exposures ($\beta = -0.005, p < 0.01$). This implies that the disclosure captures a meaningful drag on performance that is not spanned by traditional risk factors. However, this cost may serve a structural purpose. Model 3 estimates the flow-performance sensitivity of these funds. The interaction term for redemption-disclosing funds is negative and significant ($\beta = -1.917$). This indicates that their flows and returns are somewhat decoupled. This ability to separate flows and performance aligns well with the theoretical framework of Dou et al. (2020), which posits that funds facing redemption pressures will take steps to hedge or manage their flow risk.

While the fund-level evidence hints towards a flow-hedging motive, examining the data at the stock level introduces an interesting empirical puzzle. If redemption-disclosing funds possess these unique constraints, how do their portfolio allocations impact the cross-section of equities? In Table 30, we sort stocks into quintiles based on the percentage of their shares held by these specific funds.

The sorting reveals that stocks heavily owned by redemption-disclosing funds tend to be value stocks (the Book-to-Market ratio increases from 0.689 in Q1 to 0.805 in Q5) and exhibit higher loadings on the SMB factor, though there is no strictly monotonic pattern in raw market capitalization. Furthermore, these stocks also feature notably high levels of aggregate mutual fund ownership (reaching 54.3% in Q5). Crucially, as a stock becomes more heavily owned by redemption-disclosing funds, its loading on our text-based Redemption Factor rises monotonically from -1.710 in Q1 to $+0.481$ in Q5, linking institutional ownership to the factor's pricing.

However, the characteristics of these stocks appear to depart from the predictions of the flow-hedging mechanism. In that framework, funds concerned with outflows should seek out stocks with lower flow betas, which the authors show are typically also liquid, large-cap, growth stocks. Our data shows a different pattern. First, the stocks preferred by redemption-disclosing funds lean toward less liquid value and smaller-cap profiles. Second, the flow betas across our ownership quintiles are relatively flat and actually move in the opposite direction of the theory: as ownership by redemption-weary funds increases, the stock's flow beta becomes *higher* (less negative), rising from -0.195 in Q1 to -0.126 in Q5.

Taken together, the evidence presents a puzzle. High aggregate mutual fund ownership, and ownership by redemption-disclosing funds in particular, is correlated with a higher loading on the redemption factor, which in turn predicts lower returns. Yet, this phenomenon is not fully explained by the flow-hedging asset selection mechanism, given the behavior of the flow betas and the fundamental characteristics of the stocks involved. This suggests that while mutual funds actively manage their flow-performance sensitivity at the fund level, the mechanism by which their specific organizational risks are priced into the cross-section of underlying stocks may rely on channels distinct from standard flow-beta hedging.

6.2 The Anatomy of Concentration Risk

While Redemption Risk disclosures appear to reflect a defensive response to investor outflows, the evidence for Concentration Risk points to a different mechanism. One natural interpretation is the convex flow-performance relationship in the mutual fund industry. As shown by Chevalier and Ellison (1997) and Sirri and Tufano (1998), funds are rewarded disproportionately for strong performance, while poor performance does not trigger an equally strong investor response. This asymmetry can create incentives for managers to increase portfolio variance in order to raise the probability of ending up in the top performance tier. In that sense, Concentration Risk disclosures may reflect funds explicitly acknowledging a mandate that gives managers greater scope to take concentrated bets.

Table 31 provides descriptive evidence consistent with this interpretation. Funds that disclose Concentration Risk are mechanically more concentrated, with a substantially higher average Herfindahl-Hirschman Index (0.131 versus 0.077). They also exhibit much greater return volatility (0.865 versus 0.165) and flow volatility (0.794 versus 0.162) than non-disclosing funds. In addition, these funds are somewhat younger on average (12.0 versus 13.4 years), and their managers oversee a larger number of funds (506 versus 467). Taken together, these patterns suggest that concentration-disclosing funds are organized in a way that is consistent with greater risk-taking and more dispersed performance outcomes across portfolios.

To examine whether these disclosures are associated with differences in investor rewards for performance, we estimate the flow-performance specification reported in Table 32. Rather than ranking funds on raw returns, we use lagged alpha so that the performance measure more closely captures deviations from the benchmark. This is especially relevant in the present setting, since concentrated portfolios are likely to generate idiosyncratic tracking error by construction.

The regression results point to an asymmetric payoff structure. The coefficient on

`Concentration` is positive and statistically significant ($\beta = 0.118$), indicating that concentration-disclosing funds receive higher baseline flows on average. More importantly, the interaction between Concentration and top-tercile alpha is positive and significant ($\beta = 0.085$), showing that these funds receive an additional flow benefit when performance is especially strong. By contrast, the interaction with bottom-tercile alpha is statistically indistinguishable from zero, providing little evidence that they face a corresponding increase in outflows when performance is poor.

Taken together, the evidence is consistent with the idea that Concentration Risk disclosures identify funds facing stronger incentives to generate high-variance outcomes. These funds are more concentrated in their holdings, display substantially higher return volatility, and appear to benefit disproportionately from strong alpha realizations without being symmetrically penalized for weak ones. This mechanism provides a plausible link between organizational structure at the fund level and the pricing of the underlying assets. If funds with these incentives direct capital toward positions that increase portfolio variance, their demand can affect the prices of the securities they hold, helping explain the negative risk premium associated with the Concentration Risk factor.

7 Conclusion

This paper studies whether mutual fund risk disclosures reveal economically important risks stemming from the structure of delegated asset management.Mutual funds are more than carriers of assets and the risks they have. They are intermediaries whose portfolio choices are shaped by redemptions, mandates, flow-performance incentives, career concerns, and operational constraints. If these considerations affect many funds at the same time, they may become relevant not only for fund returns, but also for the prices of the underlying assets.

The evidence supports this view. I first show that prospectus risk disclosures are informative rather than boilerplate. Disclosures align with established risk dimensions such as market, size, and value: funds disclosing market risk have higher market betas, while small-cap, large-cap, value, and growth disclosures correctly map both in direction and magnitude to the SMB and HML factors. Having shown credibility on the smaller subset of canonical factors, I then show that the disclosures reveal risks and constraints that standard return-based models do not directly measure. Many of the risks are associated with organizational risks reflecting agency frictions, including redemption pressure, concentration, management risk, turnover, tracking error, mandate restrictions, and operational complexity.

These organizational risks are priced. Disclosure-sorted return spreads are not spanned by conventional factor models and are priced in the cross-section of mutual fund returns. Extending the analysis to individual stocks shows that several of the same disclosure-based risks, especially redemption and concentration risk, are priced in the broader cross-section of equity returns as well. This result is the main contribution, that risks originating inside the asset management sector do not remain confined to fund performance, but can spill over into the pricing of the securities funds hold.

Taken together, the findings show that mutual fund disclosures provide a window into the risks and constraints of the intermediary itself. These risks are not lim-

ited to firm cash-flow risk or standard covariance with known factors. They arise from the institutional structure of asset management, how funds are organized, constrained, evaluated, and compensated. The paper therefore contributes to asset pricing by showing that the relevant state variables include not only firms and macroeconomic shocks, but also the incentives and constraints of the intermediaries that allocate capital. In this sense, the paper studies the corporate finance of asset management and its implications for asset pricing. Methodologically, it provides a framework for translating qualitative, forward-looking disclosure language into structured measures of intermediary-level risk.

References

Abis, S. (2020). Man vs. machine: Quantitative and discretionary equity management. *Machine: Quantitative and Discretionary Equity Management (October 23, 2020)*.

Abis, S., & Lines, A. (2021). *Text-based mutual fund peer groups* (Tech. Rep.). Working Paper, Columbia University.

Abis, S., & Lines, A. (2022). Do mutual funds keep their promises. *Unpublished Paper*.

Barth, D., Monin, P. J., Siriwardane, E., & Sunderam, A. (2024). Hidden risk. *Finance and Economics Discussion Series*(2024-098). Retrieved from https://doi.org/10.17016/FEDS.2024.098 doi: 10.17016/FEDS.2024.098

Basak, S., & Buffa, A. M. (2019). A theory of model sophistication and operational risk. *Available at SSRN: https://ssrn.com/abstract=2737178 or http://dx.doi.org/10.2139/ssrn.2737178*.

Basak, S., & Pavlova, A. (2013, August). Asset prices and institutional investors. *American Economic Review*, *103*(5), 1728–58. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.103.5.1728 doi: 10.1257/aer.103.5.1728

Bellstam, G., Bhagat, S., & Cookson, J. A. (2021). A text-based analysis of corporate innovation. *Management Science*, *67*(7), 4004–4031.

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of political economy*, *112*(6), 1269–1295.

Brunnermeier, M. K., & Sannikov, Y. (2014, February). A macroeconomic model with a financial sector. *American Economic Review*, *104*(2), 379–421. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.104.2.379 doi: 10.1257/aer.104.2.379

Buffa, A. M., Vayanos, D., & Woolley, P. (2022). Asset management contracts and equilibrium prices. *Journal of Political Economy*, *130*(12), 3146-3201. Retrieved from https://doi.org/10.1086/720515 doi: 10.1086/720515

Bybee, L., Kelly, B., & Su, Y. (2023). Narrative asset pricing: Interpretable systematic risk factors from news text. *The Review of Financial Studies*, *36*(12), 4759–4787.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H., & Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, *19*(1), 396–455. doi: 10.1007/s11142-013-9258-3

Cen, X., Dou, W. W., Kogan, L., & Wu, W. (2023, Dec). *Fund flows and income risk of fund managers* (NBER Working Papers No. 31986). National Bureau of Economic Research, Inc. Retrieved from https://ideas.repec.org/p/nbr/nberwo/31986.html doi: None

Chevalier, J., & Ellison, G. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy*, *105*(6), 1167-1200. Retrieved from https://doi.org/10.1086/516389 doi: 10.1086/516389

Cuoco, D., & Kaniel, R. (2011). Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics*, *101*(2), 264-296. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304405X11000456 doi: https://doi.org/10.1016/j.jfineco.2011.02.012

Dou, W. W., Kogan, L., & Wu, W. (2020). Common fund flows: Flow hedging and factor pricing. *SSRN Electronic Journal*. Retrieved from https://doi.org/10.2139/ssrn.3543675 doi: 10.2139/ssrn.3543675

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56. Retrieved from https://www.sciencedirect.com/science/article/pii/0304405X93900235 doi: https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1-22. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304405X14002323 doi: https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636. Retrieved 2025-07-20, from http://www.jstor.org/stable/1831028

Gabaix, X., Koijen, R. S., Richmond, R., & Yogo, M. (2025). Asset embeddings. *NBER Working Paper 33651*. Retrieved from https://doi.org/10.3386/w33651 doi: 10.3386/w33651

Gaulin, M. (2017). Risk fact or fiction: The information content of risk factor disclosures. *Working Paper*. Retrieved from https://repository.rice.edu/handle/1911/96328

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, *57*(5), 1121–1152. Retrieved 2025-07-19, from http://www.jstor.org/stable/1913625

Haddad, V., & Muir, T. (2025). *Intermediaries and asset prices* (Tech. Rep. No. 34146). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w34146

He, Z., & Krishnamurthy, A. (2013, April). Intermediary asset pricing. *American Economic Review*, *103*(2), 732–70. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.103.2.732 doi: 10.1257/aer.103.2.732

Hwang, B.-H., & Kim, H. H. (2017). It pays to write well. *Journal of Financial Economics*, *124*(2), 373–394.

Ibert, M., Kaniel, R., Van Nieuwerburgh, S., & Vestman, R. (2017, 09). Are mutual fund managers paid for investment skill? *The Review of Financial Studies*, *31*(2), 715-772. Retrieved from https://doi.org/10.1093/rfs/hhx105 doi: 10.1093/rfs/hhx105

Kaniel, R., Lin, Z., Pelger, M., & Van Nieuwerburgh, S. (2023). Machine-learning the skill of mutual fund managers. *Journal of Financial Economics*, *150*(1), 94-138. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304405X23001253 doi: https://doi.org/10.1016/j.jfineco.2023.07.004

Kaniel, R., & Wang, P. (2025, 01). Unmasking mutual fund derivative use. *The Review of Financial Studies*, *38*(4), 1120-1166. Retrieved from https://doi.org/10.1093/rfs/hhaf001 doi: 10.1093/rfs/hhaf001

Ke, Z. T., Kelly, B. T., & Xiu, D. (2019). *Predicting returns with text data* (Tech. Rep.). National Bureau of Economic Research.

Koijen, R. S. J., & Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, *127*(4), 1475-1515. Retrieved from https://doi.org/10.1086/701683 doi: 10.1086/701683

Kostovetsky, L., & Warner, J. B. (2020). Measuring innovation and product differentiation: Evidence from mutual funds. *The Journal of Finance*, *75*(2), 779–823.

Krakow, J., & Schäfer, T. (2020). *Mutual funds and risk disclosure: Information content of fund prospectuses* (Tech. Rep.). Swiss Finance Institute.

Lopez-Lira, A. (2021). Why do managers disclose risks accurately? textual analysis, disclosures, and risk exposures. *Economics Letters*, *204*, 109896. Retrieved from https://www.sciencedirect.com/science/article/pii/S0165176521001737 doi: https://doi.org/10.1016/j.econlet.2021.109896

Lopez-Lira, A. (2023). Risk factors that matter: Textual analysis of risk disclosures for the cross-section of returns. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Pavlova, A., & Sikorskaya, T. (2022, 08). Benchmarking intensity. *The Review of Financial Studies*, *36*(3), 859-903. Retrieved from https://doi.org/10.1093/rfs/hhac055 doi: 10.1093/rfs/hhac055

Rajan, R. G. (2006). Has finance made the world riskier? *European Financial Management*, *12*(4), 499-533. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-036X.2006.00330.x doi: https://doi.org/10.1111/j.1468-036X.2006.00330.x

Sheng, J., Xu, N., & Zheng, L. (2023). Do mutual funds walk the talk? evidence from fund risk disclosure. *Evidence from Fund Risk Disclosure (October 17, 2023)*.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The Journal of Finance*, *53*(5), 1589-1622. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/0022-1082.00066 doi: https://doi.org/

10.1111/0022-1082.00066

Spindler, J. (2006). Is it time to wind up the securities act of 1933? *Regulation*, *29*(4), 48–54. (Winter 2006–2007 issue)

Ľuboš Pástor, Stambaugh, R. F., & Taylor, L. A. (2020). Fund tradeoffs. *Journal of Financial Economics*, *138*(3), 614-634. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304405X20301720 doi: https://doi.org/10.1016/j.jfineco.2020.06.005

A Figures

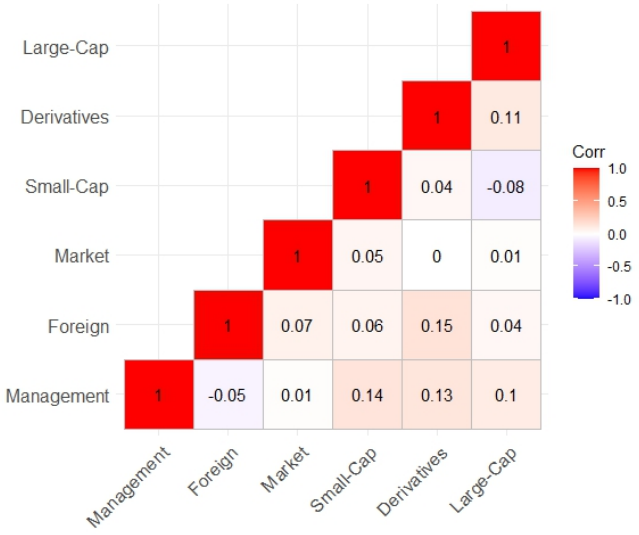


Figure 1: Correlation Matrix for k=5 Risk Topic Mentions

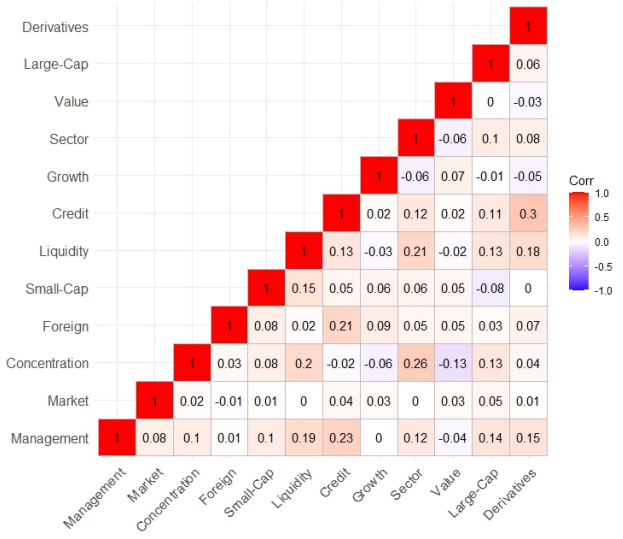


Figure 2: Correlation Matrix for k=10 Risk Topic Mentions

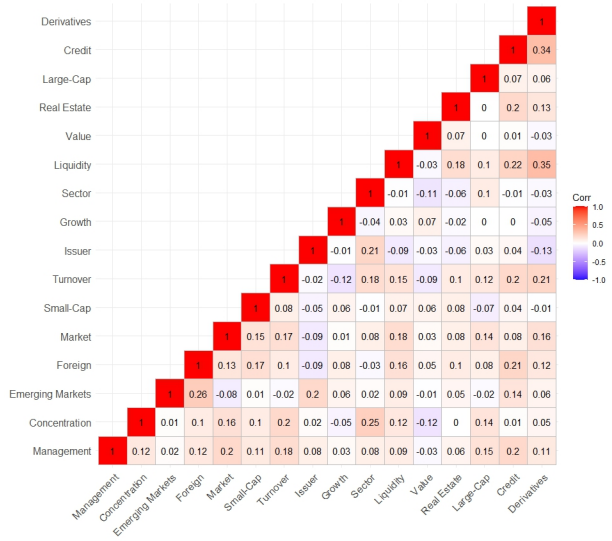


Figure 3: Correlation Matrix for k=15 Risk Topic Mentions

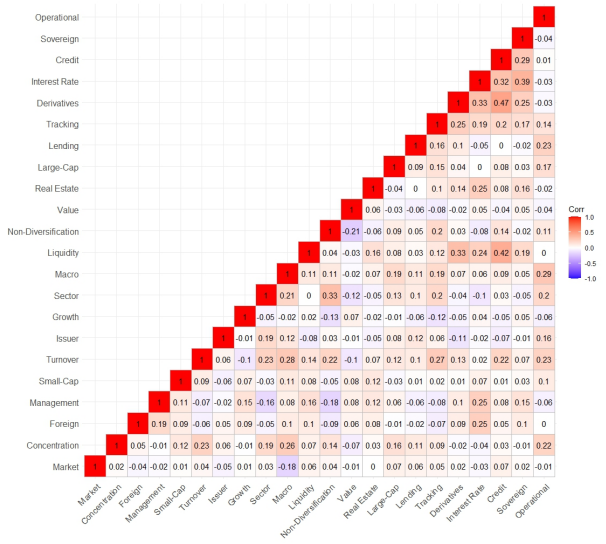


Figure 4: Correlation Matrix for k=20 Risk Topic Mentions

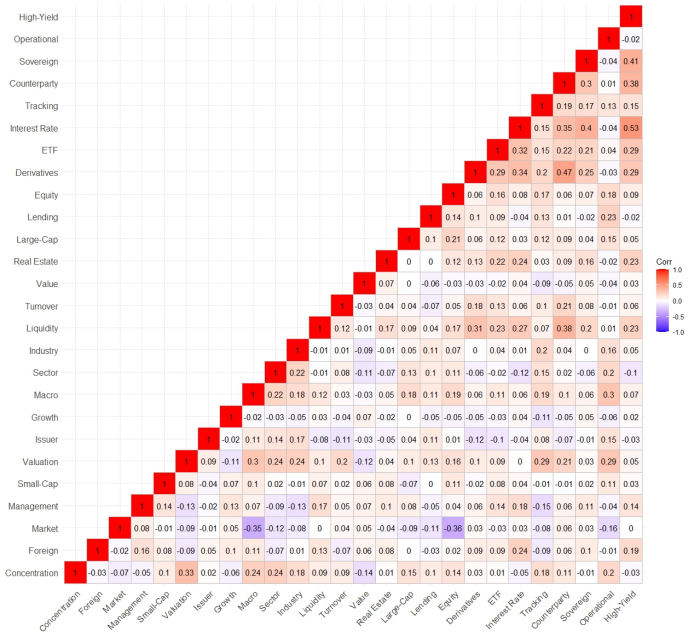


Figure 5: Correlation Matrix for k=25 Risk Topic Mentions

B Tables

B.1 Risk Disclosure Summary Statistics

Table 1: Risk Disclosure Summary Statistics

Year	Mean Risks per Fund	Mean Description Length	No. of Funds	Unique Risk Descriptions
2011	8.65	455.1	4,378	11,019
2012	9.40	478.7	5,331	15,082
2013	9.95	490.6	5,646	17,006
2014	10.55	510.5	6,038	18,200
2015	11.25	526.5	6,428	20,940
2016	12.31	546.9	6,935	23,880
2017	13.11	562.3	7,015	24,085
2018	13.52	581.0	7,191	24,239
2019	14.16	596.7	6,368	22,793
2020	14.63	627.1	7,044	24,334
2021	15.01	662.9	6,918	24,482
2022	15.57	684.5	6,996	25,797
2023	16.07	706.6	7,177	29,728
All Years	**12.89**	**579.4**	**83,465**	**130,730**

Notes: This table reports annual summary statistics of extracted risk disclosures from U.S. mutual fund prospectuses between 2011 and 2023. "Mean Risks per Fund" is the average number of distinct risk items disclosed per fund in each year. "Mean Description Length" measures the average number of characters in each risk description. "No. of Funds" is the total number of funds with available prospectus data in that year, and "Unique Risk Descriptions" counts distinct textual risk statements across all funds. The bottom row ("All Years, pooled") reports statistics computed across the full panel of fund–year observations rather than time-averaged yearly means.

B.2 Beta Regression Tables

Table 2: Regression of MKT Factor Betas on Disclosed Risk Topics (k=5)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Market Risk	0.488*** (0.008)	1.270*** (0.010)
Small-Cap Risk	0.033*** (0.003)	1.180*** (0.018)
Large-Cap Risk	-0.003 (0.004)	1.310*** (0.035)
Foreign Risk	-0.065*** (0.004)	0.676*** (0.008)
Management Risk	-0.111*** (0.007)	0.667*** (0.005)
Derivatives Risk	-0.127*** (0.004)	0.649*** (0.007)
Observations	83,278	
Adjusted R^2	0.067	0.756

OLS coefficients for regressions of MKT betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 3: Regression of SMB Factor Betas on Disclosed Risk Topics (k=5)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Small-Cap Risk	0.204*** (0.002)	1.250*** (0.012)
Market Risk	0.036*** (0.005)	0.149*** (0.006)
Management Risk	-0.032*** (0.005)	0.112*** (0.003)
Derivatives Risk	-0.036*** (0.002)	0.009 (0.005)
Large-Cap Risk	-0.083*** (0.002)	-0.459*** (0.022)
Foreign Risk	-0.159*** (0.002)	-0.257*** (0.005)
Observations	83,278	
Adjusted R^2	0.151	0.221

OLS coefficients for regressions of SMB betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 4: Regression of HML Factor Betas on Disclosed Risk Topics (k=5)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Small-Cap Risk	0.021*** (0.002)	0.026* (0.012)
Derivatives Risk	0.020*** (0.002)	0.030*** (0.005)
Management Risk	0.010* (0.005)	0.050*** (0.003)
Large-Cap Risk	0.006* (0.002)	-0.077*** (0.023)
Market Risk	-0.012* (0.005)	-0.014* (0.006)
Foreign Risk	-0.030*** (0.003)	-0.023*** (0.005)
Observations	83,278	
Adjusted R^2	0.004	0.006

OLS coefficients for regressions of HML betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 5: Regression of MKT Factor Betas on Disclosed Risk Topics (k=10)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Growth Risk	0.082*** (0.004)	1.490*** (0.034)
Market Risk	0.046*** (0.007)	1.160*** (0.014)
Value Risk	0.046*** (0.003)	1.240*** (0.027)
Small-Cap Risk	0.046*** (0.003)	1.120*** (0.018)
Large-Cap Risk	0.011** (0.004)	1.250*** (0.038)
Sector Risk	0.006 (0.003)	1.030*** (0.016)
Credit Risk	-0.038*** (0.004)	0.480*** (0.013)
Concentration Risk	-0.051*** (0.003)	0.858*** (0.020)
Redemption Risk	-0.057*** (0.004)	0.517*** (0.008)
Foreign Risk	-0.064*** (0.004)	0.668*** (0.009)
Management Risk	-0.089*** (0.005)	0.896*** (0.012)
Derivatives Risk	-0.121*** (0.003)	0.598*** (0.014)
Observations	83,278	
Adjusted R^2	0.056	0.759

OLS coefficients for regressions of MKT betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 6: Regression of SMB Factor Betas on Disclosed Risk Topics (k=10)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Small-Cap Risk	0.200*** (0.002)	1.260*** (0.012)
Sector Risk	0.040*** (0.002)	0.345*** (0.010)
Value Risk	0.038*** (0.002)	0.264*** (0.018)
Growth Risk	0.020*** (0.002)	0.203*** (0.022)
Management Risk	0.006 (0.003)	0.078*** (0.008)
Redemption Risk	0.001 (0.002)	0.122*** (0.005)
Market Risk	-0.017*** (0.004)	0.141*** (0.009)
Concentration Risk	-0.025*** (0.002)	0.004 (0.013)
Credit Risk	-0.030*** (0.002)	-0.065*** (0.008)
Derivatives Risk	-0.052*** (0.002)	-0.049*** (0.009)
Large-Cap Risk	-0.090*** (0.002)	-0.528*** (0.024)
Foreign Risk	-0.167*** (0.002)	-0.313*** (0.006)
Observations	83,278	
Adjusted R^2	0.174	0.236

OLS coefficients for regressions of SMB betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 7: Regression of HML Factor Betas on Disclosed Risk Topics (k=10)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Value Risk	0.118*** (0.002)	0.617*** (0.018)
Management Risk	0.027*** (0.003)	0.065*** (0.008)
Redemption Risk	0.022*** (0.002)	0.094*** (0.005)
Small-Cap Risk	0.018*** (0.002)	0.042*** (0.012)
Sector Risk	0.010*** (0.002)	0.001 (0.010)
Large-Cap Risk	0.006* (0.002)	-0.008 (0.024)
Concentration Risk	0.003 (0.002)	-0.049*** (0.013)
Credit Risk	0.003 (0.002)	-0.020* (0.008)
Derivatives Risk	-0.004 (0.002)	0.001 (0.009)
Market Risk	-0.022*** (0.004)	0.014 (0.009)
Foreign Risk	-0.032*** (0.002)	-0.043*** (0.006)
Growth Risk	-0.128*** (0.002)	-0.947*** (0.022)
Observations	83,278	
Adjusted R^2	0.067	0.046

OLS coefficients for regressions of HML betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 8: Regression of MKT Factor Betas on Disclosed Risk Topics (k=15)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Equity Risk	0.114*** (0.004)	1.360*** (0.021)
Growth Risk	0.084*** (0.004)	1.360*** (0.034)
Value Risk	0.040*** (0.003)	1.200*** (0.027)
Small-Cap Risk	0.038*** (0.003)	1.070*** (0.019)
Market Risk	0.031*** (0.004)	1.150*** (0.015)
Real Estate Risk	0.004 (0.004)	0.851*** (0.027)
Sector Risk	0.002 (0.004)	1.130*** (0.019)
Issuer Risk	-0.006 (0.004)	0.901*** (0.027)
Large-Cap Risk	-0.012** (0.004)	1.140*** (0.038)
Foreign Risk	-0.028*** (0.004)	0.705*** (0.012)
Redemption Risk	-0.035*** (0.004)	0.843*** (0.039)
Management Risk	-0.040*** (0.004)	0.848*** (0.016)
Concentration Risk	-0.054*** (0.003)	0.826*** (0.022)
Credit Risk	-0.055*** (0.004)	0.439*** (0.014)
Turnover Risk	-0.080*** (0.004)	0.456*** (0.008)
Emerging Markets Risk	-0.089*** (0.004)	0.497*** (0.019)
Derivatives Risk	-0.097*** (0.004)	0.639*** (0.014)
Observations	83,278	
Adjusted R^2	0.072	0.762

OLS coefficients for regressions of MKT betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 9: Regression of SMB Factor Betas on Disclosed Risk Topics (k=15)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Small-Cap Risk	0.205*** (0.002)	1.250*** (0.012)
Real Estate Risk	0.046*** (0.003)	0.350*** (0.017)
Value Risk	0.032*** (0.002)	0.241*** (0.018)
Equity Risk	0.032*** (0.002)	0.230*** (0.014)
Sector Risk	0.030*** (0.002)	0.340*** (0.012)
Redemption Risk	0.027*** (0.002)	0.544*** (0.025)
Management Risk	0.018*** (0.003)	0.044*** (0.010)
Growth Risk	0.018*** (0.002)	0.166*** (0.022)
Market Risk	0.010*** (0.003)	0.176*** (0.010)
Issuer Risk	-0.002 (0.003)	0.124*** (0.017)
Turnover Risk	-0.006* (0.002)	0.071*** (0.005)
Concentration Risk	-0.024*** (0.002)	-0.039** (0.014)
Credit Risk	-0.038*** (0.002)	-0.102*** (0.009)
Derivatives Risk	-0.045*** (0.002)	-0.045*** (0.009)
Large-Cap Risk	-0.097*** (0.002)	-0.567*** (0.025)
Emerging Markets Risk	-0.100*** (0.002)	-0.451*** (0.012)
Foreign Risk	-0.128*** (0.002)	-0.297*** (0.008)
Observations	83,278	
Adjusted R^2	0.195	0.244

OLS coefficients for regressions of SMB betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 10: Regression of HML Factor Betas on Disclosed Risk Topics (k=15)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Value Risk	0.120*** (0.002)	0.596*** (0.018)
Real Estate Risk	0.023*** (0.003)	0.086*** (0.018)
Management Risk	0.022*** (0.003)	0.026* (0.010)
Small-Cap Risk	0.020*** (0.002)	0.019 (0.012)
Turnover Risk	0.014*** (0.002)	0.079*** (0.005)
Market Risk	0.014*** (0.003)	0.106*** (0.010)
Equity Risk	0.013*** (0.002)	0.073*** (0.014)
Concentration Risk	0.009*** (0.002)	0.002 (0.014)
Sector Risk	0.006** (0.002)	-0.070*** (0.013)
Issuer Risk	0.006* (0.003)	-0.014 (0.017)
Large-Cap Risk	0.004 (0.002)	-0.048 (0.025)
Credit Risk	0.002 (0.002)	-0.017 (0.009)
Derivatives Risk	-0.002 (0.002)	0.005 (0.009)
Redemption Risk	-0.003 (0.002)	-0.032 (0.025)
Emerging Markets Risk	-0.007** (0.002)	-0.082*** (0.013)
Foreign Risk	-0.034*** (0.002)	-0.040*** (0.008)
Growth Risk	-0.128*** (0.002)	-0.963*** (0.022)
Observations	83,278	
Adjusted R^2	0.069	0.047

OLS coefficients for regressions of HML betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 11: Regression of MKT Factor Betas on Disclosed Risk Topics (k=20)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Market Risk	0.079*** (0.005)	1.210*** (0.017)
Growth Risk	0.063*** (0.004)	1.360*** (0.034)
Small-Cap Risk	0.038*** (0.003)	1.090*** (0.018)
Sector Risk	0.036*** (0.004)	1.130*** (0.021)
Value Risk	0.036*** (0.004)	1.190*** (0.028)
Real Estate Risk	0.021*** (0.004)	0.888*** (0.027)
Management Risk	0.020*** (0.004)	0.945*** (0.024)
Concentration Risk	0.013*** (0.003)	1.010*** (0.027)
Lending Risk	0.011* (0.005)	1.180*** (0.049)
Operational Risk	0.001 (0.005)	1.060*** (0.044)
Large-Cap Risk	-0.008 (0.004)	1.170*** (0.048)
Issuer Risk	-0.010* (0.004)	0.824*** (0.029)
Macro Risk	-0.017*** (0.004)	0.739*** (0.015)
Redemption Risk	-0.020*** (0.004)	0.886*** (0.039)
Interest Rate Risk	-0.054*** (0.004)	0.481*** (0.017)
Turnover Risk	-0.063*** (0.004)	0.478*** (0.015)
Derivatives Risk	-0.064*** (0.004)	0.552*** (0.015)
Tracking Risk	-0.065*** (0.004)	0.555*** (0.011)
Credit Risk	-0.079*** (0.005)	0.498*** (0.054)
Foreign Risk	-0.079*** (0.004)	0.702*** (0.010)
Non-Diversification Risk	-0.099*** (0.004)	0.787*** (0.032)
Sovereign Risk	-0.102*** (0.007)	-0.373*** (0.102)
Observations	83,278	
Adjusted R^2	0.085	0.760

OLS coefficients for regressions of MKT betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 12: Regression of SMB Factor Betas on Disclosed Risk Topics (k=20)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Small-Cap Risk	0.202*** (0.002)	1.200*** (0.012)
Sector Risk	0.049*** (0.003)	0.384*** (0.013)
Real Estate Risk	0.035*** (0.003)	0.376*** (0.018)
Management Risk	0.028*** (0.002)	0.173*** (0.016)
Value Risk	0.024*** (0.002)	0.257*** (0.018)
Redemption Risk	0.020*** (0.003)	0.542*** (0.025)
Operational Risk	0.011*** (0.003)	0.359*** (0.028)
Growth Risk	0.008*** (0.003)	0.261*** (0.022)
Issuer Risk	0.006* (0.003)	0.122*** (0.018)
Turnover Risk	0.002 (0.002)	0.114*** (0.009)
Concentration Risk	-0.002 (0.002)	0.190*** (0.017)
Sovereign Risk	-0.008 (0.004)	-0.169** (0.065)
Credit Risk	-0.011*** (0.003)	-0.020 (0.035)
Macro Risk	-0.012*** (0.002)	0.071*** (0.010)
Market Risk	-0.013*** (0.003)	0.081*** (0.011)
Lending Risk	-0.017*** (0.003)	-0.043 (0.031)
Tracking Risk	-0.020*** (0.002)	0.009 (0.007)
Interest Rate Risk	-0.043*** (0.003)	-0.106*** (0.011)
Derivatives Risk	-0.043*** (0.003)	-0.047*** (0.010)
Non-Diversification Risk	-0.048*** (0.003)	-0.154*** (0.021)
Large-Cap Risk	-0.105*** (0.003)	-1.050*** (0.031)
Foreign Risk	-0.181*** (0.002)	-0.353*** (0.007)
Observations	83,278	
Adjusted R^2	0.188	0.241

OLS coefficients for regressions of SMB betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.

***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 13: Regression of HML Factor Betas on Disclosed Risk Topics (k=20)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Value Risk	0.113*** (0.002)	0.622*** (0.018)
Operational Risk	0.045*** (0.003)	0.386*** (0.028)
Macro Risk	0.033*** (0.002)	0.187*** (0.010)
Real Estate Risk	0.023*** (0.003)	0.111*** (0.018)
Concentration Risk	0.014*** (0.002)	0.110*** (0.017)
Small-Cap Risk	0.010*** (0.002)	0.034** (0.012)
Tracking Risk	0.010*** (0.002)	0.049*** (0.007)
Credit Risk	0.005 (0.003)	-0.021 (0.035)
Management Risk	0.003 (0.002)	0.026 (0.016)
Large-Cap Risk	0.003 (0.003)	0.029 (0.031)
Issuer Risk	0.002 (0.003)	0.017 (0.018)
Sector Risk	0.002 (0.003)	-0.109*** (0.013)
Derivatives Risk	0.001 (0.003)	0.030** (0.010)
Redemption Risk	-0.002 (0.002)	-0.032 (0.025)
Turnover Risk	-0.004 (0.002)	0.037*** (0.009)
Interest Rate Risk	-0.004 (0.003)	-0.020 (0.011)
Lending Risk	-0.006* (0.003)	-0.052 (0.031)
Market Risk	-0.007* (0.003)	-0.012 (0.011)
Sovereign Risk	-0.009* (0.004)	-0.034 (0.066)
Non-Diversification Risk	-0.035*** (0.003)	-0.151*** (0.021)
Foreign Risk	-0.041*** (0.002)	-0.081*** (0.007)
Growth Risk	-0.129*** (0.002)	-0.907*** (0.022)
Observations	83,278	
Adjusted R^2	0.076	0.055

OLS coefficients for regressions of HML betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.

***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 14: Regression of MKT Factor Betas on Disclosed Risk Topics (k=25)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Equity Risk	0.100*** (0.004)	1.380*** (0.024)
Growth Risk	0.075*** (0.004)	1.350*** (0.035)
Value Risk	0.046*** (0.003)	1.200*** (0.028)
Sector Risk	0.029*** (0.004)	1.200*** (0.026)
Small-Cap Risk	0.025*** (0.003)	1.060*** (0.019)
Management Risk	0.015*** (0.003)	0.967*** (0.028)
Redemption Risk	0.015*** (0.004)	0.989*** (0.040)
Real Estate Risk	0.008 (0.004)	0.853*** (0.027)
ETF Risk	0.004 (0.004)	0.771*** (0.035)
Issuer Risk	-0.002 (0.004)	0.943*** (0.026)
Lending Risk	-0.002 (0.005)	1.160*** (0.048)
Operational Risk	-0.005 (0.005)	1.090*** (0.044)
Industry Risk	-0.007 (0.005)	1.000*** (0.034)
Market Risk	-0.009* (0.004)	1.080*** (0.020)
Large-Cap Risk	-0.015*** (0.004)	1.180*** (0.038)
High-Yield Risk	-0.024*** (0.006)	0.266*** (0.059)
Macro Risk	-0.028*** (0.004)	0.675*** (0.015)
Concentration Risk	-0.036*** (0.004)	0.899*** (0.023)
Turnover Risk	-0.045*** (0.004)	0.860*** (0.046)
Interest Rate Risk	-0.055*** (0.004)	0.521*** (0.020)
Derivatives Risk	-0.063*** (0.004)	0.528*** (0.015)
Valuation Risk	-0.072*** (0.004)	0.457*** (0.016)
Tracking Risk	-0.078*** (0.004)	0.527*** (0.011)
Counterparty Risk	-0.078*** (0.005)	0.473*** (0.054)
Foreign Risk	-0.083*** (0.004)	0.692*** (0.010)
Sovereign Risk	-0.096*** (0.007)	-0.379*** (0.101)
Observations	83,278	
Adjusted R^2	0.089	0.761

OLS coefficients for regressions of MKT betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 15: Regression of SMB Factor Betas on Disclosed Risk Topics (k=25)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Small-Cap Risk	0.194*** (0.002)	1.260*** (0.012)
Real Estate Risk	0.045*** (0.003)	0.360*** (0.018)
Value Risk	0.038*** (0.002)	0.284*** (0.018)
Sector Risk	0.034*** (0.003)	0.385*** (0.017)
Redemption Risk	0.027*** (0.003)	0.555*** (0.025)
Industry Risk	0.026*** (0.003)	0.233*** (0.022)
Turnover Risk	0.025*** (0.003)	0.346*** (0.030)
Equity Risk	0.023*** (0.002)	0.281*** (0.015)
Growth Risk	0.014*** (0.003)	0.170*** (0.022)
ETF Risk	0.014*** (0.003)	0.212*** (0.023)
Management Risk	0.014*** (0.002)	0.130*** (0.018)
Operational Risk	0.008* (0.003)	0.326*** (0.029)
Issuer Risk	0.005 (0.003)	0.162*** (0.017)
Sovereign Risk	-0.008 (0.005)	-0.198** (0.065)
Counterparty Risk	-0.008** (0.003)	-0.023 (0.035)
High-Yield Risk	-0.012** (0.004)	-0.177*** (0.038)
Lending Risk	-0.013*** (0.003)	0.005 (0.031)
Market Risk	-0.015*** (0.002)	0.013 (0.013)
Macro Risk	-0.021*** (0.003)	0.008 (0.010)
Valuation Risk	-0.021*** (0.003)	0.077*** (0.010)
Concentration Risk	-0.023*** (0.002)	0.022 (0.015)
Tracking Risk	-0.028*** (0.002)	-0.009 (0.007)
Interest Rate Risk	-0.038*** (0.003)	-0.076*** (0.013)
Derivatives Risk	-0.043*** (0.003)	-0.059*** (0.010)
Large-Cap Risk	-0.099*** (0.002)	-0.487*** (0.024)
Foreign Risk	-0.174*** (0.002)	-0.356*** (0.007)
Observations	83,278	
Adjusted R^2	0.187	0.242

OLS coefficients for regressions of SMB betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

Table 16: Regression of HML Factor Betas on Disclosed Risk Topics (k=25)

	Disclosure Mention (Dummy)	**Disclosure Intensity (Proportion)**
Risk Topic	Coefficient	Coefficient
Value Risk	0.119*** (0.002)	0.671*** (0.018)
Industry Risk	0.046*** (0.003)	0.347*** (0.022)
Operational Risk	0.044*** (0.003)	0.387*** (0.029)
Macro Risk	0.034*** (0.002)	0.174*** (0.010)
Real Estate Risk	0.025*** (0.003)	0.095*** (0.018)
Small-Cap Risk	0.015*** (0.002)	0.056*** (0.012)
Tracking Risk	0.008*** (0.002)	0.044*** (0.007)
Management Risk	0.008*** (0.002)	0.049** (0.018)
Counterparty Risk	0.007* (0.003)	-0.002 (0.034)
Valuation Risk	0.006* (0.002)	0.064*** (0.010)
Equity Risk	0.005* (0.002)	0.094*** (0.015)
ETF Risk	0.003 (0.003)	0.065** (0.023)
Derivatives Risk	0.002 (0.003)	0.032** (0.010)
Large-Cap Risk	0.001 (0.002)	-0.008 (0.024)
Interest Rate Risk	0.000 (0.003)	-0.029* (0.013)
Issuer Risk	-0.000 (0.003)	-0.030 (0.017)
Concentration Risk	-0.002 (0.002)	-0.026 (0.015)
Redemption Risk	-0.002 (0.002)	-0.016 (0.025)
Sovereign Risk	-0.006 (0.005)	-0.034 (0.065)
Market Risk	-0.006** (0.002)	-0.032* (0.013)
Lending Risk	-0.008* (0.003)	-0.034 (0.031)
Turnover Risk	-0.019*** (0.003)	-0.186*** (0.030)
High-Yield Risk	-0.022*** (0.004)	-0.001 (0.038)
Foreign Risk	-0.042*** (0.002)	-0.078*** (0.007)
Sector Risk	-0.046*** (0.003)	-0.379*** (0.017)
Growth Risk	-0.124*** (0.002)	-0.896*** (0.022)
Observations	83,278	
Adjusted R^2	0.080	0.063

OLS coefficients for regressions of HML betas on disclosure mention (dummy) and intensity (proportion). Topics are sorted by the mention (dummy) coefficient. Standard errors are in parentheses.
***, **, and * denote significance at the 0.1%, 1%, and 5% levels, respectively.

B.3 GRS

Table 17: GRS Test of Textual HMN Factors Using Fama–French 6 Factors (2011–2023)

HMN Factor Set	GRS Statistic	p-value
$\text{HMN}_{k=5}$	2.39	0.0264
$\text{HMN}_{k=10}$	2.79	0.0008
$\text{HMN}_{k=15}$	3.30	5.3×10^{-6}
$\text{HMN}_{k=20}$	3.40	1.4×10^{-7}
$\text{HMN}_{k=25}$	3.18	1.1×10^{-7}

Notes: HMN_k denotes the textual "Have-Minus-NotHave" factor set constructed from k risk topics obtained via k-means clustering of mutual-fund risk disclosures. The GRS test examines whether the intercepts from regressing the Fama–French six factors (MKT, SMB, HML, RMW, CMA, MOM) on each HMN factor set are jointly zero. Significant rejections imply that HMN factors contain systematic information not spanned by the six standard factors.

Table 18: GRS Test of Textual HMN Factors Using JKP Daily Anomaly Returns (2011–2023)

HMN Factor Set	GRS Statistic	p-value
$\text{HMN}_{k=5}$	7.13	1.45×10^{-7}
$\text{HMN}_{k=10}$	4.44	4.32×10^{-7}
$\text{HMN}_{k=15}$	4.77	3.31×10^{-10}
$\text{HMN}_{k=20}$	4.34	6.68×10^{-11}
$\text{HMN}_{k=25}$	4.31	2.25×10^{-12}

Notes: Each row corresponds to the set of textual "Have-Minus-NotHave" (HMN) risk factors derived from k-means topic modeling with $k \in \{5, 10, 15, 20, 25\}$. The Gibbons–Ross–Shanken (1989) test evaluates the null hypothesis that the alphas from regressing the JKP daily anomaly returns on each HMN factor set are jointly zero. Rejecting the null indicates that the HMN textual factors contain return-relevant information not spanned by the JKP anomalies.

B.4 Fund Fama-MacBeth Specifications

Table 19: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Fund Returns (k=5)

	Fama–French 6 (λ)						*Textual Risk Factors (λ)*						R^2
	Mkt	SMB	HML	RMW	CMA	Mom	Mgmt	Market	SmCap	LgCap	Forgn	Deriv	
FF6 only	0.118	-0.008	-0.042	0.049	-0.002	0.186							0.54
	[2.84]**	[-0.30]	[-1.09]	[1.77]	[-0.08]	[3.45]***							
Textual only							-0.014	0.050	0.000	0.006	-0.020	-0.020	0.59
							[-2.57]*	[1.96]*	[0.05]	[1.23]	[-2.23]*	[-2.98]**	
Both	0.123	-0.018	-0.022	0.014	0.002	0.163	-0.011	0.051	0.001	0.005	-0.022	-0.021	0.70
	[2.88]**	[-0.67]	[-0.60]	[0.57]	[0.09]	[3.57]***	[-2.28]*	[2.03]*	[0.14]	[1.04]	[-2.46]*	[-3.26]**	

Table 20: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Fund Returns (k=10)

	FF6	*Textual Risk Factors (λ)*											R^2
Model	Included	Deriv	Forgn	Growth	Sector	Value	Liqd	Mgmt	Market	LgCap	Conc	Credit	SmCap	
FF6 only	✓													0.54
Textual only		-0.020	-0.021	0.020	-0.006	0.006	-0.013	-0.012	0.006	0.004	-0.015	-0.013	0.002	0.70
		[-2.26]*	[-2.25]*	[2.27]*	[-1.59]	[1.09]	[-3.43]***	[-2.25]*	[1.55]	[0.85]	[-3.50]***	[-2.50]*	[0.24]	
Both	✓	-0.021	-0.022	0.020	-0.007	0.006	-0.014	-0.013	0.006	0.004	-0.015	-0.014	0.001	0.76
		[-2.38]*	[-2.38]*	[2.30]*	[-1.77]	[1.09]	[-3.72]***	[-2.36]*	[1.50]	[0.81]	[-3.44]***	[-2.73]**	[0.20]	

Table 21: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Fund Returns (k=15)

	FF6	*Textual Risk Factors (λ)*																	$\bar{R}^2$
Model	Included	Deriv	EmrgM	Equity	Forgn	Growth	Sector	Value	Issuer	Liqd	Mgmt	Market	LgCap	Conc	Trnovr	REIT	Credit	SmCap	
FF6 only	✓																		0.54
Textual only		-0.021	-0.025	0.009	-0.021	0.020	-0.007	0.007	-0.001	-0.015	-0.006	-0.007	0.003	-0.016	-0.016	-0.002	-0.016	0.001	0.76
		[-2.37]*	[-2.71]**	[1.90]	[-2.65]**	[2.26]*	[-1.72]	[1.14]	[-0.28]	[-2.98]**	[-1.71]	[-2.90]**	[0.66]	[-3.62]***	[-2.95]**	[-0.32]	[-2.60]**	[0.21]	
Both	✓	-0.021	-0.025	0.009	-0.021	0.020	-0.007	0.007	0.000	-0.015	-0.006	-0.007	0.003	-0.016	-0.016	-0.002	-0.016	0.001	0.79
		[-2.41]*	[-2.68]**	[1.97]*	[-2.60]**	[2.30]*	[-1.79]	[1.14]	[0.01]	[-3.05]**	[-1.72]	[-2.80]**	[0.69]	[-3.62]***	[-3.04]**	[-0.33]	[-2.59]**	[0.18]	

Table 22: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Fund Returns (k=20)

	FF6	*Textual Risk Factors (λ)*																						$\bar{R}^2$
Model	Included	Conc	Credit	Oprtnl	Deriv	Forgn	Growth	Sector	IntRate	Value	Issuer	LgCap	Liqd	Mgmt	Macro	Market	SmCap	NonDiv	Trnovr	REIT	SecLend	Track	SovDbt	
FF6 only	✓																							0.54
Textual only		-0.008	-0.026	-0.003	-0.021	-0.023	0.021	-0.004	-0.019	0.007	-0.001	0.001	-0.018	0.001	-0.017	0.010	0.001	-0.022	-0.014	-0.007	-0.006	-0.019	-0.029	0.78
		[-1.86]	[-2.28]*	[-0.63]	[-2.16]*	[-2.21]*	[2.12]*	[-0.85]	[-2.15]*	[1.09]	[-0.30]	[0.26]	[-3.36]***	[0.58]	[-3.36]***	[2.31]*	[0.15]	[-3.47]***	[-3.32]***	[-0.88]	[-1.59]	[-2.59]**	[-2.11]*	
Both	✓	-0.008	-0.027	-0.003	-0.022	-0.022	0.021	-0.003	-0.019	0.007	-0.000	0.001	-0.018	0.001	-0.017	0.010	0.001	-0.021	-0.014	-0.007	-0.006	-0.018	-0.030	0.81
		[-1.75]	[-2.35]*	[-0.88]	[-2.26]*	[-2.18]*	[2.14]*	[-0.78]	[-2.19]*	[1.02]	[-0.21]	[0.23]	[-3.45]***	[0.46]	[-3.38]***	[2.12]*	[0.08]	[-3.34]***	[-3.37]***	[-0.89]	[-1.68]	[-2.58]**	[-2.19]*	

Table 23: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Fund Returns (k=25)

	FF6	*Textual Risk Factors (λ)*																								$\bar{R}^2$
Model	Included	Indust	Cpty	HiYld	Oprtnl	Deriv	Equity	ETF	Forgn	Growth	Sector	IntRate	Value	Issuer	Liqd	Mgmt	Macro	Market	LgCap	Conc	Trnovr	REIT	SecLend	SmCap	Track	SovDbt	Valuation	
FF6 only	✓																											0.54
Textual only		-0.022	-0.026	-0.023	-0.003	-0.022	0.005	-0.013	-0.022	0.021	0.008	-0.020	0.007	-0.000	-0.013	0.003	-0.016	-0.003	0.001	-0.013	-0.006	-0.007	-0.007	0.000	-0.017	-0.028	-0.018	0.81
		[-2.84]**	[-2.29]*	[-2.09]*	[-0.64]	[-2.18]*	[1.32]	[-2.54]*	[-2.16]*	[2.12]*	[1.26]	[-2.17]*	[1.04]	[-0.20]	[-3.42]***	[1.48]	[-2.84]**	[-0.82]	[0.18]	[-3.27]**	[-1.02]	[-0.84]	[-1.91]	[0.05]	[-2.49]*	[-2.04]*	[-3.94]***	
Both	✓	-0.022	-0.027	-0.023	-0.004	-0.022	0.005	-0.014	-0.022	0.021	0.008	-0.020	0.007	-0.000	-0.013	0.002	-0.016	-0.003	0.001	-0.013	-0.007	-0.007	-0.007	0.001	-0.017	-0.028	-0.018	0.83
		[-2.82]**	[-2.29]*	[-2.08]*	[-0.96]	[-2.20]*	[1.32]	[-2.63]**	[-2.16]*	[2.09]*	[1.39]	[-2.17]*	[1.07]	[-0.20]	[-3.45]***	[1.39]	[-2.86]**	[-0.88]	[0.19]	[-3.26]**	[-1.07]	[-0.84]	[-1.91]	[0.07]	[-2.49]*	[-2.05]*	[-3.96]***	

B.5 Stock Fama-MacBeth Specifications

Table 24: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Stock Returns (k=5)

	Fama–French 6 (λ)						*Textual Risk Factors (λ)*						R^2
	Mkt	SMB	HML	RMW	CMA	Mom	Mgmt	Market	SmCap	LgCap	Forgn	Deriv	
FF6 only	0.066	-0.024	-0.046	0.028	-0.009	0.261							0.08
	[1.34]	[-0.78]	[-1.24]	[0.99]	[-0.37]	[5.24]***							
Textual only							-0.007	-0.013	-0.004	0.008	-0.018	-0.008	0.07
							[-1.23]	[-0.46]	[-0.51]	[1.43]	[-1.57]	[-1.12]	
Both	0.043	-0.017	-0.045	0.015	-0.002	0.246	-0.010	0.002	0.001	0.003	-0.020	-0.012	0.12
	[0.93]	[-0.62]	[-1.23]	[0.57]	[-0.09]	[5.13]***	[-1.78]	[0.06]	[0.17]	[0.57]	[-1.85]	[-1.80]	

Table 25: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Stock Returns (k=10)

	FF6	*Textual Risk Factors (λ)*												R^2
Model	Included	Deriv	Forgn	Growth	Sector	Value	Liqd	Mgmt	Market	LgCap	Conc	Credit	SmCap	
FF6 only	✓													0.08
Textual only		-0.004	-0.023	0.024	-0.009	0.002	-0.006	-0.001	0.007	0.004	-0.014	-0.006	-0.001	0.12
		[-0.42]	[-2.09*]	[2.52*]	[-2.09*]	[0.32]	[-1.53]	[-0.19]	[2.09*]	[0.80]	[-2.82**]	[-1.17]	[-0.07]	
Both	✓	-0.006	-0.020	0.025	-0.010	0.003	-0.008	-0.004	0.006	0.001	-0.013	-0.008	0.003	0.15
		[-0.71]	[-1.97*]	[2.73**]	[-2.30*]	[0.44]	[-2.00*]	[-0.69]	[1.77	[0.20]	[-2.85**]	[-1.57]	[0.39]	

Table 26: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Stock Returns (k=15)

	FF6	*Textual Risk Factors (λ)*																	$\bar{R}^2$
Model	Included	Deriv	EmrgM	Equity	Forgn	Growth	Sector	Value	Issuer	Liqd	Mgmt	Market	LgCap	Conc	Trnovr	REIT	Credit	SmCap	
FF6 only	✓																		0.08
Textual only		-0.006	-0.030	0.003	-0.025	0.024	-0.007	0.003	-0.003	-0.006	-0.002	-0.004	0.003	-0.015	-0.009	-0.004	-0.007	-0.002	0.14
		[-0.62]	[-2.87**]	[0.60]	[-2.75**]	[2.59**]	[-1.43]	[0.49]	[-1.38]	[-1.20]	[-0.62]	[-1.43]	[0.59]	[-3.14**]	[-1.51]	[-0.47]	[-1.12]	[-0.22]	
Both	✓	-0.007	-0.031	0.003	-0.026	0.024	-0.007	0.005	-0.004	-0.007	-0.004	-0.004	0.001	-0.016	-0.011	-0.005	-0.009	0.001	0.18
		[-0.79]	[-3.02**]	[0.73]	[-2.85**]	[2.58**]	[-1.54]	[0.78]	[-1.65]	[-1.32]	[-0.92]	[-1.56]	[0.22]	[-3.25**]	[-1.86]	[-0.54]	[-1.39]	[0.11]	

Table 27: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Stock Returns (k=20)

	FF6	*Textual Risk Factors (λ)*																					$\bar{R}^2$
Model	Included	Conc	Credit	Oprtnl	Deriv	Forgn	Growth	Sector	IntRate	Value	Issuer	LgCap	Liqd	Mgmt	Macro	Market	SmCap	NonDiv	Trnovr	REIT	SecLend	Track	SovDbt	
FF6 only	✓																							0.08
Textual only		-0.011	-0.007	-0.004	-0.005	-0.020	0.026	-0.004	-0.006	0.001	-0.004	-0.001	-0.008	0.004	-0.016	0.001	0.002	-0.018	-0.006	-0.011	-0.008	-0.011	-0.006	0.17
		[-2.21*]	[-0.59]	[-0.99]	[-0.51]	[-1.71]	[2.52]*	[-0.82]	[-0.66]	[0.21]	[-1.38]	[-0.22]	[-1.44]	[1.92]	[-2.72**]	[0.11]	[0.20]	[-2.56*]	[-1.20]	[-1.12]	[-1.85]	[-1.38]	[-0.41]	
Both	✓	-0.012	-0.009	-0.006	-0.007	-0.024	0.026	-0.004	-0.008	0.003	-0.005	-0.003	-0.009	0.004	-0.017	0.001	0.003	-0.019	-0.007	-0.011	-0.009	-0.012	-0.008	0.20
		[-2.40*]	[-0.74]	[-1.27]	[-0.69]	[-2.10*]	[2.50*]	[-0.80]	[-0.84]	[0.49]	[-1.77]	[-0.46]	[-1.60]	[1.86]	[-3.07**]	[0.16]	[0.38]	[-2.75**]	[-1.51]	[-1.13]	[-2.03*]	[-1.59]	[-0.60]	

Table 28: Fama-MacBeth Estimates of Risk Premia for Textual Factors on Stock Returns (k=25)

	FF6	*Textual Risk Factors (λ)*																								$\bar{R}^2$
Model	Included	Indust	Cpty	HiYld	Oprtnl	Deriv	Equity	ETF	Forgn	Growth	Sector	IntRate	Value	Issuer	Liqd	Mgmt	Macro	Market	LgCap	Conc	Trnovr	REIT	SecLend	SmCap	Track	SovDbt	Valuation	
FF6 only	✓																											0.08
Textual only		-0.017	-0.009	-0.007	-0.006	-0.006	0.001	-0.004	-0.020	0.028	0.002	-0.007	0.001	-0.003	-0.008	0.005	-0.018	0.002	-0.002	-0.013	0.011	-0.010	-0.008	0.000	-0.010	-0.006	-0.012	0.20
		[-2.03]*	[-0.72]	[-0.59]	[-1.34]	[-0.57]	[0.18]	[-0.81]	[-1.78]	[2.67**]	[0.36]	[-0.72]	[0.17]	[-1.08]	[-2.00*]	[2.72**]	[-2.82**]	[0.65]	[-0.29]	[-2.80**]	[1.64]	[-0.97]	[-1.80]	[-0.02]	[-1.46]	[-0.44]	[-2.41*]	
Both	✓	-0.018	-0.010	-0.009	-0.007	-0.007	0.001	-0.006	-0.023	0.026	0.003	-0.008	0.003	-0.004	-0.009	0.005	-0.019	0.002	-0.003	-0.013	0.010	-0.011	-0.008	0.001	-0.011	-0.008	-0.014	0.23
		[-2.10*]	[-0.83]	[-0.76]	[-1.62]	[-0.74]	[0.20]	[-1.01]	[-2.09*]	[2.62**]	[0.40]	[-0.87]	[0.51]	[-1.52]	[-2.16*]	[2.75**]	[-3.05**]	[0.69]	[-0.46]	[-2.88**]	[1.47]	[-1.11]	[-1.94]	[0.15]	[-1.58]	[-0.59]	[-2.61**]	

B.6 Redemption Risk: Mechanism and Validation Tables

Table 29: The Anatomy of Redemption Risk: Fund Characteristics and Regressions

Panel A: Univariate Fund Characteristics		
Variable	**Non-Disclosing (0)**	**Disclosing (1)**
Total Net Assets (TNA, $M)	3,018.70	1,524.40
Common Equity (%)	86.46	80.23
Cash (%)	2.14	5.44
Turnover Ratio	0.47	0.86
Annual Return (%)	9.93	8.03
Monthly Flow (%)	1.44	17.57
Flow Volatility (%)	11.42	66.28

Panel B: Liquidity, Returns, and Flow-Performance Regressions			
	(1) Portfolio Liquidity (L)	**(2) Annual Return**	**(3) Monthly Flow**
Redemption	-0.056^{**}	-0.005^{***}	0.685^{*}
	(0.025)	(0.001)	(0.378)
Fund Return			1.351
			(1.140)
Return $\times$ Redemption			-1.917^{**}
			(0.820)
Risk Factor Controls	No	Yes (FF6)	No
Time Fixed Effects	Year	No	Month
Fund Fixed Effects	No	No	Yes
Observations	60,227	83,825	966,851
Adjusted R^2	0.002	0.102	-0.003

*Note: Panel A presents the time-series averages of fund characteristics for funds that do not disclose (Redemption = 0) and do disclose (Redemption = 1) redemption risk in their prospectuses. Panel B presents panel regressions at the fund-year and fund-month level. Model 1 regresses the standardized Ľuboš Pástor et al. (2020) portfolio liquidity measure (L) on the disclosure dummy. Model 2 regresses annual fund returns on the disclosure dummy, controlling for standard aggregate risk exposures. Model 3 regresses monthly fund flows on contemporaneous returns interacted with the disclosure dummy. Standard errors are clustered at the fund portfolio level. *, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.*

Table 30: Stock Characteristics Sorted by Redemption Fund Ownership

	Low	**Q2**	**Q3**	**Q4**	**High**
Exposures					
Redemption Factor Beta	-1.710	-0.431	-0.562	-0.229	0.481
DKW Common Flow Beta	-0.195	-0.169	-0.134	-0.100	-0.126
Market Beta	0.774	0.928	0.942	0.932	0.795
SMB Beta	0.494	0.569	0.695	0.787	0.723
Firm Characteristics					
Log Size (Market Cap)	13.10	14.20	13.90	13.50	12.70
Book-to-Market (BM)	0.689	0.616	0.664	0.693	0.805
Operating Profitability (OP)	-3.090	-0.297	-0.516	0.059	-0.088
Aggregate MF Ownership	21.4%	2.1%	1.9%	1.9%	54.3%

Note: This table presents the average characteristics of stocks sorted into quintiles based on the percentage of their shares held by mutual funds that explicitly disclose Redemption Risk. Variables are reported as time-series averages of the cross-sectional quintile means. The sample restricts to firm-years with available ownership data and characteristics. Aggregate Ownership represents the total mutual fund ownership share of the stock's market capitalization.

B.7 Concentration Risk: Mechanism and Validation Tables

Table 31: Institutional and Performance Characteristics of Concentration-Disclosing Funds

Variable	**Non-Disclosing**	**Disclosing**
Total Net Assets (TNA, $M)	2,914.00	1,358.00
Portfolio HHI	0.0769	0.1310
Fund Age (Years)	13.40	12.00
Family Size (Number of Funds)	383	285
Manager Workload (Number of Funds)	467	506
Annual Return (%)	9.62	9.14
Return Volatility	0.1650	0.8650
Annual Alpha	-0.000036	-0.000036
Flow Volatility	0.1620	0.7940
Number of Fund-Years	36,739	46,457

Note: This table presents the time-series averages of institutional, portfolio, and performance characteristics for mutual funds that do not disclose (0) and do disclose (1) Concentration Risk in their prospectuses. Fund Age is measured in years since the first offer date. Family Size represents the average number of unique funds managed under the same management company code. Manager Workload represents the average number of unique funds advised by the same portfolio manager. Return and flow volatility are presented as raw decimals.

Table 32: Alpha-Based Tournament Incentives: The Convexity of Fund Flows

	Dependent Variable: Annual Flow
Concentration	0.118^{**}
	(0.038)
Bottom Alpha Tercile (Dummy)	0.313
	(0.422)
Top Alpha Tercile (Dummy)	-0.026
	(0.025)
Bottom Alpha $\times$ Concentration	-0.262
	(0.427)
Top Alpha $\times$ Concentration	0.085^{*}
	(0.043)
Time Fixed Effects	Year
Observations	72,562
Adjusted R^2	0.000

*Note: This table presents panel regressions estimating the flow-performance relationship using risk-adjusted performance (Alpha) to test the convex tournament hypothesis. The dependent variable is standardized, winsorized annual fund flows. Independent variables are indicator variables for the top and bottom terciles of lagged annual alpha (with the middle tercile serving as the omitted baseline), interacted with the Concentration disclosure dummy. Standard errors are clustered at the fund portfolio level. *, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.*

C Risk Topic Descriptions

This appendix provides summarized descriptions of the risk topics identified through k-means clustering of the risk disclosure embeddings. Each subsection corresponds to a different number of clusters (k). For clarity, each topic's full name is followed by the abbreviated label (e.g., *Mgmt*) used for that factor in the corresponding Fama-MacBeth results tables.

C.1 Descriptions for k=5 Topics

Table 33: Risk Topic Summaries (k=5)

Risk Topic	Description
Management Risk *(Mgmt)*	The Fund is subject to management risk, as its performance depends on the investment adviser's skill in security selection, strategy, and portfolio construction. The adviser's investment analysis may not produce the intended results, and its judgments about the attractiveness or value of a particular security, industry, or sector may prove to be incorrect, potentially causing the Fund to underperform its benchmark or other funds with similar objectives. This risk may be heightened if the adviser's strategy leads to a concentrated portfolio, where the Fund invests a significant portion of its assets in a limited number of issuers, making it more susceptible to adverse events affecting those holdings. Furthermore, the Fund may invest in securities that are or become illiquid, which could make it difficult to sell those assets at a favorable time or price, especially to meet redemption requests, potentially forcing the Fund to sell them at a loss.

Table 33 – continued from previous page

Risk Topic	Description
Foreign Risk *(Forgn)*	Investments in foreign securities involve special risks not typically associated with U.S. investments, including adverse political, economic, regulatory, and social developments. Foreign markets, particularly emerging markets, may be less liquid, more volatile, and subject to less stringent accounting and disclosure standards than U.S. markets; these risks are often magnified in emerging markets, which can be subject to greater instability and may be considered speculative. Furthermore, the value of the fund's assets is subject to currency risk, as the value of foreign currencies may fluctuate significantly and unpredictably relative to the U.S. dollar. A decline in the value of a foreign currency against the U.S. dollar will reduce the value of the fund's investments denominated in that currency, potentially offsetting investment gains or increasing losses.

Table 33 – continued from previous page

Risk Topic		Description
Derivatives *(Deriv)*	**Risk**	The Fund may use complex financial instruments, such as derivatives, and engage in securities lending, which involve risks that may be different from, and possibly greater than, the risks of investing directly in securities. The use of derivatives can create investment leverage, magnifying gains or losses and potentially resulting in losses that exceed the Fund's initial investment. These instruments are subject to counterparty risk, where the other party to a transaction may be unable or unwilling to fulfill its contractual obligations. Similarly, securities lending involves the risk that a borrower may fail to return loaned securities, and the Fund could suffer a loss if the value of the collateral provided declines or if cash collateral is invested unsuccessfully. Both strategies may also be subject to liquidity risk, making it difficult to close out a position at a favorable time or price.

Table 33 – continued from previous page

Risk Topic	Description
Market Risk *(Market)*	The value of equity securities fluctuates, often rapidly and unpredictably, due to factors affecting both the broader market and individual companies. Stock markets as a whole are volatile and can decline in response to adverse economic, political, or regulatory developments, and they tend to move in cycles of rising and falling prices. The value of an individual security can also fall due to factors unique to its issuer, such as poor financial results or industry-specific challenges, causing it to underperform the market. Furthermore, different investment styles, such as growth or value, carry distinct risks and may shift in and out of favor with investors, impacting performance.
Small-Cap Risk *(Sm-Cap)*	Investments in small- and mid-capitalization companies involve greater risks than those associated with larger, more established companies. The securities of these companies tend to be more volatile and less liquid, trading in lower volumes, which can make it difficult to sell shares at a desired time or price. Compared to their larger counterparts, these companies may have more limited product lines, markets, and financial resources, less management depth, and shorter operating histories, making them more vulnerable to adverse business, economic, or market conditions. As a result, the securities of small- and mid-sized companies may underperform the market as a whole and are subject to more abrupt and erratic price movements.

Table 33 – continued from previous page

Risk Topic	Description
Large-Cap Risk *(LgCap)*	The risk that investments in large, established companies may underperform the broader market or securities of smaller companies. Larger companies may be less able to adapt to changing market conditions and respond quickly to new competitive challenges, such as changes in technology or consumer tastes. Additionally, due to their maturity, these companies may be subject to more limited growth potential and unable to achieve the high growth rates of successful smaller companies. As a result, large-capitalization stocks can fall out of favor with investors for extended periods, causing their returns to trail those of other market segments.

C.2 Descriptions for k=10 Topics

Table 34: Risk Topic Summaries (k=10)

Risk Topic	Description
Redemption Risk *(Liqd)*	The fund's shares may trade in the secondary market at prices that differ from their Net Asset Value (NAV), sometimes significantly. This risk is heightened by the potential for a lack of an active trading market, trading halts, or early market closures, which can lead to shares trading at a premium or discount to NAV. The fund also faces liquidity risk, as certain investments may become illiquid or difficult to trade during periods of market stress or heavy redemptions, potentially forcing sales at a loss or below perceived value. Furthermore, the fund is subject to operational risks, such as cybersecurity breaches, which could disrupt business operations and negatively impact the fund and its shareholders.
Management Risk *(Mgmt)*	The fund's performance is subject to the risks associated with its investment strategy. For actively managed funds, there is the risk that the manager's security selections and investment decisions will fail to produce the desired results, leading to underperformance against the fund's benchmark or peers. For passively managed or index-tracking funds, the primary risk is an imperfect correlation, or tracking error, between the fund's performance and that of its target index. A passive fund will not take defensive positions in declining markets and will continue to hold securities regardless of their investment merit, which can lead to losses.

Table 34 – continued from previous page

Risk Topic	Description
Credit Risk *(Credit)*	The value of a security can decline due to factors specific to its issuer, such as a deterioration in its financial condition. This creates credit risk, which is the risk that an issuer will be unable or unwilling to make timely interest or principal payments, potentially leading to a default. The fund is also subject to risk from its securities lending activities, where it could experience losses if a borrower fails to return a loaned security or if the value of the collateral held declines. These risks apply to various instruments, including bonds, preferred stocks, and convertible securities.
Sector Risk *(Sector)*	The fund may concentrate its investments in a particular industry or group of related industries, such as information technology, financials, or real estate. This makes the fund's performance highly sensitive to economic, political, regulatory, or market developments affecting that specific sector. Adverse events within the concentrated area, such as unfavorable government regulation or rapid product obsolescence, could have a much greater negative impact on the fund's value than on a more broadly diversified portfolio.

Table 34 – continued from previous page

Risk Topic	Description
Derivatives Risk *(Deriv)*	The use of derivatives, such as futures, options, and swaps, involves risks beyond those of direct security investments. These instruments can be highly volatile and create economic leverage, which magnifies both gains and losses and may result in losses exceeding the amount initially invested. The fund also faces counterparty risk, which is the risk that the other party in a derivative transaction will be unable or unwilling to fulfill its contractual obligations. A counterparty's failure to perform could result in a significant financial loss to the fund.
Market Risk *(Market)*	The value of equity securities fluctuates, sometimes rapidly and unpredictably, in response to a variety of factors. These can include broad economic, political, or market developments that affect the entire stock market, as well as events that are specific to a particular company or industry. Stock markets tend to move in cycles, with periods of rising and falling prices. A general market decline could cause the value of the fund's investments to fall, regardless of the fundamental performance of individual holdings.

Table 34 – continued from previous page

Risk Topic	Description
Concentration Risk *(Conc)*	As a non-diversified fund, the fund may invest a larger portion of its assets in a smaller number of issuers, making its performance more susceptible to the financial health of those few companies. Furthermore, the fund may concentrate its investments in a particular industry, sector, or geographic region, making it more vulnerable to adverse economic, political, or regulatory events affecting that specific area. This lack of diversification can lead to greater price volatility compared to a more broadly diversified fund.
Foreign Risk *(Forgn)*	Investing in foreign securities involves special risks not typically associated with U.S. investments, including political and economic instability, less stringent regulatory and accounting standards, and lower market liquidity. These risks are often magnified in emerging markets. Additionally, the fund is subject to currency risk, as the value of its foreign investments can be adversely affected by fluctuations in currency exchange rates. A decline in the value of foreign currencies relative to the U.S. dollar will reduce the fund's returns.

Table 34 – continued from previous page

Risk Topic	Description
Small-Cap Risk *(SmCap)*	Securities of small- and mid-sized companies involve greater risks than those of larger, more established companies. These companies may have more limited financial resources, narrower product lines, and less experienced management. As a result, their stock prices tend to be more volatile and less liquid, and they may be more susceptible to adverse economic or market developments, potentially leading to greater fluctuations in the fund's value.
Large-Cap Risk *(LgCap)*	The risk that investments in large, established companies may underperform the broader market or securities of smaller companies. Due to their size and maturity, larger companies may be less able to adapt to changing market conditions and respond quickly to new competitive challenges, such as changes in technology or consumer tastes. Consequently, these companies may be subject to more limited growth potential, and their returns may trail other market segments for extended periods.

Table 34 – continued from previous page

Risk Topic	Description
Value Risk *(Value)*	A value-oriented strategy focuses on securities believed to be trading at a price below their intrinsic worth. This approach carries the risk that the market may not recognize a security's value for a long period, or at all, and the stock price may not appreciate as anticipated. A stock judged to be undervalued may actually be appropriately priced or even decline further. As an investment style, value investing can fall out of favor with investors, causing the fund to underperform the broader market or funds with different investment styles for extended periods.
Growth Risk *(Growth)*	A growth-oriented strategy focuses on companies with high potential for future earnings growth, whose stock prices are often based on high expectations. This style of investing can be more volatile than other strategies, as a company's stock price may decline significantly if it fails to meet these earnings expectations. Growth stocks often pay little to no dividends, which can reduce the cushion against price declines during market downturns, and the growth style itself may fall out of favor with investors, leading to periods of underperformance.

C.3 Descriptions for k=15 Topics

Table 35: Risk Topic Summaries (k=15)

Risk Topic	Description
Redemption Risk *(Liqd)*	The risk that the Fund may be unable to sell a security at an advantageous time or price. During periods of market stress or heavy redemptions, certain investments may become illiquid or difficult to trade, which could force the Fund to sell them at a loss or for less than their perceived value. This is particularly true for restricted securities or derivatives traded in over-the-counter markets. Such illiquidity can negatively impact the Fund's performance, hinder its ability to meet shareholder redemption requests, and make it more challenging to value its holdings accurately.
Credit Risk *(Credit)*	The risk of loss from lending portfolio securities and from the creditworthiness of debt issuers. When lending securities, the Fund faces the risk that the borrower may fail to return the securities in a timely manner or at all, and that the value of the collateral held could decline. Additionally, the Fund is subject to credit risk, which is the possibility that an issuer of a debt security will be unable to make timely interest or principal payments, or that the issuer's perceived financial health will deteriorate, causing the security's value to fall. These risks can be influenced by changes in interest rates, which affect the value of debt instruments and the cost of borrowing.

Table 35 – continued from previous page

Risk Topic	Description
Management Risk *(Mgmt)*	The risk that the Fund's investment strategy may not produce the intended results, leading to underperformance. For actively managed funds, this includes the possibility that the investment manager's security selections, asset allocation decisions, and market judgments are incorrect. For passively managed funds that track an index, the risk is that the fund will not take defensive positions in declining markets and will continue to hold securities regardless of their investment merit, potentially leading to losses and tracking errors relative to its underlying index.
Concentration Risk *(Conc)*	The risk of loss due to the Fund's focus on a limited number of investments. As a non-diversified fund, it may invest a larger percentage of its assets in the securities of a single issuer, which means changes in the market value of that one investment could cause greater NAV fluctuations than in a diversified fund. This risk is magnified if the Fund concentrates its investments in a particular industry, sector, or geographic region, making it more susceptible to adverse economic, political, or regulatory events affecting that specific area.

Table 35 – continued from previous page

Risk Topic	Description
Derivatives Risk *(Deriv)*	The risk associated with using financial instruments such as futures, options, and swaps. The value of a derivative can be highly volatile and may not perfectly correlate with its underlying asset, potentially resulting in losses that exceed the amount initially invested. The use of derivatives can also create leverage, magnifying both gains and losses. Furthermore, the Fund is exposed to counterparty risk, where the other party in a derivative transaction may be unwilling or unable to fulfill its contractual obligations, which could lead to significant financial loss.
Emerging Markets Risk *(EmrgM)*	The risk that investing in securities of foreign issuers, particularly those in emerging markets, may involve greater volatility and risk than U.S. investments. Foreign markets can be affected by adverse political, economic, or regulatory developments, currency exchange rate fluctuations, and less-developed legal and accounting standards. These risks are often magnified in emerging markets, which may experience greater instability, lower trading volumes, and less liquidity, potentially leading to speculative price movements and a higher risk of loss.

Table 35 – continued from previous page

Risk Topic	Description
Sector Risk *(Sector)*	The risk that market conditions or economic developments could significantly affect a single industry or a group of related industries. When the Fund concentrates its investments in a specific sector, such as Technology, Financials, or Healthcare, its performance becomes closely tied to the factors impacting that area. These factors can include government regulation, intense competition, product obsolescence, and changes in consumer demand, which may cause the Fund's value to be more volatile than that of a fund with broader diversification across multiple sectors.
Foreign Risk *(Forgn)*	The risk that the value of the Fund's investments will be adversely affected by factors specific to foreign markets. Investments in non-U.S. securities are subject to risks from political and economic instability, less stringent regulatory and accounting standards, and potentially lower market liquidity. Additionally, the Fund is exposed to currency risk, as the value of securities denominated in a foreign currency can decline if that currency depreciates against the U.S. dollar. Currency exchange rates can be unpredictable and volatile, potentially reducing the Fund's returns.

Table 35 – continued from previous page

Risk Topic	Description
Issuer Risk *(Issuer)*	The risk that the value of an individual security may decline for reasons directly related to the company that issued it, even when the broader market is performing well. Factors such as poor management performance, a deteriorating financial condition, reduced demand for the issuer's products or services, or negative investor perceptions can cause a security's price to be more volatile than the market as a whole and lead to underperformance.
Turnover Risk *(Trnovr)*	The risk that external market mechanics and internal processes can negatively impact the Fund. Market trading risks include the potential for Fund shares to trade at a significant premium or discount to their Net Asset Value (NAV), especially during periods of high volatility, and the possibility of trading halts. Operational risk stems from potential failures or breaches in the systems of the Fund or its service providers, including cybersecurity threats, human error, or technology failures, which could disrupt business operations and result in financial losses.

Table 35 – continued from previous page

Risk Topic	Description
Equity Risk *(Equity)*	The risk that stock markets will be volatile and that the value of equity securities will fluctuate, sometimes rapidly and unpredictably. These price movements may result from factors affecting individual companies, specific industries, or the market as a whole, such as adverse political, economic, or regulatory developments. Because stock markets tend to move in cycles, the Fund's investments may decline in value and underperform other asset classes over short or extended periods.
Real Estate Risk *(REIT)*	The risk that the Fund's investments in real estate securities, including Real Estate Investment Trusts (REITs), will be adversely affected by factors impacting the real estate market. These risks include declines in property values, changes in interest rates, increases in property taxes, and shifts in zoning laws or local economic conditions. REITs are also subject to management risk and may have limited diversification. As an investor in REITs, the Fund indirectly bears its proportionate share of the expenses of the underlying real estate companies.

Table 35 – continued from previous page

Risk Topic	Description
Market Risk *(Market)*	The risk that the value of securities will decline due to factors affecting the overall financial markets. Market values can fluctuate, sometimes rapidly and unpredictably, in response to general economic conditions, changes in interest rates, political events, or shifts in investor sentiment. Events such as global pandemics, geopolitical conflicts, or actions taken by central banks can disrupt economies and markets, causing the Fund's investments to lose value regardless of the performance of individual issuers.
Large-Cap Risk *(LgCap)*	The risk that investments in large, established companies may underperform the broader market or securities of smaller companies. Larger companies may be less able to adapt to changing market conditions and respond quickly to new competitive challenges, such as changes in technology or consumer tastes. Additionally, due to their maturity, these companies may be subject to more limited growth potential and unable to achieve the high growth rates of successful smaller companies. As a result, large-capitalization stocks can fall out of favor with investors for extended periods, causing their returns to trail those of other market segments.

Table 35 – continued from previous page

Risk Topic	Description
Small-Cap Risk *(Sm-Cap)*	The risk that securities of small and mid-sized companies may be more volatile and less liquid than those of larger, more established companies. These companies often have more limited product lines, fewer financial resources, and less management depth, making them more vulnerable to adverse business or economic developments. As a result, their stock prices may fluctuate more significantly than the overall market, and they may underperform larger companies during certain market cycles.
Value Risk *(Value)*	The risk that stocks believed to be undervalued may not appreciate as anticipated or may continue to be out of favor with the market for long periods. A value-oriented strategy may cause the Fund to underperform growth-style funds or the broader market, particularly when market and economic conditions favor companies with rapid earnings growth. There is also the risk that a security judged to be undervalued is actually appropriately priced, and its value may not increase as expected or could decline further.

Table 35 – continued from previous page

Risk Topic	Description
Growth Risk *(Growth)*	The risk that the prices of growth stocks may be more volatile than other types of stocks and may fall sharply if the issuing company fails to meet investor expectations for future earnings or revenue. Growth companies often trade at higher valuations and may not pay dividends, making their stock prices highly sensitive to market sentiment and changes in economic conditions. This investment style can go in and out of favor, and during periods when the market does not favor growth stocks, the Fund may underperform funds that follow a different investment style.

C.4 Descriptions for k=20 Topics

Table 36: Risk Topic Summaries (k=20)

Risk Topic	Description
Market Risk *(Market)*	The risk that the value of equity securities will fluctuate, sometimes rapidly and unpredictably, in response to developments affecting individual companies, entire industries, or the securities market as a whole. Stock markets tend to move in cycles, with periods of rising and falling prices. Factors such as adverse political or economic conditions, changes in interest rates, or negative investor sentiment can cause a general market downturn, which could cause the value of a fund's investments to decline, resulting in a loss.

Table 36 – continued from previous page

Risk Topic	Description
Issuer Risk *(Issuer)*	The risk that the value of an individual security may decline for reasons directly related to its issuer, such as poor management performance, declining financial health, reduced demand for its products or services, or other adverse business developments. An issuer-specific event can cause a security's value to underperform the broader market, even when general market conditions are favorable, which would negatively affect the fund's performance.
Sector Risk *(Sector)*	The risk that a fund's performance may be driven by the performance of a particular industry or economic sector. When a fund concentrates its investments in a specific sector, it becomes more susceptible to adverse economic, business, regulatory, or other developments affecting that sector. This lack of diversification can result in greater price volatility than a fund that is invested more broadly across various industries.
Credit Risk *(Credit)*	The risk that a party to a financial contract, such as a derivative or repurchase agreement, will be unable or unwilling to meet its contractual obligations. If a counterparty becomes bankrupt or otherwise fails to perform, the fund could experience a loss of the full value of the transaction, delays in accessing its collateral, or be unable to realize any gains on the investment, resulting in a decline in the fund's value.

Table 36 – continued from previous page

Risk Topic	Description
Macro Risk *(Macro)*	The risk that events in the U.S. and global financial markets will cause unusual market volatility and negatively impact performance. Economies and financial markets are increasingly interconnected, and political, economic, or social instability in one country or region can adversely affect markets worldwide. Events such as war, acts of terrorism, natural disasters, or public health crises like pandemics can disrupt global supply chains, reduce consumer demand, and lead to significant and prolonged declines in the value of the fund's investments.
Concentration Risk *(Conc)*	The risk of loss due to a fund's investments being concentrated in a particular issuer, industry, sector, or geographic region. By focusing on a specific area, the fund's performance becomes closely tied to the economic, political, and market conditions of that area. As a result, the fund may be subject to greater volatility and its value may be more susceptible to adverse developments affecting that area of focus than a more broadly diversified fund.

Table 36 – continued from previous page

Risk Topic	Description
Foreign Risk *(Forgn)*	The risk that investments in foreign securities, particularly in emerging markets, may be subject to greater volatility and different risks than U.S. investments. These risks include adverse political and economic developments, less-regulated markets, potential trading and settlement difficulties, and fluctuations in currency exchange rates. A decline in the value of a foreign currency relative to the U.S. dollar will reduce the value of an investment denominated in that currency. These factors can make foreign investments, especially those in emerging markets, less liquid and more unpredictable.
Management Risk *(Mgmt)*	The risk that the investment manager's strategies, analyses, and investment decisions will not produce the desired results. An investment manager's judgments about the attractiveness, value, or potential appreciation of a particular security or asset class may prove to be incorrect, causing the fund to underperform its benchmark, other funds with similar objectives, or the market as a whole, and may result in investment losses.

Table 36 – continued from previous page

Risk Topic	Description
Sovereign Risk *(SovDbt)*	The risk that the value of government debt securities will fluctuate. While U.S. Treasury securities are backed by the full faith and credit of the U.S. government, securities issued by government agencies and instrumentalities may not be, subjecting them to greater credit risk. Sovereign debt issued by foreign governments involves the risk that the governmental entity may be unable or unwilling to pay interest and repay principal when due because of political, economic, or other factors.
Lending Risk *(SecLend)*	The risk of loss associated with lending portfolio securities to qualified institutions. This includes the risk that the borrower may fail to return the securities in a timely manner or at all (borrower default risk), and the risk that the fund may suffer losses on the investment of cash collateral received for the loan (investment risk). A decline in the value of the collateral or delays in its recovery could negatively impact the fund's performance.
Non-Diversification Risk *(NonDiv)*	The risk that a fund that is non-diversified may invest a larger percentage of its assets in the securities of a smaller number of issuers. As a result, the fund is more susceptible to the performance of a single issuer, and a decline in the market value of that one security could cause a greater fluctuation in the fund's share price than would occur in a diversified fund. This concentration can lead to increased volatility.

Table 36 – continued from previous page

Risk Topic	Description
Operational Risk *(Oprtnl)*	The risk of loss resulting from inadequate or failed internal processes, people, and systems, or from external events. This includes human error, processing failures, and technology disruptions. A cybersecurity breach could allow an unauthorized party to gain access to fund assets or sensitive data, leading to financial loss, data corruption, or loss of operational functionality for the fund, its service providers, or the issuers in which it invests.
Tracking Risk *(Track)*	The risk associated with an indexing strategy where the fund seeks to track the performance of a benchmark index, regardless of its investment merit. Unlike an actively managed fund, a passive fund will not take defensive positions in declining markets. The fund's performance may not perfectly match that of its index due to fees, expenses, and transaction costs, a difference known as tracking error. Additionally, errors in the index's construction or calculation may adversely impact the fund.
Derivatives Risk *(Deriv)*	The risk that the use of derivative instruments—such as futures, options, or swaps—may not perform as expected and can lead to losses that are greater than the amount originally invested. Derivatives can be highly volatile, involve leverage, and may be subject to counterparty, liquidity, and valuation risks. Leverage can magnify both gains and losses, causing the fund's value to be more volatile than if it had not been leveraged.

Table 36 – continued from previous page

Risk Topic	Description
Redemption Risk *(Liqd)*	The risk that a particular security or asset may be difficult to purchase or sell at a desired time and price, potentially preventing the fund from liquidating a position without negatively impacting its market value. Illiquid securities may be hard to value and may need to be sold at a loss to meet redemption requests or other cash needs, a risk that can be magnified during periods of market turmoil.
Real Estate Risk *(REIT)*	The risk that investments in real estate securities, including Real Estate Investment Trusts (REITs), will be subject to risks similar to those associated with direct ownership of real estate. These risks include declines in property values, risks related to general and local economic conditions, changes in interest rates and property taxes, and tenant defaults. REITs also depend on management skill and may have limited diversification, which can increase their volatility.
Turnover Risk *(Trnovr)*	The risk that a fund's shares may trade in the secondary market at a price that is different from its Net Asset Value (NAV), known as a premium or discount. This can be caused by a lack of an active trading market, disruptions in the creation and redemption process, or periods of high market volatility. Additionally, market closures or trading halts can prevent the fund from rebalancing its portfolio and may result in trading losses.

Table 36 – continued from previous page

Risk Topic	Description
Interest Rate Risk *(IntRate)*	The market value of fixed-income securities generally declines when interest rates rise, with longer-duration instruments being more sensitive to such fluctuations. Changes in interest rates also affect debt instruments through prepayment risk, where falling rates lead to early repayment and reinvestment at lower yields, and extension risk, where rising rates extend the life of assets, delaying reinvestment at potentially higher rates. Convertible securities and preferred stocks also see their values influenced by movements in general interest rate levels.
Value Risk *(Value)*	The risk that a stock judged to be undervalued may not appreciate as anticipated or may decline in price. A security's intrinsic value may not be recognized by the broader market for a long period of time, or the factors expected to increase its price may not occur. Additionally, value stocks as a category can perform differently from the market as a whole and may underperform growth stocks or other investment styles for extended periods.

Table 36 – continued from previous page

Risk Topic	Description
Growth Risk *(Growth)*	The risk that the prices of growth stocks may be more volatile than other types of stocks because their valuations often reflect high expectations for future earnings. If a growth company's earnings fail to meet these expectations, its stock price may decline significantly. Growth stocks may also lack the dividend yield that can cushion prices in a market downturn and, as an investment style, may underperform the broader market or value stocks for extended periods.
Small-Cap Risk *(SmCap)*	Investments in small- and mid-capitalization companies may exhibit higher volatility, lower liquidity, and sharper price fluctuations than those in larger, more established firms. Smaller issuers often have shorter operating histories, less predictable earnings, and more limited managerial and financial resources. Their narrower product offerings can make them more sensitive to adverse economic or business conditions.
Large-Cap Risk *(LgCap)*	Holdings in large-capitalization firms may lag behind the broader market or smaller companies during particular phases of the market cycle. Larger, mature enterprises tend to have slower growth prospects and may respond less swiftly to shifts in economic conditions, consumer preferences, or technological change than smaller, more dynamic competitors.

C.5 Descriptions for k=25 Topics

Table 37: Risk Topic Summaries (k=25)

Risk Topic	Description
Redemption Risk *(Liqd)*	The risk that a particular investment may be difficult to purchase or sell at a desirable time or price, or that there may not be an active trading market. This illiquidity makes it challenging to value the asset accurately and may force the Fund to sell holdings at a loss to meet shareholder redemption requests, a risk heightened during periods of market turmoil or adverse economic conditions.
Counterparty Risk *(Cpty)*	The risk that a party with whom the Fund enters into an agreement (such as a derivative contract, repurchase agreement, or securities lending transaction) will be unable or unwilling to honor its contractual obligations. If a counterparty defaults or becomes bankrupt, the Fund may suffer a loss of its investment, incur delays in accessing collateral, or be unable to realize gains on the transaction, potentially leading to a decline in the Fund's value. This risk is inherent in over-the-counter transactions and arrangements involving posted collateral.

Table 37 – continued from previous page

Risk Topic	Description
Sovereign Risk *(SovDbt)*	The risk that investments in government debt securities (sovereign debt), whether foreign or domestic, will fluctuate in value. While U.S. Treasury securities are backed by the full faith and credit of the U.S. government, securities issued by government agencies and instrumentalities may not be, subjecting them to greater credit risk. Foreign sovereign debt carries the additional risk that the issuing government may be unwilling or unable to pay interest and repay principal when due because of political, economic, or cash flow factors, and legal recourse may be limited in the event of default.
Growth Risk *(Growth)*	The risk that the prices of growth stocks will be more volatile than other types of stocks because their high valuations reflect expectations for future earnings growth. If a growth company fails to meet these expectations, its stock price may decline significantly. Growth stocks may also lack the dividend yield that can cushion prices in a market downturn, and the growth investment style may underperform the broader market or value stocks for extended periods.

Table 37 – continued from previous page

Risk Topic	Description
Industry Risk *(Indust)*	The risk that the Fund's performance will be highly sensitive to adverse economic, political, or regulatory developments affecting a specific industry or group of related industries, such as Consumer Discretionary, Industrials, Materials, or Energy. Concentration in a single industry makes the Fund more susceptible to factors unique to that sector—including commodity price volatility, changes in consumer demand, intense competition, supply chain disruptions, and government regulation—which can lead to greater volatility than a more broadly diversified portfolio.
Lending Risk *(SecLend)*	The risk that the Fund will incur a loss from lending its portfolio securities to institutions. This risk stems from the possibility that the borrower may fail to return the securities when due (borrower default risk), delays in recovering the loaned securities, or losses arising from the reinvestment of cash collateral if the value of those investments declines. The Fund may also be required to sell other portfolio holdings at a loss to satisfy obligations to the defaulting borrower.

Table 37 – continued from previous page

Risk Topic	Description
Equity Risk *(Equity)*	The risk that the value of equity securities (including common and preferred stocks) will fluctuate, sometimes rapidly and unpredictably, in response to factors affecting individual companies, specific industries or sectors, or general market and economic conditions. Equity markets are volatile and tend to move in cycles, and since equity holders rank junior to debt holders, they entail a greater risk of loss, including the potential for significant decline over short or extended periods.
Operational Risk *(Oprtnl)*	The risk of loss or adverse impact to the Fund arising from inadequate or failed internal processes, human error, errors of service providers, or systems and technology failures. This includes Cybersecurity Risk, where intentional or unintentional breaches of the Fund's or its service providers' electronic systems—such as through hacking or malicious software—may lead to unauthorized access to assets, data corruption, loss of proprietary information, or a disruption of operational capacity.

Table 37 – continued from previous page

Risk Topic	Description
Tracking Risk *(Track)*	The risk that the Fund's performance will not perfectly match that of its target benchmark index, resulting in tracking error. As a passively managed fund, it must hold securities regardless of individual investment merit and generally will not take defensive positions in declining markets. Tracking error may be caused by operating expenses, transaction costs, valuation differences, differences between the Fund's portfolio and the Index composition, and errors in the Index's calculation or methodology.
Value Risk *(Value)*	The risk that the value-oriented strategy employed by the Fund may cause it to underperform the broader market or funds that use a different style, particularly when growth stocks are in favor. Stocks believed to be undervalued may remain inexpensive for extended periods, may never realize their anticipated intrinsic value, or the determination of undervaluation may prove incorrect, leading to a decline in price.
Derivatives Risk *(Deriv)*	The risk that the use of complex financial instruments, such as futures, options, and swaps, may expose the Fund to risks beyond those of direct investments. Derivatives can be highly volatile and may create significant investment leverage, magnifying both gains and losses and potentially resulting in losses that exceed the amount initially invested. Additional risks include imperfect correlation with underlying assets, illiquidity, valuation difficulties, counterparty default, and potential adverse tax consequences.

Table 37 – continued from previous page

Risk Topic	Description
Macro Risk *(Macro)*	The risk that the Fund's investments will be adversely affected by events or conditions concentrated in a specific country, region, or set of interconnected global markets, including the U.S. This encompasses exposure to economic, political, and social instability; adverse regulatory or legal changes; security concerns such as terrorism or conflict; and widespread public health crises (e.g., pandemics or epidemics) or natural disasters that cause unusual market volatility and disrupt global supply chains and economies.
Concentration Risk *(Conc)*	The risk that the Fund, by investing a relatively large portion of its assets in a small number of issuers, a specific industry, sector, or geographic region, will be more susceptible to adverse economic, market, or regulatory events affecting that limited area of focus. This lack of broad diversification can lead to greater price volatility and cause the Fund's performance to be disproportionately impacted by the performance of a single security or issuer.
High-Yield Risk *(HiYld)*	The risk that the Fund's investments in debt securities rated below investment grade (commonly known as "high-yield" or "junk bonds") will be subject to greater credit risk, default risk, and price volatility than higher-rated securities. These securities are considered speculative, are more sensitive to adverse economic and issuer-specific developments, and may be less liquid and more difficult to value accurately, potentially leading to substantial losses.

Table 37 – continued from previous page

Risk Topic	Description
Foreign Risk *(Forgn)*	The risk that investments in foreign securities, including those in emerging markets, may be adversely affected by greater political, economic, and social instability, less stringent regulatory standards, and less liquidity than U.S. investments. This risk is compounded by currency risk, where adverse fluctuations in foreign exchange rates against the U.S. dollar can reduce the value of foreign-denominated investments or lead to losses, and by custodial and settlement risks inherent in less-developed markets.
Turnover Risk *(Trnovr)*	The risk that the Fund's investment strategy will result in frequent buying and selling of portfolio securities, leading to a high portfolio turnover rate (e.g., 100% or more). This increased trading volume results in correspondingly higher transaction costs, including brokerage commissions, which may negatively affect the Fund's performance. High turnover can also increase the realization and distribution of capital gains, leading to higher tax liabilities for shareholders.
ETF Risk *(ETF)*	The risk that the Fund's investments in Exchange-Traded Funds (ETFs) or other investment companies will subject the Fund to the risks of the ETF's underlying holdings and its operational structure. ETF shares may trade at a discount or premium to their Net Asset Value (NAV), introducing secondary market trading risk. Additionally, the Fund indirectly bears a proportionate share of the ETF's operating expenses, creating a potential duplication of fees.

Table 37 – continued from previous page

Risk Topic	Description
Real Estate Risk *(REIT)*	The risk that investments in real estate securities, including Real Estate Investment Trusts (REITs), will be subject to risks similar to direct real estate ownership. These include declines in property values, sensitivity to interest rate changes, fluctuations in rental income and occupancy rates, changes in property taxes and zoning laws, and adverse economic conditions in local or general real estate markets. Additionally, REITs are subject to risks related to management quality and favorable tax treatment requirements.
Interest Rate Risk *(IntRate)*	The risk that the value of fixed-income and related securities, including convertible and preferred stocks, will decline due to a rise in interest rates. Securities with longer maturities or durations are more sensitive to these changes. This risk also includes prepayment risk (when falling rates cause principal to be repaid early, forcing reinvestment at lower yields) and extension risk (when rising rates slow repayments, locking the Fund into lower-yielding securities longer than anticipated).
Issuer Risk *(Issuer)*	The risk that the value of an individual security or investment may decline due to factors specific to the issuer, even if the overall market is performing well. These issuer-specific events include poor management decisions, declining financial condition, adverse competitive pressures, reduced demand for products or services, credit rating downgrades, or other internal factors that cause the issuer's price to be more volatile or to underperform the broader market.

Table 37 – continued from previous page

Risk Topic	Description
Valuation Risk *(Valuation)*	The risk that the Fund's price upon selling a security or asset may differ significantly from the value reflected on the Fund's books, particularly for instruments traded in low-volume, volatile, or non-U.S. markets, or those valued using fair value methodologies. This risk is compounded by authorized participant concentration risk, where a limited number of intermediaries perform creation/redemption, potentially causing Fund shares to trade at a significant premium or discount to NAV, especially during market disruptions.
Sector Risk *(Sector)*	The risk that the Fund's concentration in technology and financial sectors will make its performance sensitive to adverse developments affecting these industries. This includes intense competition, rapid product obsolescence, reliance on intellectual property rights in the Technology sector, and vulnerability in the Financials sector to extensive government regulation, credit market liquidity issues, volatile interest rates, and loan defaults.

Table 37 – continued from previous page

Risk Topic	Description
Market Risk *(Market)*	The risk that the value of the Fund's investments will fluctuate, sometimes rapidly or unpredictably, due to general market conditions and volatility. Stock markets tend to move in cycles, and price declines can result from factors affecting the overall economy (e.g., interest rate changes, political events, public health crises) or a particular industry or sector. This risk implies that the Fund could lose value over short or extended periods, regardless of the intrinsic quality of the individual securities held.
Management Risk *(Mgmt)*	The risk that the Fund's investment strategy, decision-making, and techniques, including the investment manager's selection of securities or reliance on analytical models, will fail to produce the desired results or cause the Fund to underperform its benchmark or peers. Since the Fund is actively managed, its performance depends heavily on the accuracy of the manager's judgments about markets, interest rates, values, or trends, which may prove incorrect.
Small-Cap Risk *(SmCap)*	The risk that investments in small- and mid-capitalization companies will be subject to greater volatility, less liquidity, and potentially more abrupt price movements than investments in larger, more established companies. Smaller companies generally have less stable operating histories, less predictable earnings, more limited financial and managerial resources, and narrower product lines, making them more vulnerable to adverse business or economic developments.

Table 37 – continued from previous page

Risk Topic	Description
Large-Cap Risk *(LgCap)*	The risk that investments in large-capitalization companies may underperform the broader market or smaller companies during certain market cycles. Larger, more established companies may experience slower growth rates and may be less agile in adapting to changing economic conditions, consumer tastes, or technological innovation compared to smaller, faster-growing firms.

D Appendix Figures

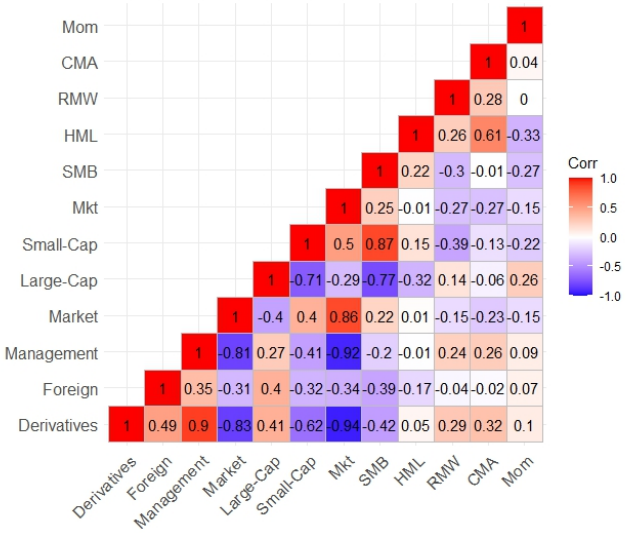


Figure 6: Correlation Matrix for k=5 Textual Factors and FF6 Factors

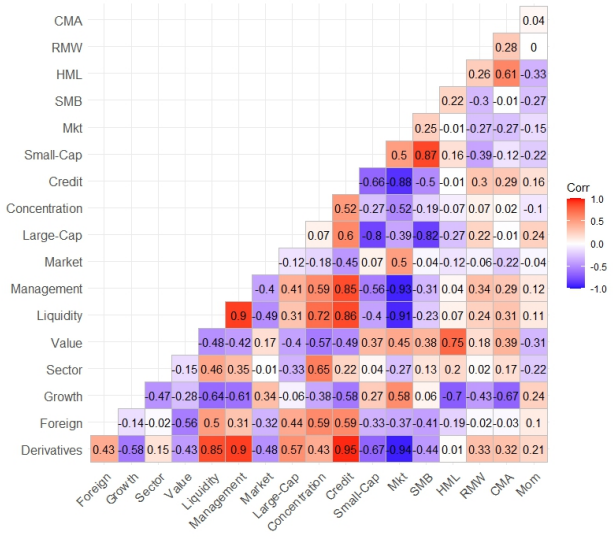


Figure 7: Correlation Matrix for k=10 Textual Factors and FF6 Factors

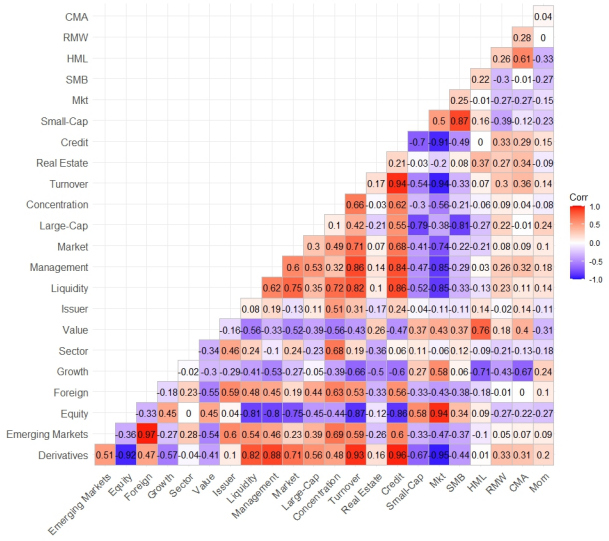


Figure 8: Correlation Matrix for k=15 Textual Factors and FF6 Factors

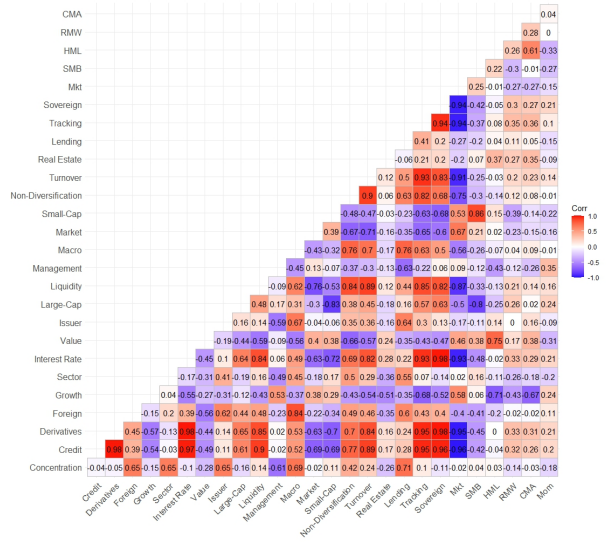


Figure 9: Correlation Matrix for k=20 Textual Factors and FF6 Factors

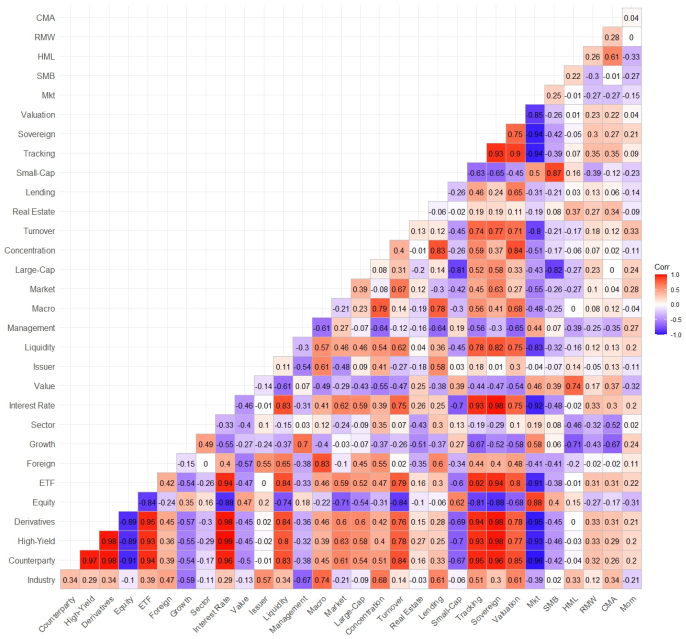


Figure 10: Correlation Matrix for k=25 Textual Factors and FF6 Factors