NEW YORK UNIVERSITY
JOURNAL OF LAW & BUSINESS

VOLUME 22 SUMMER 2026 NUMBER 3

THE FUNCTION AND FORM OF BOILERPLATE IN MUTUAL FUND DISCLOSURES

ANNE TUCKER*
SUSAN NAVARRO SMELCER**
YUSEN XIA***

The mutual fund market—where and how most Americans save for retirement—is governed largely by a disclosure regime administered by the Securities and Exchange Commission (SEC). These disclosures are often dismissed as rote boilerplate, assumed to convey little meaningful information. This Article challenges that view. We argue that mutual fund disclosures, considered in the aggregate, encode valuable information about fund strategy, innovation, and portfolio change. Funds invest substantial resources in drafting disclosure language, and the SEC actively enforces compliance—facts that underscore the informational and legal significance of linguistic choices.

We develop a game-theoretic model of disclosure strategy that explains why funds rely on standardized language to communicate shared legal meanings while selectively deviating from boilerplate to signal strategic change.

* Tucker (JD) is a Professor of Law and the Robert Cotten Alston Chair in Corporate Law at the University of Georgia School of Law, ORCID iD: 0000-0001-9817-230X.

** Smelcer (JD, PhD) is an Assistant Professor, Wake Forest University, Department of Politics and International Affairs; ORCID iD: 0000-0002-0184-0517.

*** Xia (PhD) is the Anne and Michael D. Easterly Distinguished Professor and Director, Institute for Insight at the Georgia State University Robinson College of Business; ORCID iD: 0000-0003-2360-5574.

The authors thank our graduate research assistants, Sahil Dharamup, Joe Griswell, Kellyann Malone, and the students who participated in the 2023 Spring mutual fund lab. We also want to thank readers of the article, including participating faculty at the 2023 Vanderbilt Securities Law Conference, 2024 Winter Deals Conference, 2024 Georgia State University College of Law Workshop, and the 2024 CELS Conference, including our discussant, Reilly Steel, and the 2025 University of Georgia School of Law Faculty Workshop. All mistakes are our own.

We then test the model's predictions using original quantitative analyses of mutual fund disclosures from 2011 to 2022. The evidence reveals both the informational function and form of boilerplate. Funds use boilerplate strategically, particularly in describing investment strategies; disclosure language varies systematically by fund type and exhibits strong clustering within fund families—what we call "house" language. Most importantly, departures from boilerplate language predict economically and statistically significant changes in portfolio composition. Far from being empty gestures at compliance, boilerplate language serves as a low-cost signaling mechanism at the intersection of market behavior and regulatory oversight.

Introduction

Mutual fund disclosures read, unsurprisingly, like the legal documents they are—full of terms of art, legalese, and stock language.[1] They contain information such as past performance, fees, and discussions of the funds' investment strategies and the attendant risks.[2] Readers may be quick to dismiss the disclosures as "boilerplate"—a four-letter word in the law.[3] Calling language boilerplate belittles the drafter's skill and trivializes its informative value.[4] Yet boilerplate plays an undeniable role across practice areas, from complaint templates to standard contract terms and form documents—even before AI-assisted

1. Disclosures, like contracts, can overload investors with information, be excessively long, include obsolete terms, and accumulate legalese, known as "legal encrustation." For a discussion of the dark side of boilerplate, see Stephen Choi et al., *Who Calls the Shots? How Mutual Funds Vote on Director Elections*, 3 Harv. Bus. L. Rev. 35, 40–41 (2013).

2. Wulf A. Kaal & Bentley J. Anderson, *Unconstrained Mutual Funds and Retail Investor Protection*, 36 Rev. Banking & Fin. L. 817, 830–33 (2017) [hereinafter *Unconstrained*]. Mutual funds are subject to additional SEC filing requirements, as well as annual and semiannual financial statements. These regulations are beyond the scope of this Article. *See, e.g.*, Letter from Barry D. Miller, Associate Director of Office of Legal and Disclosure, SEC (July 30, 2010) [hereinafter ICI Letter] (detailing disclosure obligations beyond the prospectus); Enhanced Reporting of Proxy Votes by Registered Management Investment Companies: Reporting of Executive Compensation Votes by Institutional Investment Managers, 17 C.F.R. pt. 200, Release Nos. 33-11131 (Nov. 2022).

3. *Cf.* Marcel Kahan & Michael Klausner, *Standardization and Innovation in Corporate Contracting*, 83 Va. L. Rev. 713, 719–720 (1997) (describing learning benefits and network benefits of boilerplate); Douglas G. Baird, *The Boilerplate Puzzle*, 104 Mich. L. Rev. 933, 949–50 (2006) (discussing the boilerplate in commercial transactions); Gregory Klass, *Boilerplate and Party Intent*, 82 Law & Contemp. Probs. 106 (2019) (describing problems of discerning party intent with boilerplate).

4. For example, the SEC warns operating companies (firms) to "avoid 'boilerplate' risk factors. We believe a discussion of risk in purely generic terms does not tell investors how the risk may affect their investment . . . ." SEC Plain English Disclosures 17 C.F.R. pts. 228, 229, 230, 239 and 274 [Release Nos. 33-7497; 34-39593; IC-23011]. *See, e.g.*, Reid Kress Weisbord & David Horton, *Boilerplate and Default Rules in Wills Law: An Empirical Analysis*, 103 Iowa L. Rev. 663 (2018) (describing boilerplate in wills); Lauren Henry Scholz, *Fiduciary Boilerplate: Locating Fiduciary Relationships in Information Age Consumer Transactions*, 46 J. Corp. L. 143 (2020) (describing consumer boilerplate); Orly Lobel, *Boilerplate Collusion: Clause Aggregation, Antitrust Law & Contract Governance*, 106 Minn. L. Rev. 877 (2021) (describing the role of boilerplate in antitrust).

drafting.[5] No one reinvents the wheel with pay-by-the-hour legal services.

Mutual fund disclosures require a special kind of drafting to satisfy the Securities and Exchange Commission's (SEC) regulatory requirements[6] and meet investors' (and their intermediaries')[7] information needs.[8] Funds must draft compliant disclosures that tell the markets what they do and then follow the rules of their own game. As we show in a regulatory enforcement game, boilerplate can be an attractive drafting strategy for funds, regulators, and investors.[9]

We also empirically examine the role that boilerplate plays in mutual fund disclosures. To do so, we use an original dataset of all mutual fund summary prospectuses filed between 2011

---

5. *See, e.g.*, James M. Lacko & Janis K. Pappalardo, *The Failure and Promise of Mandated Consumer Mortgage Disclosures: Evidence from Qualitative Interviews and a Controlled Experiment with Mortgage Borrowers*, 100 AM. ECON. REV. 516 (2010); Janis K. Pappalardo, *Product Literacy and the Economics of Consumer Protection Policy*, 46 J. CONSUMER AFF. 319 (2012).

6. The Securities and Exchange Commission (SEC) requires information to register and monitor funds through both initial and annual disclosures. 15 U.S.C.A. § 77g (2022). The SEC uses the information to let in new funds, with initial registration, and monitor funds thereafter with annual disclosures. Form N–1A, Registration Statement of Open-End Management Investment Companies, 17 C.F.R. § 239.15A (2021), https://www.sec.gov/files/formn-1a.pdf [hereinafter SEC FORM N-1A]. A mutual fund's prospectus is perhaps the most significant source of information about funds. Every fund must file this document with the SEC to initially register and then annually renew the fund.

7. Intermediaries include broker/dealers and employer-plan sponsors who provide the investment portals through which most individuals invest. *See* Susan Navarro Smelcer, Anne Tucker & Yusen Xia, *Aggregated Risks: Mutual Fund Disclosures & Market Risk*, 49 J. OF CORP. L. 49, 63–64 (2023) [hereinafter *Aggregated Risks*] (discussing plan sponsors); VANGUARD, BEST PRACTICES FOR PLAN FIDUCIARIES 22 (2022) (discussing plan fiduciaries).

8. Margaret Giles, *What to Invest in During High Inflation*, MORNINGSTAR (Jan. 12, 2023), https://www.morningstar.com/markets/what-invest-during-high-inflation (discussing Morningstar investment analysis). *See, e.g.*, Anne M. Tucker, Yusen Xia, *Promise & Peril of Plain English: Mutual Fund Disclosure Readability*, 13 HARV. BUS. L. REV. 59, 71–75 (2023) [hereinafter *Promise & Peril of Plain English*] (discussing ordinary investors' information needs). "Average" investors' relative disinterest in the nitty-gritty of mutual fund disclosures suggests the importance of other audiences for the information. Beyond the SEC, sophisticated intermediaries access, absorb, and act upon disclosed information on the investors' behalf. The benefits of more "investor friendly" disclosures seem to accrue substantially to the regulators and intermediaries, but benefits only indirectly trickle down to individual investors.

9. *See, e.g.*, Jeremy McClane, *Boilerplate and the Impact of Disclosure in Securities Dealmaking*, 72 VAND. L. REV. 191, 233 (2019) (describing incentives to recycle previously approved SEC language).

and 2022 to observe how much boilerplate exists. We define boilerplate as repeated and recycled language across filings and measure it using two similarity measurements that account for word importance (cosine TF/IDF) and word order (cosine 5-gram shingles).[10] We find a lot of boilerplate—on average, around 80% of disclosure language is recycled year-to-year in the narrative sections, such as a fund's Investment Strategy and Principal Risk sections. Consistent with our regulatory model, drafting mutual fund disclosures resembles other areas of the law: to draft tomorrow's disclosure, funds look to the last one filed.

Our findings are expected, consistent with our research experience reading thousands of these documents, and intuited by anyone who has ever attempted to draft or read even a summary prospectus. The puzzle is that the SEC rules expressly prohibit boilerplate in disclosures—or at least too much boilerplate.[11] Of course, not all disclosure language is repetitive.[12] It would be a poor regulatory tool if it were. Funds toe an invisible line, changing some disclosure language year to year but not most.[13]

A central tension emerges between disclosure language stability—boilerplate—on one hand, and change, on the other. Scholars have observed and partially resolved this tension in other legal drafting tasks. Lawyers drafting contracts, for example, may choose boilerplate provisions to reduce drafting costs,[14] facilitate shared meaning with the other party,[15] and increase

10. *See infra* Section II.

11. 17 C.F.R. § 230.421 (2025) (requiring prospectuses to be written in clear, plain English and discouraging vague boilerplate disclosure).

12. *See infra* Section III.

13. It would be odd—and perhaps even alarming—if a fund's description of its principal strategies and resulting risks radically changed from year to year in an investment meant to last a lifetime. Some boilerplate is therefore expected, if not welcomed. As we note in our findings in Section III, principal risk statements change more significantly than investment strategy sections.

14. For example, boilerplate avoids drafting costs by not "reinventing the wheel," and risk-averse lawyers are comforted by language that is road-tested and trusted by the SEC and investors. *See generally*, Jessica Jeffers & Anne Tucker, *Shadow Contracts*, 1 U. Chicago BUS. L. Rev. 269, 282 (2022) (discussing the role of transactional costs in drafting new language).

15. In contracts and other disclosure regimes, the predominance of boilerplate language promotes stability, facilitates shared meaning—especially with complex topics—and signals expectations. David Horton & Reid Kress Weisbord, *Boilerplate No Contest Clauses*, 82 Law & Contemp. Probs. 69, 72–73 (2019).

interpretability by future judges.[16] Lawyers change contract language to tailor the contract to a given deal and avoid too much boilerplate, which can lead to impenetrable, illogical, and inefficient drafting.[17] But drafters and readers alike use boilerplate as a roadmap to construct, read, and understand contracts. There is no working model, however, for mandatory mutual fund disclosures. Our work fills this gap by offering both a theoretical model and novel empirical evidence of how boilerplate works in disclosures.

To begin, we review the regulatory requirements that govern mutual fund disclosure, such as updating disclosure language in response to external events or when investment strategies drift.[18] We then offer a game-theoretic account of disclosure drafting and the incentives of funds, the SEC, and investors.[19] The game highlights the uncertainty that funds, regulators, and investors face. How much information is the right amount? Boilerplate can be a useful drafting strategy for the three players because repeated language reinforces *shared meaning* over time, avoids costly mistakes with terms of art, and reduces transaction drafting costs. Identifying strategic reasons for stability also highlights the inverse: change.

We identify four hypotheses of disclosure drafting that we take to the data: (1) boilerplate advances shared meaning over time, (2) boilerplate reduces transaction costs, but (3) boilerplate gives way to new language when conditions on the ground change, such as new strategies or (4) when the fund changes its portfolio holdings.

---

16. Repeated exposure to contractual boilerplate increases consistency and reliability in judicial interpretation. Henry E. Smith, *Modularity in Contracts: Boilerplate and Information Flow*, 104 MICH. L. REV. 1175, 1188 (2006) (acknowledging that boilerplate facilitates judicial interpretation); Jeremy McClane, *Boilerplate Semantics: Judging Natural Language in Standard Deal Contracts*, 2020 WIS. L. REV. 595, 624–25 (2020) [hereinafter *Boilerplate Semantics*] (acknowledging the role of courts in interpreting boilerplate); John Coyle & W. Mark C. Weidemaier, *Interpreting Contracts Without Context*, 67 AM. U. L. REV. 1673 (2018).

17. Disclosures, like contracts, can overload investors with information, be excessively long, include obsolete terms, and accumulate legalese, known as "legal encrustation." For a discussion of the dark side of boilerplate, see Choi et al., *supra* note 1, at 40–41. *See also* 2011 SEC staff comments to the ARROW INVESTMENTS TRUST 3 (on file with author).

18. Susan Navarro Smelcer, Anne Tucker & Yusen Xia, *Regulating Dynamic Risk in Changing Market Conditions*, 13 WM. & MARY BUS. L. REV. 775, 788–89 (2022).

19. *See infra* Section I.

Our findings support the shared meaning hypothesis. Different types of funds—as indicated by asset class—use boilerplate in unique ways. For example, fixed-income (FI) funds employ a different level of boilerplate, on average, as compared to every other type of fund.[20] We also find that disclosures by funds in the same asset class—that is, funds with similar investment strategies—are more likely to share language than disclosures from different asset classes. This is true of both the investment strategy and risk sections.

Next, evidence of boilerplate's ability to reduce transaction costs can be found in funds' observable use of "house language." Funds within the same family share more language compared to similarly situated funds in other complexes.

We also observe when disclosure language changes. Funds with innovative or evolving strategies, such as environmental, social, and governance (ESG) funds, are more likely to revise their investment strategies annually compared to funds with more established strategies.

Importantly, major changes in disclosure language signal more than mere wordsmithing. We find that major changes to disclosure language signal statistically significant changes in the mix of a fund's portfolio. That is, lower levels of repeated language are correlated with meaningful and statistically significant changes in assets (equity securities, debt securities, cash, and other holdings).

This Article proceeds as follows: Section I describes the regulatory architecture of mutual fund disclosures, introduces a game-theoretic model, and identifies four resulting hypotheses. Section II describes our data and methodology. Section III outlines our main boilerplate findings. Section IV concludes with three use cases for boilerplate disclosure language focusing on regulators, funds, investors, and their intermediaries.

I.
Fund Disclosures & Boilerplate

This Part briefly introduces the regulations governing mutual fund disclosures before introducing a regulatory enforcement game to understand drafting incentives. It concludes

20. Fixed-income funds, which invest in debt and the bond markets, are broadly considered more stable as a form of guaranteed or principal-protecting assets, compared to stocks like domestic equity funds.

with our hypotheses that boilerplate is a drafting strategy that promotes shared meaning and reduces transaction costs. But boilerplate gives way to updated language when investment strategies evolve.

A. *Disclosure Rules*

Mutual funds must file a registration statement (SEC Form N-1A) to form a new open-ended fund whenever the fund materially changes and at least annually thereafter.[21] Funds must also make material updates when investment strategies change and drift,[22] and funds update the prospectus to keep it "accurate in all material respects."[23]

The SEC uses disclosed information (and more) to ensure fund compliance, monitor market risks, prevent fraud, and promote investor confidence.[24] The last goal—investor confidence—and its companion, investor *protection*,[25] have traditionally given the SEC political authority to regulate markets and mandate information from issuers.[26] Admittedly, few individual investors read detailed fund disclosures, but investment

---

21. SEC Form N-1A, *supra* note 6; John McGuire, *Registering Investment Companies under Form N-1A*, Mutual Funds Today: Current Issues and Developments, CW009 ALI-CLE (2014); *Unconstrained*, *supra* note 2, at 830 (*citing to* 15 U.S.C. §§ 80a-1(b)(1)–(3), (5), (7); § 80a-8(b)(1)(E)).

22. Changes to "investment objective and policies" that change a fund's focus "require a post-effective amendment filed under Rule 485(a), rather than a Rule 497 prospectus." Investment Company Act Release No. 19,382 (Apr. 6, 1993).

23. Michael Glazer, *Prospectus Disclosure and Delivery Requirements*, Ch.4, Mutual Funds & Exchange Traded Funds, 4–32 (2015).

24. *See, e.g.*, Geoffrey A. Manne, *The Hydraulic Theory of Disclosure Regulation and Other Costs of Disclosure*, 58 Ala. L. Rev. 473, 479 (2007) (describing the SEC regulatory regime as one primarily built around disclosure).

25. SEC Form N-1A, *supra* note 6, at General Instructions C.1.(b) (naming ordinary investors as an audience). The SEC is concerned with ordinary investors in regulating firms, as well as funds. *See e.g.*, Usha Rodrigues, *Dictation & Delegation*, 92 Ind. L.J. 435, 495-497 (describing the SEC's rulemaking under Title III of the JOBS Act and the focus on protecting ordinary investors).

26. *Promise & Peril of Plain English*, *supra* note 8, at 63. The SEC's political authority to regulate markets is based upon the hypothetical individual investor who reads disclosures. *Id.* at 62–63. But, within a well-regulated market, information *to investors* has diminishing returns beyond a fund's crucial attributes. This is especially true where intermediaries and regulators are on the scene. Crucial information such as security type, general strategy, risk profile, fees, and performance track record (if any) is detailed in the summary prospectus. Perhaps this explains why so few individual investors read disclosures. SEC Form N-1A, *supra* note 6.

intermediaries consume, interpret, and act on disclosed information on the investors' behalf.[27]

The SEC requires three types of disclosures from funds.[28] First, the SEC mandates certain check-the-box disclosures of hard facts like fee tables and the inclusion of mandatory language.[29] These disclosures provide information about voting records, taxes, portfolio holdings, fees, and past performance.[30] Second, the SEC imposes principles-based requirements to list and describe the "principal" investment strategies and risks.[31] Third, the SEC requires additional open-ended disclosures like updating for material strategy changes or shifts in external market conditions.[32]

SEC regulations also dictate the language and style of the disclosure. Funds must produce clear and comprehensive communications[33] written in plain English,[34] and subject to other guidelines around formatting, ordering, and writing.[35]

---

27. McClane, *supra* note 9, at 219–20 (expressing skepticism that individual investors read disclosures and discussing how disclosed information reaches investors, through analytics and other intermediaries). Investment intermediaries and professionals include broker-dealers, plan administrators, and industry analysts like those at Morningstar. *Aggregated Risks*, *supra* note 7, at 63–64.

28. Smelcer, Tucker & Xia, *supra* note 18, at 787; *Aggregated Risks*, *supra* note 7, at 61–62.

29. Smelcer, Tucker & Xia, *supra* note 18, at 787–88.

30. *Unconstrained*, *supra* note 2, at 830–32. Mutual funds are subject to additional SEC filing requirements, as well as annual and semiannual financial statements. These regulations are beyond the scope of this Article.

See also ICI Letter, *supra* note 2; SEC Enhanced Reporting of Proxy Votes by Registered Management Investment Companies: Reporting of Executive Compensation Votes by Institutional Investment Managers, 17 C.F.R. pt. 200, 232, 240, 249, 270 and 274 (2024).

31. Smelcer, Tucker & Xia, *supra* note 18, at 789.

32. *Id.* at 789–90. Our study focuses primarily on the narrative discussions and requirements to update disclosures, as opposed to check-the-box disclosures. The prospectus is a comprehensive, annual, widely available document. Parts are automatically provided to investors. For a discussion of the prospectus, see *id.*, at 790–91; s*ee also Unconstrained*, *supra* note 2, at 832; ICI Letter, *supra* note 2.

33. SEC Form N-1A, *supra* note 4, at General Instructions C.1.(b) (naming ordinary investors as an audience). The SEC is concerned with ordinary investors in regulating firms, as well as funds.

34. SEC Plain English Disclosures, 17 C.F.R. pts. 228, 229, 230, 239 and 274 (1998); *see also Promise & Peril of Plain English*, *supra* note 8, at 70–71.

35. *See, e.g.*, Form and Content of Prospectuses, 17 C.F.R. § 230.421(d)(2)(i)–(vi) (2023) (outlining style guides); *see also* SEC. & EXCH COMM'N, Enhanced Disclosure and New Prospectus Delivery Option for Registered Open-End Management Investment Companies (2009), https://

The SEC justifies its style and content rules in service of better investor comprehension.[36]

Similarly, the SEC cautions funds to avoid "[v]ague boilerplate explanations that are readily subject to differing interpretations."[37] SEC examiners also flag mutual fund filings for the improper use of generic language,[38] laundry lists,[39] or confusing language.[40] SEC enforcement actions attempt to police the overuse of boilerplate.[41] Regulators have issued

---

www.sec.gov/rules/final/2009/33-8998-secg.htm [https://perma.cc/JM2R-R8CG] (describing summary prospectus expectations).

36. *See* Section I.B.

37. SEC Plain English Disclosures, *supra* note 34. The SEC similarly warns operating companies to "avoid 'boilerplate' risk factors. We believe a discussion of risk in purely generic terms does not tell investors how the risk may affect their investment . . . ." *Id.*

38. A 2013 staff comment to the GuideStone Funds filing noted the Staff's concern with registrants providing derivatives disclosure that was either too brief or generic. The Staff stated its view that appropriate derivatives disclosure will be "tailored specifically to how a fund expects to be managed and should address those strategies that the fund expects to be the most important means of achieving its objectives and that it anticipates will have a significant effect on its performance." The disclosure "should describe the purpose that the derivatives are intended to serve in the portfolio . . . and the extent to which derivatives are expected to be used." Funds should avoid enumerating a generic "laundry list" of derivatives. (on file with authors)

39. BrandywineGLOBAL - Global Total Return ETF.pdf (on file with authors)—"The disclosure in the Principal investment strategies" section includes a laundry list of fixed income securities and debt instruments in which the Fund may invest.

40. Under the heading, Principal Investment Strategies, the disclosure states, "The Longview Extended Commodity Index is a composite index of commodity sector returns. This statement may be confusing to some shareholders. How can an index be comprised of sector returns? It may be helpful to provide the components of the index." 2011 SEC staff comments to the ARROW INVESTMENTS TRUST.pdf, p.3 (on file with author).

As in contracts, the overuse of boilerplate erodes consumer comprehension, leading to legal encrustation where the "intelligibility of language deteriorates significantly as legal jargon is added to standard formulations . . . ." Stephen J. Choi, Mitu Gulatti & Robert E. Scott, *The Black Hole Problem in Commercial Boilerplate*, 67 DUKE L.J. 1, 5 (2017).

41. The SEC has tried several times to limit boilerplate recitations in securities disclosures, and in December 2015, Congress took aim at boilerplate by mandating that the SEC revise its disclosure regulations to eliminate such language as much as possible. Fixing America's Surface Transportation Act of 2015 (FAST Act), Pub. L. No. 114-94, § 72002, 129 Stat. 1312, 1784–85 (2015) ("[T]he Securities and Exchange Commission shall take all such actions to revise regulation S-K . . . to eliminate provisions of regulation S-K, required for all issuers, that are duplicative, overlapping, outdated, or unnecessary . . . ."); *see also* Business and Financial Disclosure Required by Regulation S-K, 81 Fed. Reg. 23,915, 23,917 (Apr. 22, 2016) ("We are specifically seeking comment

orders asking some funds to revise "nearly identical language" across disclosures that made it difficult to "distinguish between the investments . . . ."[42]

Some of the SEC prohibitions are followed,[43] like when SEC Examiners reject filings that copy and paste the same language across different funds with different strategies because the boilerplate is noncompliant.[44] But, as we show in Section III, the SEC tolerates extensive boilerplate.

SEC regulations introduce several tensions: language stability and change; compliance and non-compliance; and information for regulators and investors. Resolving these tensions is not obvious from the face of the regulations alone. We turn to a theoretical account to fill in the missing incentives and motivations.

B. *Disclosure in a Regulatory Oversight Game*

We can characterize funds' competing incentives as a game of regulatory oversight. Both the SEC and investors require different levels of disclosed information from the fund. The fund must choose how much and which type of information to disclose.[45] The cost of doing so grows necessarily with the

---

on . . . whether, and if so how, we could revise our requirements to . . . promote efficiency, competition, and capital formation . . . .").

42. SEC Comment Letter, American Funds Retirement Income Portfolio Series, File Nos.: 333-203797; 811-23053 (May 29, 2015) (2015 WL 14055764). If risk statements are found to be "identical," the SEC rejects the filing and demands "fund-specific risk disclosure to permit investors to understand how the risks differ among the Funds." SEC Comment Letter, Unified Series Trust, File Nos. 811-21237; 333-100654 (Nov. 21, 2002) (on file with authors).

43. *See e.g.*, *Promise & Peril of Plain English*, *supra* note 8, at 78 (describing results that funds follow some SEC guidelines, but not all).

44. "The Principal Investment Strategies section for CLS Global Diversified Equity Fund, CLS Growth and Income Fund, and CLS Global Growth Fund each contain three substantially identical paragraphs that seek to describe the Fund's risk level relative to a blended benchmark, the advisor's strategy to control risk, and the advisor's active management strategy." Letter from Thompson Hine LLP to Keith Gregory File Nos. 333-20635 and 811-08037, (Oct. 31, 2014) (2014 WL 13088193).

45. *See, e.g.*, Geoffrey A. Manne, *The Hydraulic Theory of Disclosure Regulation and Other Costs of Disclosure*, 58 Ala. L. Rev. 473, 479 (2007) (SEC regulation depends upon disclosure). A mutual fund's prospectus is perhaps the most significant source of information about funds. Every fund must file this document with the SEC to initially register and then annually renew the fund. Form N–1A, Registration Statement of Open-End Management Investment Companies, 17 C.F.R. § 239.15A (2021).

amount and variety of information disclosed. After receiving the fund's disclosure, both the SEC and investors must decide whether the fund has met their requirements and whether to "punish" the fund.[46]

Imagine that the fund can choose any level of disclosure d with cost c_d.[47] The cost of writing a disclosure is such that, as the level of disclosure increases, so too does the cost.[48] The SEC requires disclosures that reveal some level of information α about the fund's compliance. This preference α encompasses the minimum information required both to inform investors of the relevant risk and allow the SEC to monitor the markets pursuant to its mandate in the 1933, 1934, and 1940 Acts.[49]

---

46. The SEC acts as a third-party arbitrator sitting between the fund and investors, much like a judge in interpreting a contract. *See generally* Russell B. Stevenson, *The SEC and the New Disclosure*, 62 CORN. L. REV. 50 (1976) (describing the SEC's enforcement of disclosure regulations). The SEC reviews, accepts (or rejects), and enforces initial and updated disclosures.

> The Disclosure Review and Accounting Office (DRAO) is primarily responsible for the analysis and review of filings for mutual funds, ETFs, closed-end funds, variable insurance products, unit investment trusts, and similar investment companies under the Federal securities laws. These responsibilities include the review of initial registration statements, post-effective amendments thereto, proxy statements, and financial statements as required by the Sarbanes-Oxley Act of 2002.

Division of Investment Management: Division of Disclosure Review and Accounting Office, SEC, (Apr. 6. 2023), https://www.sec.gov/division-investment-management-about-disclosure-review-and-accounting-office.

47. Funds exercise discretion in writing disclosures, especially in complying with principles-based and open-ended SEC regulations. *See, e.g.*, Tamar Frankel, *The Failure of Investor Protection by Disclosure*, 81 U. CIN. L. REV. 421, 426 (2012). Disclosure drafters face uncertainty about how to comply or what counts when regulations task the disclosing party with interpreting the rules. McClane, *supra* note 9, at 217–218.

48. *See generally* Jessica Jeffers & Anne Tucker, *Shadow Contracts*, 1 U. CHICAGO BUS. L. REV. 269, 276 (2022) (discussing the role of transactional costs in drafting new language); *see also* McClane, *supra* note 9, at 233 (describing temptations for companies and counsel to use market-tested and SEC approved language).

49. Congress granted the Securities and Exchange Commission authority over securities market regulation and disclosure through the Securities Act of 1933, 15 U.S.C. §§ 77a–77aa; the Securities Exchange Act of 1934, 15 U.S.C. §§ 78a–78qq; and the Investment Company Act of 1940, 15 U.S.C. §§ 80a-1–80a-64. "Disclosure has become the main regulatory mechanism for the securities markets. Disclosure offers two main benefits. First, it enables investors to make their decisions on whether and how to trade in securities (third-party promises). Second, disclosure can deter issuers and brokers from misleading investors." Frankel, *supra* note 47, at 426.

Investors, however, require additional information beyond the SEC's minimum requirement α to make informed investment choices.[50] Let us call this higher level of disclosure β, such that $\beta > \alpha$.

After observing the fund's disclosure d, the SEC and investors must decide how to respond. The SEC suffers a cost r if it receives a disclosure that fails to meet its minimum information requirements to regulate and monitor markets. The SEC can avoid r by initiating an enforcement proceeding. But enforcement is costly to both the fund and the SEC, and both parties accrue cost e when an enforcement proceeding is initiated.[51]

50. Scholars have documented retail investors' reliance on past performance and Morningstar Ratings to make investment decisions. *See, e.g.*, Itzhak Ben-David et al., *What Do Mutual Fund Investors Really Care About?*, 35 REV. FIN. STUD. 1723 (2022) (finding evidence that investors chase returns and look to rating agencies); Richard B. Evans & Yang Sun, *Models or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment*, 34 REV. FIN. STUD. 67 (2021) (using changes in Morningstar rating systems to create a natural experiment to observe the positive relationship between investor fund flows and Morningstar ratings); *see also*, Sankar Srinivasan, BEGINNERS GUIDE TO MUTUAL FUNDS (2023); Diane K. Schooley et al., *Generation X: Understanding Their Risk Tolerance and Investment Behavior*, 16 J. FIN. PLANNING 58 (Sept. 2003); *How to Determine Your Risk. Tolerance Level*, CHARLES SCHWAB (Mar. 13, 2026) [hereinafter *Charles Schwab Risk*], https://www.schwab.com/learn/story/how-to-determine-your-risk-tolerance-level.

51. *See* Enforcement Data, on file with author (describing the observations of a manual review of over 100 SEC comments on mutual fund registration statements, which revealed enforcement patterns around boilerplate, vague language, duplicative language, laundry lists, and generic statements). Disclosure defects can be costly—we found enforcement cases ranging from $10–40M. *See* Press Release, SEC, *SEC Charges Investment Adviser and Distributor in Connection with Improper Fees for Mutual Fund Investors* (Sept. 21, 2015), https://www.sec.gov/newsroom/press-releases/2015-198 (announcing a $40 million settlement with First Eagle Investment Management and FEF Distributors for failing to disclose conflicts related to securities lending activities); Press Release, SEC, *SEC Charges Invesco* Advisers for Making Misleading Statements About Supposed Investment Considerations (Nov. 8, 2024), https://www.sec.gov/newsroom/press-releases/2024-179 (announcing a $17.5 million settlement with Invesco Advisers, Inc. for making misleading statements about the percentage of company-wide assets under management that integrated environmental, social, and governance (ESG) factors in investment decisions); Press Release, SEC, *SEC Charges BlackRock Advisors for Failing to Disclose Conflict of Interest* (Apr. 20, 2015), https://www.sec.gov/newsroom/press-releases/2015-71 (announcing a $12 million settlement with BlackRock Advisors, LLC for failing to disclose a conflict involving a former employee).

When the fund loses an enforcement action, it accrues an additional cost c_e to remediate its defective disclosure.[52]

Investors have a different choice to make. Investors, like the SEC, can seek to impose legal liability for a failure to provide sufficient information.[53] But the most important punitive tool available to investors is the ability to exit, which the fund experiences as negative fund flow.[54] Assume that investors suffer some cost i when either (1) the fund has not provided sufficient information, but investors do not exit; or (2) the fund has provided sufficient information,[55] and the investors choose to exit anyway. When the fund's investors choose to exit, the fund experiences a loss of fees f.[56]

---

52. Breaching these obligations or having materially incomplete disclosures can result in enforcement actions or private legal actions. *See, e.g.*, In re Marion Merrell Dow, Inc., Sec. Litig., No. 92-0609-CV-W-6, 1993 U.S. Dist. LEXIS 14197, at *28–30 (W.D. Mo. Oct. 4, 1993) (finding that a drug company's prospectus included warnings that were too generic and failed to specify the magnitude of risk).

53. *Aggregated Risks, supra* note 7, at 113. Breaching these obligations or having materially incomplete disclosures can result in enforcement actions or private legal action. *In re* Marion Merrell Dow, Inc., *supra* note 52, at *28–30 (criticizing a disclosure as too generic).

Funds also face the threat of investor actions. Fund disclosures establish a quasi-contractual relationship with investors. Securities Act of 1933, 15 U.S.C. § 77k. Misrepresentations and omissions in registration statements filed for a registered public offering are actionable under Section 11. *See, e.g.*, Omnicare, Inc. v. Laborers Dist. Council Constr. Indus. Pension Fund, 135 S. Ct. 1318, 1323 (2015) (alleging that registration statement contained an untrue statement of material fact or omitted a material fact under section 11).

Section 12(a) addresses misrepresentations in a prospectus or oral communication with investors. 15 U.S.C. § 77l.

Investors can sue if a fund deviates from its stated strategies, fails to disclose primary risks, or omits crucial information about fees, rights, and relationships with other financial intermediaries. *See, e.g.*, *In re* Countrywide Fin. Corp. Sec. Litig., 588 F. Supp. 2d 1132, 1162 (C.D. Cal. 2008) (referencing 15 U.S.C. § 77k(e) for damages under Section 11 liability); *see also In re* MetLife Demutualization Litig., 624 F. Supp. 2d 232, 271 (E.D.N.Y. 2009) (referencing 15 U.S.C. § 77*l*(a)(2) for section 12 damages).

54. *See, e.g.*, Anne M. Tucker, *Locked In: The Competitive Disadvantage of Citizen Shareholders*, 25 YALE L.J. FORUM 163, 172 (2015) (describing investor exit costs and constraints).

55. Investors have insufficient information when $d < \beta$, and sufficient information when $\beta \leq$ d.

56. Keith C. Brown et al., *Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry*, 51 J. FIN. 85, 88 (1996) (describing the relationship between the number of fund investors and mutual fund managers compensation); *see also* Eric D. Roiter, *Disentangling Mutual Fund Governance from Corporate Governance*, 6 HARV. BUS. L. REV. 1

But investors are not the only actors watching these disclosures. Other funds are also monitoring their competitors' disclosures. Disclosures revealing more specific information about the funds' strategy allow others to more easily copy those strategies, all else being equal.[57]

This revelation can be costly.[58] These "copycats" may entice investors away from the fund, resulting in a negative fund flow. We conceptualize the cost of copycatting as γf—some proportion (γ) of the total fees earned by the fund (γ).[59] We also assume that γ increases with disclosure quality d, such that more fulsome disclosures result in higher levels of copycatting and negative fund flow. Functionally, this means that funds will always experience some negative fund flow, *even when investors are satisfied with the fund's disclosure.*

This configuration of incentives forces funds to balance the cost of writing a fulsome disclosure with the possibility of enforcement by the SEC, exit by investors, and the possibility of copycatting.[60] The fund will choose to write a disclosure d^+ that

---

(2016) (explaining the relationship between negative fund flows and loss of advisor fees).

57. Disclosure risks revealing private, proprietary information such as research and skill, that once disclosed, can be copied. This phenomenon has been documented in the hedge fund industry. *See* Zhewei Shi, *The Impact of Portfolio Disclosure on Hedge Fund Performance*, 126 J. FIN. ECON. 36, 36–53 (2017) (showing that hedge fund performance declines and returns correlated with peers' increases after the initiation of 13F filings, suggesting disclosure erodes proprietary advantages); Vikas Agarwal et al., *Uncovering Hedge Fund Skill from the Portfolio Holdings They Hide*, 68 J. FIN. 739, 739–83 (2013) (finding that hedge funds' confidential holdings outperform their disclosed ones). The literature has been extended to mutual fund portfolios. *See* Thomas C. Hagenberg, *Transaction-Level Transparency and Portfolio Mimicking*, 79 J. ACCT. & ECON. 101713 (2025) (finding that enhanced transparency facilitates portfolio mimicking, especially among smaller insurers copying more sophisticated peers); Russ Wermers et al., *Predicting Stock Performance with Mutual Fund Portfolio Disclosure*, Univ. of MD., Smith Sch. of Bus. (2008) (demonstrating that disclosed mutual fund holdings contain predictive information).

58. Zhaoyang Chen et al., *Testing the Effect of Portfolio Holdings Disclosure in an Environment Absent of Mandatory Disclosure*, 57 ACCTS. & FIN. 701, 701–23 (2017) (showing that more frequent disclosure reduces alpha due to front-running and copycat behavior); Vikas Agarwal et al., *Mandatory Portfolio Disclosure, Stock Liquidity, and Mutual Fund Performance*, 70 J. FIN. 2733, 2733–76 (2015) (finding that more frequent disclosure reduces performance by revealing proprietary strategies and weakening incentives for information gathering).

59. By definition, $\gamma \in [0,1]$.

60. Funds balance SEC mandates to make complete disclosures (or risk liability) with practical market concerns, hoping not to scare away potential investors. *Aggregated Risks*, *supra* note 7, at 61.

satisfies both the SEC and investors when enforcement costs and the potential for negative fund flows are sufficiently large as compared to the cost of writing the disclosures.[61]

When the possibility of losing fees is relatively low (or the fees themselves are relatively low) as compared with the cost of writing d^+, the fund will write a less informative disclosure.[62] The fund will choose between writing a compliant disclosure $d^{"}$[63] with cost $c_d^{"}$, and a noncompliant disclosure $d^{'}$ with cost $c_d^{'}$.

The SEC will choose to enforce when the fund's disclosure is noncompliant ($d^{'}$) and will not enforce otherwise. Knowing this, the fund will comply with the SEC's regulations when the cost of doing so is less than the cost of a successful enforcement action.[64] When this condition does not hold, however, the fund will choose to submit a non-compliant disclosure.[65]

To this point, we have assumed a game characterized by complete information where the SEC can perfectly interpret the signal sent by the fund through its disclosure level d, and where funds perfectly understand the SEC's minimum compliance requirements (α). But this is likely not the case. The SEC's preference for compliant disclosures (α) in real life is achieved, in part, by requiring that funds disclose information through principles-based and open-ended discussions.[66] Funds must exercise judgment and discretion over what information is, in fact, required by the SEC and investors when drafting

---

61. In other words, the fund will produce $\alpha < \beta \leq d^+$ with cost $c_{d^+}^+$. From a technical perspective, this is satisfied when the difference between writing a disclosure that satisfies only the SEC ($d^{"}$ with cost $c_d^{"}$) and writing a disclosure that satisfies both investors and the SEC (d^+ with cost c_d^+) is sufficiently small, such that $f(1-\gamma^+) > c_d^+ - c_d^{"}$ and $c_d^+ - c_d^{'} < r + e + f(1-\gamma^+)$.

62. We can express this condition formally as $f(1-\gamma^+) \leq c_d^+ - c_d^{"}$.

63. A compliant disclosure is expressed as $\alpha \leq d^{"} < \beta$.

64. Formally, this can be expressed as $e + r > c_d^+ - c_{d'}^{'}$.

65. The full proof for this unique, pure-strategy equilibrium is included in Appendix B.

66. For a discussion of principles-based and open-ended disclosures, which require narrative responses, compared to check-the-box responses, such as stating a tax rate, *see Changing Market Conditions*, *supra* note 18, at 785–932. Narrative discussions force issuers to share in the regulatory burden of disclosure by interpreting what information is required, and complying with evolving standards as opposed to clear, disclosure requirements that require the regulator to define the disclosure universe ex-ante and bear the sole burden of determining what information is relevant.

the "principal" strategy and risk discussions.[67] In addition, the fund's disclosure d may include concepts or strategies that are unfamiliar to the SEC or investors and, as a result, risk misinterpretation.[68]

1. *Boilerplate Promotes Shared Meaning*

Boilerplate can serve a particularly useful function given the costs and uncertainties of drafting compliant disclosures, as demonstrated in the model. First, boilerplate can lower the costs of writing disclosures because repeated language can transmit *shared meaning*. Legally significant language repeated in disclosures over time can reliably express complex legal and financial concepts to SEC examiners and investors.[69]

SEC examiners, like judges, benefit from repeated exposure to similar language and disclosure structures when evaluating whether disclosures are compliant.[70] The SEC examiner builds knowledge and experience with the meaning and effects of repeated language. Once the SEC accepts a formulation of *derivative investments* or *emerging market risks*, that language, when it reappears in whole or part, is easier for the examiner to comprehend and approve.

Not all fund investment strategies are equal. Some are more complex than others, and if so, the task of writing compliant disclosures is most costly and ripe for boilerplate. A parallel can be drawn between complex contracts—known for inviting more boilerplate—and complex investment strategies like

67. *Id.* (discussing different SEC disclosure requirements).

68. SEC review suggests a stamp of approval. Stephen J. Choi et al., *supra* note 40, at 4–5 (2017). Recycling approved language presents a cheap way to effectively communicate strategies and mitigate regulatory and litigation risks. Introducing new language or novel concepts likely incurs additional costs. *Id.*

69. McClane, *supra* note 9, at 275 (discussing boilerplate and investor comprehension).

70. Repeated exposure to contractual boilerplate increases consistency and reliability in judicial interpretation. "[J]udges can more easily afford to be legal generalists . . . when a great deal of contractual language they have to deal with is modular boilerplate." Smith, *supra* note 16, at 1188. *Cf. Boilerplate Semantics*, *supra* note 16, at 624–25 (cautioning that when courts see meaning where individuals see none, the language is susceptible to many legitimate readings and boilerplate can erode consent and let differences in intent/meaning persist through bad, boilerplate drafting).

derivatives and hedging.[71] Consider Guggenheim Multi-hedge Strategies Fund, which repeated derivative strategy language verbatim in the 2021 and 2024 summary prospectuses.[72] The principal investment strategy discussion was, in total, over 1,100 words describing sophisticated strategies like *merger arbitrage*, *global macro investments*, and *long/short equity positions*.[73] Contrast Guggenheim's investment strategy discussion with the short and pithy under-100-word descriptions of the S&P index funds (described below).[74] We posit that recycling boilerplate descriptions of complex strategies efficiently communicates hard-to-distill financial concepts into concrete language or at least formulations familiar to investors and the SEC. Standardized descriptions of investment strategies can also help obscure proprietary information, reducing the risk of copycat behavior by competitors.

---

71. For a discussion of complex investment strategies like derivatives, see Kelly S. Kibbie, *Dancing with the Derivatives Devil: Mutual Funds' Dangerous Liaison with Complex Investment Contracts and the Forgotten Lessons of 1940*, 9 HASTINGS BUS. L.J. 195, 200 (2013).

72. Both prospectuses include the following language:

> Investments in derivative instruments, such as futures, options, and forward contracts and swap agreements, have the economic effect of creating financial leverage in the Fund's portfolio because such investments may give rise to losses that exceed the amount the Fund has invested in those instruments. Financial leverage will magnify, sometimes significantly, the Fund's exposure to any increase or decrease in prices associated with a particular reference asset resulting in increased volatility in the value of the Fund's portfolio. The value of the Fund's portfolio is likely to experience greater volatility over short-term periods. While such financial leverage has the potential to produce greater gains, it also may result in greater losses, which in some cases may cause the Fund to liquidate other portfolio investments at a loss to comply with limits on leverage and asset segregation requirements imposed by the Investment Company Act of 1940 (the "1940 Act") or to meet redemption requests. The Fund's use of credit risks. The Fund may use leverage to the extent permitted by applicable law by entering into borrowing transactions (principally lines of credit) for investment purposes.

Guggenheim Funds Summary Prospectus, April 30, 2021, available at https://www.sec.gov/Archives/edgar/data/899148/000119312521144436/d156314d497k.htm [hereinafter Guggenheim Prospectus dated April 30, 2021]; Guggenheim Funds Summary Prospectus, May 1, 2024 [hereinafter Guggenheim Prospectus dated May 1, 2024].

73. Guggenheim Prospectus dated April 30, 2021, *supra* note 72; Guggenheim Prospectus dated May 1, 2024, *supra* note 72.

74. *See infra* notes 78–79 and accompanying text.

Some disclosure continuity from year to year, especially related to strategy,[75] is a byproduct of the SEC's regulatory design.[76] Funds operate based on stated principal investment strategies that remain mostly stable over time.[77] If Fund A is an S&P 500 index fund, its investment strategy description likely identifies the index and describes related risks from tracking errors and stock market volatility in similar (or possibly even identical) ways over time.[78] Vanguard's S&P 500 Index fund, for example, uses the same principal investment strategy statement from 2021 through 2024.[79] This repeated *boilerplate* language is a feature, not a bug. If the strategy is stable and consistent, there is neither a regulatory nor a fund incentive to change the language.

Later, when disclosures are out in the world, familiar language helps intermediaries and investors confidently comprehend and navigate the investment landscape.[80] If an investor wants an index fund versus a growth fund or an emerging

---

75. SEC Form N-1A, *supra* note 6 at Item 4: Principal Strategy.

76. *See* discussion *infra* Section III (finding significant year-to-year language stability in filings). The SEC disclosure rules do not anticipate entirely new disclosures every year for the same fund, nor should there be substantive differences between similarly situated funds.

77. SEC Form N-1A, *supra* note 6 at Item 4(a): Principal Investment Strategies of the Fund; *see also infra* Part III.A.1 Table 4 (establishing that funds recycle 85% or more of language between years).

78. For example, Vanguard's S&P 500 index fund states:

> The Fund employs an indexing investment approach designed to track the performance of the S&P 500 Index (the Index), a widely recognized benchmark of U.S. stock market performance that is dominated by the stocks of large U.S. companies. The Fund attempts to replicate the target index by investing all, or substantially all, of its assets in the stocks that make up the Index, holding each stock in approximately the same proportion as its weighting in the Index.

Vanguard, Vanguard 500 Index Fund Summary Prospectus (2024).

79. The 2021 principal investment strategy language states:

> The Fund employs an indexing investment approach designed to track the performance of the Standard & Poor's 500 Index, a widely recognized benchmark of U.S. stock market performance that is dominated by the stocks of large U.S. companies. The Fund attempts to replicate the target index by investing all, or substantially all, of its assets in the stocks that make up the Index, holding each stock in approximately the same proportion as its weighting in the Index.

Vanguard, Vanguard 500 Index Fund Summary Prospectus (2021), https://www.sec.gov/Archives/edgar/data/36405/000168386321002805/f8919d1.htm.

80. McClane, *supra* note 9, at 275 (discussing boilerplate and investor comprehension).

market fund, it is easier to shop when investments are described in familiar terms.[81]

Consider similar alternative S&P 500 funds offered by other complexes. These funds share investment strategies, and those similarities are evident in the language. S&P 500 funds' common language includes describing their "normal" investment practice of investing "80% of assets in the index"[82] or "all or substantially all assets,"[83] seeking to "replicate the performance of the index" subject to weighting,[84] and describing the index as comprised of "common stocks publicly traded in the United States."[85]

Funds also invest in the same external markets alongside other funds, which means that risks affecting one fund likely affect others.[86] For example, S&P 500 funds disclose stock market risks,[87] risks associated with replicating an index,[88] and risks of investing in large capitalization companies.[89] To this

---

81. Contracts scholars recognize that boilerplate facilitates shared meaning among parties to a deal. This is especially important as deals become more complex. Smith, *supra* note 16, at 1176; *see, e.g.*, Erik F. Gerding, *Contract as Pattern Language*, 88 WASH. L. REV. 1323, 1326 (2013) (describing complex transactions as being drafted by piecing together language from smaller standardized transactions).

82. FIDELITY, FIDELITY S&P 500 INDEX FUND APRIL 29, 2025, Summary Prospectus (2025), https://fundresearch.fidelity.com/mutual-funds/summary/315911750 [hereinafter Fidelity S&P 500]; CHARLES SCHWAB, SUMMARY PROSPECTUS, FEBRUARY 26, 2026, SCHWAB S&P 500 INDEX FUND [hereinafter Schwab S&P 500], https://connect.rightprospectus.com/Schwab/TVT/808509855/SP?site=FundDocs.

83. VANGUARD, VANGUARD S&P 500 INDEX FUND, SUMMARY PROSPECTUS, APRIL 29, 2025 (2025), AVAILABLE AT https://www.vanguard.com/pub/Pdf/p040.pdf [hereinafter Vanguard S&P 2025 Index].

84. Schwab S&P 500, *supra* note 82, at 1; Vanguard S&P 2025 Index, *supra* note 83, at 2.

85. Vanguard S&P 2025 Index, *supra* note 83, at 1; Fidelity S&P 500, *supra* note 80, at 3.

86. *Aggregated Risks*, *supra* note 7, at 64–66.

87. Vanguard S&P 2025 Index, *supra* note 83, at 8 (highlighting the chance that stock prices overall will decline. Stock markets tend to move in cycles, with periods of rising prices and periods of falling prices. The Fund's target index tracks a subset of the U.S. stock market, which could cause the Fund to perform differently from the overall stock market. In addition, the Fund's target index may, at times, become focused in stocks of a particular market sector, which would subject the Fund to proportionately higher exposure to the risks of that sector); *see also* Schwab S&P 500, *supra* note 82, at 2.

88. Vanguard S&P 2025 Index, *supra* note 83, at 3.

89. "Investment style risk, which is the chance that returns from large-capitalization stocks will trail returns from the overall stock market. Large-cap stocks tend to go through cycles of doing better—or worse—than other

end, prior empirical studies have found that unrelated funds, across a range of investment strategies, used similar language and structures to describe external market risks like spiking inflation.[90]

These shared meanings can evolve into terms of art that provide flexibility.[91] Contract drafters, for example, intentionally introduce ambiguity and flexibility into contract language to provide breathing room for future performance in the face of unknowns.[92] Provisions regarding reasonable conduct, professional standards, timely delivery, or ordinary course of business are recycled boilerplate terms that deliberately incorporate flexibility.[93] Ex-post flexibility through recycled boilerplate benefits parties who prefer to delay defining fixed standards and judges who must interpret contracts in the context of messy, real-world facts.

Funds' legal teams are similarly constrained in their knowledge about the future state of the world in which the fund will operate. Funds face uncertainty about the information

---

segments of the stock market or the stock market in general. These periods have, in the past, lasted for as long as several years." Vanguard S&P 2025 Index, *supra* note 83, at 3.

90. *Aggregated Risks*, *supra* note 7, at 55; *id.* at 83 (quoting Am. Beacon Fund, Prospectus Materials (Form N-1A) 5 (Feb. 28, 2022)).

91. Parties drafting a contract have incomplete information about future states of the world (prices, markets, scarcity, and performance) and mitigate some of that uncertainty with incomplete contract terms and flexible language. *See, e.g.*, Albert Choi & George Triantis, *Strategic Vagueness in Contract Design: The Case of Corporate Acquisitions*, 119 YALE L.J. 848, 881–96 (2010) (describing efficiencies from contractual vagueness and incompleteness); Albert Choi & George Triantis, *Completing Contracts in the Shadow of Costly Verification*, 37 J. LEGAL STUD. 503, 517–20 (2008) (analyzing contractual litigation costs as a factor in drafting).

92. *See, e.g.*, *Boilerplate Semantics*, *supra* note 16, at 625 (linking some boilerplate writing to "the need to draft relatively broadly favors the use of opaque constructions, in which different imagined states of the world can be accommodated").

93. These phrases are legally acceptable and enforceable despite their vagueness and indefiniteness. *See, e.g.*, Robert E. Scott & George Triantis, *Anticipating Litigation in Contract Design*, 115 YALE L.J. 814, 844 (discussing implied contractual standards. Such language will not destroy the contract's enforceability, invite egregious litigation, or scare parties off by an unfinished deal); Kroboth v. Brent, 625 N.Y.S.2d 748, 749 (3d Dep't 1995) (describing the enforceability of best efforts); *see also* Coady Corp. v. Toyota Motor Distribs., 361 F.3d 50, 59 (1st Cir. 2004) (describing reasonableness); Nat'l Data Payment Sys., Inc. v. Meridian Bank, 212 F.3d 849, 854 (3d Cir. 2000) (describing diligence); T.S.I. Holdings, Inc. v. Jenkins, 924 P.2d 1239, 1250 (Kan. 1996) (describing reasonable diligence).

required by the SEC and investors. Disclosure drafters, like transactional attorneys, mitigate uncertainty (to some degree) with flexible strategy[94] and risk language.[95] Funds need flexibility to respond to new risks and invest in evolving markets, while avoiding the risk of sharing too much and inadvertently revealing proprietary information. But investors need to know enough about the fund to invest, and the SEC needs to know enough to effectively monitor it.[96] When drafters find the goldilocks of disclosure language that is SEC compliant and satisfies investors, without sharing too much, funds have incentives to continue its use.[97]

This result can be readily observed. Prior empirical studies have found that unrelated funds dramatically increased climate-change-related disclosures in 2020.[98] Many opted for a generic formulation like the following:

> The value of the Fund's investments may go up or down due to general market conditions that are not specifically related to the particular issuer, such as real or perceived adverse economic conditions, changes in the general outlook for revenues or corporate earnings, changes in interest or currency rates, regional or global instability, *natural or environmental disasters* . . . .[99]

In fact, among the funds that discussed the climate change-related risks in 2020, nearly 30% used a similar

---

94. The uncertainty may arise in how an asset manager will allocate the portfolio, which itself turns on unknown market conditions and opportunities. For a discussion on uncertainty and risks in mutual funds, see *Aggregated Risks*, *supra* note 7, at 66–70.

95. "Uncertainty may arise from unknown risks, like a pandemic or inflationary spikes. The nature of uncertainty implies that funds may not fully understand what risks are possible. That is, funds cannot fully account for all tail events. . . . Especially severe public health events, such as the dramatic onset of COVID-19, constitute uncertainties for which funds could not adequately assess or disclose risk prior to its onset." *Id.* at 69.

96. For a discussion of disclosures, aided by boilerplate, providing information to regulators, investors and the market writ large, see McClane, *supra* note 9, at 216–17.

97. *Changing Market Conditions*, *supra* note 18, at 849.

98. *Aggregated Risks*, *supra* note 7, at 91 (noting that climate change disclosures increased over time, and dramatically in 2020).

99. *Id.* at 95 (quoting AIM Sector Funds (Invesco Sector Funds), Prospectus Materials (Form N-1A) 2 (Aug. 26, 2021)) (emphasis added).

approach.[100] One fund's artful, flexible articulation of a risk that is acceptable to the SEC will be quickly followed in spirit, if not the letter, by other funds.[101]

2. *Boilerplate Reduces Transaction Costs*

Beyond increasing understanding, boilerplate can also reduce transaction costs by lessening the burden of writing disclosures year after year.[102] Previously accepted language—either its own or its competitor's—reduces a fund's uncertainty about complaint disclosures (α) and acts as a safe harbor.[103] Recycling SEC-approved language presents a cheap way to effectively communicate strategies and mitigate regulatory and investor risks.

Boilerplate can reduce the overall cost of disclosure (c_d) while increasing the amount of information transmitted. This means that funds are more likely to produce disclosures that satisfy both the SEC and investors. As in other areas of the law, why reinvent the wheel? Reusing language—at least to a certain point—allows funds to write higher quality disclosures at a lower cost.

---

100. Unrealized risks receive more generic treatment. In the face of materialized risks that have come to fruition, funds opt for focused and specific statements, like in 2022, when 65% of funds made COVID-specific or focused risk disclosures in the waning global pandemic. *Id.* at 87.

An example of a focused climate change disclosure is:

> Certain issuers, industries, and regions may be adversely affected by the impacts of climate change, including on the demand for and the development of goods and services and related production costs, and the impacts of legislation, regulation and international accords related to climate change, as well as any indirect consequences of regulation or business trends driven by climate change.

Id. at 90.

101. *See, e.g.*, *id.* at 64–66 (discussing how funds learn from other funds, primarily through disclosures).

102. Disclosure drafters may be uncertain how to comply or what counts when regulations task the disclosing party with interpreting the rules. McClane, *supra* note 9, at 217–218; Last year's language and competitors' existing filings create a compliance benchmark: such language has been approved (if tacitly) by the SEC. Recycling previously accepted language and using familiar phrasing or boilerplate aids in compliance. *Id.* at 196 (discussing boilerplate in terms of compliance and "because there are only so many ways to say certain things").

103. Stephen Choi et al., *Who Calls the Shots? How Mutual Funds Vote on Director Elections*, 3 HARV. BUS. L. REV. 35, 43–46 (2013) (describing the review process by examiners and the uncertainty generated by such reviews).

As in contracts, recycling language with shared meaning reduces drafting costs.[104] While disclosures are not negotiated between the fund and an investor,[105] funds write disclosures through layers of internal review from asset management, compliance, legal, and outside counsel.[106] Transaction costs to change a fund's disclosure, particularly the strategy section, involve legal time (in-house) and costs (outside counsel).[107] Independent directors for funds ultimately approve disclosure filings and add yet another layer of review and cost.[108] Substantive changes to disclosures are summarized and reviewed by board members, and sometimes their counsel.[109] Recycling familiar language likely greases the wheels of this multi-tiered internal review process.

Boilerplate also serves a fund's internal operational efficiency. Internal fund actors, charged with implementing the disclosed information, may also prefer repeated, boilerplate language. Consider, for example, the job of a fund complex's compliance officer (like a Vanguard compliance officer). Unchanged disclosures leave working compliance systems

---

104. Contractual boilerplate serves practical purposes, such as reducing transaction costs by avoiding the need to reinvent the wheel. For a discussion of boilerplate and transaction costs see Kahan & Klausner, *supra* note 3, at 718. *See also* Choi et al., *supra* note 103, at 4–5.

105. Contractual boilerplate commonly occurs in arms-length negotiations between sophisticated parties, often negotiated by individuals on behalf of a larger organization. Horton & Weisbord, *supra* note 15, at 72–73.

106.

> The overly complex and often difficult to interpret rules, releases, and guidance often require the assistance of outside counsel. Even as a rule or interpretation evolves or is revised it is still often necessary to go back to outside counsel to ensure that such changes have not created conflicts or unintended consequences for the Firm. The Firm has had to obtain the services of a third-party vendor in order to fulfill its AML responsibilities. The Firm has also had to devote considerable time, money, and human resources to working with its partn company's AMI, unit and audit/compliance groups to educate them in the differences between bank and broker-dealer responsibilities under the various rules and regulations.

The Cost of Compliance: 2013 KPMG/AIMA/MFA Global Hedge Fund Survey, KPMG Int'l (2013) [hereinafter Cost of Compliance], https://assets.kpmg.com/content/dam/kpmg/pdf/2013/10/the-cost-of-compliance-v2.pdf.

107. *Id.*

108. Michael Glazer, *Prospectus Disclosure and Delivery Requirements*, Mutual Funds & Exchange Traded Funds, 4-28 (2015).

109. Timothy D. Lytton & Anne M. Tucker, *The Hidden Work of Disclosures*, 104 Oregon L. Rev. (forthcoming 2026), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6667859.

intact without incurring the costs of updating them.[110] Further, using the same disclosure language across similar funds in Vanguard increases the officer's confidence that the compliance test in place, matches the disclosure language that the right test is in place.[111]

House language for similar strategies and risk would likewise lighten the loads of marketing teams, outside counsel, and independent directors reviewing disclosures.[112] Portability of language from one fund to another is a feature of disclosure boilerplate, analogous in function to contractual boilerplate.[113]

3. *Boilerplate Has Lower Utility When a Fund's Strategy Changes*

The regulatory game specified above suggests incentives for language stability, given both the SEC's uncertainty about interpreting disclosures (d) and the fund's uncertainty about minimum compliance requirements (α). In the game above, boilerplate is preferable to funds, the SEC, and investors when investment strategies remain the same year-to-year, when describing common risks or strategies, and when markets are stable.[114]

But when conditions on the ground change, boilerplate is no longer a smart strategy. Using boilerplate may reduce drafting costs, but repeated language may fail to adequately convey the fund's current investment strategies and risks, exposing the fund to potential SEC enforcement and investor exit.[115]

When boilerplate language no longer reflects the funds' current strategy and risks, funds have an incentive to draft new language. In other words, funds incur the full cost of drafting the disclosure—a cost that was previously reduced by using boilerplate language. Prior empirical work has demonstrated, for example, that funds respond to changing market conditions

---

110. *Cost of Compliance*, *supra* note 106, at 25.

111. For a discussion on operational economies of scale and interdependence in fund operations within the same fund complex or family, see John Morley, *The Separation of Funds and Managers: A Theory of Investment Fund Structure and Regulation*, 123 YALE L. J. 1228, 1259 (2014).

112. Glazer, *supra* note 108, at 4-28.

113. Boilerplate also includes copied and pasted language from one agreement to another, especially in standardized form contracts used in business-to-business transactions. Horton & Weisbord, *supra* note 15, at 72–73.

114. *See supra* notes 42-45, and accompanying text.

115. In the model, we expressed this as $\alpha < \beta \leq d^{+}$ with cost c_d^{+}.

(e.g., COVID-19 and inflation spikes) with new and updated disclosure language.[116]

C. *Boilerplate Hypotheses*

We empirically test four hypotheses to evaluate our theoretical and practical understanding of boilerplate and language change. If boilerplate plays a meaningful role in creating flexibility and promoting a shared meaning, similarly situated funds should exhibit more similar language, all else being equal. In particular, funds with similar investment styles should be more alike in their use of boilerplate than those with dissimilar investment styles. For example, boilerplate used by Index (I) funds should differ from boilerplate used by Domestic Equity (DE) funds, DE funds should differ from Foreign Equity (FE) funds, and so on.

(1) Shared Meaning Hypothesis: Funds with similar investment styles will exhibit more similar language than disclosures by funds with dissimilar investment styles.

But boilerplate also has value in reducing operational and transaction costs. Reusing language makes the act of complying with disclosure regulations less costly. As a result, we expect that funds in the same fund family (that is, funds managed by the same investment company) would resemble each other more than funds outside of their fund family, all else being equal. For example, we expect two DE funds in the Vanguard complex to share more language compared to other DE funds in other complexes, like Fidelity and BlackRock.

(2) Transaction Cost Hypothesis: Funds within the same fund family should share more language than funds outside of their family.

When investment strategies change, however, so too should funds' use of boilerplate. Boilerplate can predictably provide

116. *Aggregated Risks*, *supra* note 7, at 77 (describing dynamic inflation disclosures) and 85 (describing dynamic COVID-19 disclosures). In additional unpublished research, the authors tracked how funds incorporated new investment strategies and risk statements related to artificial intelligence, finding increases in strategy disclosures (up to 72 unique statements in 2023) and risk disclosures (up to 36 unique risk statements). Notes on file with authors.

operational flexibility only when the underlying investment practices themselves are stable. When practices evolve, disclosure language must balance accuracy and flexibility, but boilerplate may not effectively accomplish this.

We argue that disclosure language changes are most likely when (1) funds are executing an emerging or evolving strategy (such as a strategy emphasizing innovation-related or ESG assets), and (2) when a fund changes the mix of assets in its investment portfolio. Each of these events mutes the beneficial aspects of boilerplate language.

(3) Emerging/Evolving Strategy Hypothesis: Strategy disclosures for funds pursuing emerging or evolving investment strategies will contain less boilerplate than fund disclosures with more traditional investment strategies.

(4) Portfolio Change Hypothesis: Strategy disclosures made following portfolio changes—that is, changes in the mix of equity securities, debt securities, and cash—will exhibit less boilerplate than funds with a stable distribution of portfolio assets.

Section II describes our data and details our methodology. Using these data, we examine each of these hypotheses in Section III. Section IV discusses three use cases for our methodology, focusing on the SEC, internal fund audiences, and investors and their intermediaries.

II.
Data & Methodology

In this Section, we turn to the empirical examination of boilerplate and change in mutual fund disclosures. We conceptualize boilerplate as repetitive language in disclosures that carries over from year to year.[117] We then propose two different

117. Repetitive language suggests text that can function across changing economic conditions and investment strategies—operating, much like contractual boilerplate, as a template that retains meaning through repeated use. Measuring boilerplate using year-to-year text similarity measures is particularly appropriate here, since general language is stable enough to track over time. Scholars working with cross-sectional data, where individual units appear only once, similarly rely on text similarity measures such as cosine similarity. *See, e.g.*, McClane, supra note 9, at 220–28.

methods of measuring similarity across disclosures using a dataset of mutual fund disclosure summary prospectuses to assess our hypotheses. We find, as a baseline matter (and unsurprisingly), that mutual fund disclosures exhibit a high degree of boilerplate.[118] We use empirical confirmation of this aspect of disclosures as a jumping-off point for evaluating our hypotheses in Part III.

A. *Mutual Fund Disclosure Data 2011-2022*

We use the SEC's Mutual Fund Prospectus Risk/Return Summary Data Sets to access the Investment Strategy and Principal Risks of all registered, open-end investment companies between 2011 and 2022.[119] We begin with a dataset of 203,540 disclosures spread over our 12-year period of interest. Our unit of observation is the fund-year, as identified by a filing's accession number.[120]

These 203,540 disclosures include two types of duplicates. First, the data contains submissions for multiple series of the same fund. We consider these to be duplicates because different series of the same fund will have identical investment strategies and risks; only fees differ.[121] As a result, the prospectuses of different series of the same fund will be identical. Retaining filings for each series would not only create meaningless duplication

---

118. Our results are complementary to earlier empirical findings of "considerable boilerplate" in registration statements (on average 47%). McClane, *supra* note 9, at 229.

119. We access SEC Summary Prospectus Data sets (2010-2024) available at SEC, *Mutual Fund Prospectus Risk/Return Summary Data Sets*, https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets (last visited July 31, 2026); *see also* SEC, *Mutual Fund Prospectus Risk/Return Summary Data Sets*, https://www.sec.gov/dera/data/rr1.pdf (last visited July 31, 2026). The data set includes fund filings using SEC form 497, which contains the definitive materials filed under paragraph (a), (b), (c), (d), (e) or (f) of Securities Act Rule 497 (prospectus) and 485BPOS, which is the Post-effective amendment filed pursuant to Securities Act Rule 485(b). *Id.* at 2. Amendments filed under 485(b) are effective immediately and contain non-material changes. Amendments filed under 485(a), in contrast, are material amendment and require additional SEC review. Glazer, *supra* note 108, at 4-30-4-32.

120. The accession number comprises the Central Index Key (CIK) uniquely identifying the entity submitting the filing, the year filed, and "the sequential count of filings submitted from that CIK." SEC, *Accessing EDGAR Data*, https://www.sec.gov/os/accessing-edgar-data (last visited July 6, 2023).

121. SEC, *Introduction to Investing: Mutual Fund Classes*, https://www.investor.gov/introduction-investing/investing-basics/glossary/mutual-fund-classes (last visited July 6, 2023).

but would also bias our results toward larger funds. Second, funds may update their prospectuses throughout the year. The SEC requires a fund to update its prospectus in the event of a material change in its investment strategy or principal risks.[122]

Removing these duplicates reduces the total number of observations to 144,619. We then further remove duplicates of filings with identical accession numbers.[123] This results in 31,870 observations.

Finally, we remove funds that file a prospectus in only one year. We need more than one year of data to make year-to-year comparisons.[124] This leaves 23,828 disclosures of funds' principal investment strategies and risks filed between 2011 and 2022. We document our data loss in Table 1.

TABLE 1: DATA LOSS RESULTING FROM CLEANING STEPS

Data Processing Step	*Number of Observations*
Initial number of observations	203,540
Removing multiple series of the same fund and keeping only the most recent filing	144,619
Removing duplicate accession numbers	31,870
Removing series with only one year of data	23,828

122. Securities Act of 1933, 15 U.S.C. § 10(a)(3); 17 C.F.R. § 230.496; *see also Importance of Delivering Timely and Material Information to Investment Company Investors*, SEC: DIV. OF INVESTMENT MGMT. (Apr. 14, 2020), https://www.sec.gov/investment/delivering-timely-material-information (explaining that updating prospectuses with material information "is particularly relevant for investment companies that continuously offer their shares, such as open-ended funds").

123. The count drops precipitously because risk and strategy contents are included in several formal filings—shareholder reports, full prospectus, and summary prospectus. Dropping duplicates across the field keep us from double counting the same language which would skew the similarly scores towards even more boilerplate.

124. Measuring boilerplate in this way excludes funds short-lived funds filing disclosures only in a single year. We have no empirical or theoretical reason to believe that such funds would use boilerplate in a systematically different way than other types of funds. To the extent that such short-lived funds have distinct investment strategies, these funds might use "standard" or "general" language differently than funds designed to exist over multiple years. Such funds, however, lie outside our empirical definition of boilerplate.

B. *Using Text Similarity Measures to Observe Boilerplate*

To study boilerplate in mutual fund disclosures, we construct series-year dyads to calculate year-over-year similarity between sequentially filed disclosures.[125] For example, we pair the most recent and complete disclosures of BNY Mellon Large Cap Securities Fund from 2011 and 2012, 2012 and 2013, 2013 and 2014, and so on.

We calculate pairwise similarity as the cosine similarity based on (1) paired term frequency-inverse document frequency (TF-IDF) matrices; and (2) 5-gram shingles.[126]

Cosine similarity methods can be used to determine the similarity of two different vectors or sets.[127] It measures distance as the cosine angle between the two vectors and is calculated as the dot product of the vectors divided by the magnitude:

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}$$

This method requires that a vector of tokens be converted to numerical data. As a result, cosine similarity provides a good solution for comparing both the term frequency and the TF-IDF vectors of two texts. Term frequency denotes the number of

125. *GSU Mutual Fund Research*, GEORGIA STATE UNIVERSITY, *https://sites.google.com/view/gsu-mutual-fund-research* (authors' research website hosting replication code and research protocols).

126. Our findings are robust to different measurements of this concept. As a sensitivity check, we also conduct our analysis using Jaccard similarity and two different token lengths: unigram (bag-of-words) and 5-gram shingles. Jaccard similarity can be used with a bag-of-words approach using either unigrams (which does not preserve context and word order) or with any size of n-gram shingle (which does preserve context and word order). This is especially useful for discrete, qualitative data. But Jaccard similarity is ill-suited for comparing quantitative data with continuous measurements. Unlike cosine similarity, Jaccard similarity takes a set-based approach, calculated as the intersection between two texts, A and B, divided by the union of the texts. In other words, Jaccard similarity is the proportion of tokens two texts have in common. As a result, Jaccard similarity is less flexible than cosine similarity measures. Results using the Jaccard measures reflected the patterns and relationships in the cosine scores. For simplicity, we report only the cosine scores. For replication code, including the Jaccard scores, see *id.*

127. We apply these methods to text, but they can be used to measure commonalities between any two collections of objects, numbers, or words.

times a unique token appears in the text. TF-IDF, however, weights term frequency by the number of documents in which the term appears, providing a measure of term frequency in the context of the corpus as a whole. We calculate TF-IDF as

$$n_{t|d} * \log\left(\frac{n_d}{n_{t|d}}\right)$$

where $n_{t|d}$ is the number of tokens t in a particular disclosure, d, n_d is the number of disclosures in the corpus, and $n_{d|t}$ is the number of disclosures d containing the token t.

As a robustness check, we conduct our analysis using two different constructions of disclosure tokens. TF-IDF utilizes a bag-of-words approach using unigrams. Word order is not preserved, and similarity is calculated according to weighted term frequency. A term's weight is determined by how frequently it occurs in a document as compared to all other documents.

Our other method preserves context lost by the bag-of-words approach by employing 5-gram shingles.[128] We first split each disclosure into sentences. Then, for each sentence, we capture each sequence of five words using a sliding window. For example, a fund may disclose the following: "The value of your investment may fluctuate significantly." This sentence would produce the following 5-gram shingles: "the-value-of-your-investment," "value-of-your-investment-may," "of-your-investment-may-fluctuate," and "your-investment-may-fluctuate-significantly." This approach allows us to account not only for changed tokens but also for changed *meanings*.

A summary and comparison of these methods is provided in *Table 2*.

128. We split each string into sequences of five consecutive words (so-called "5-gram shingles"), shifting the window by 1 word each time. In a last step, we compute the similarity score as the cosine similarity between the two sets of shingles for each pair of contracts. Because the shingling preserves the order in which words appear, this process is designed to capture textual similarity as a form of copying and pasting. *See, e.g.*, Adam Badawi et al., *The Value of M&A Drafting*, Duke L.Sch. Pub. L. & Legal Theory Ser. No. 2023-14, at 14 (Jan. 25, 2023).

TABLE 2: SUMMARY AND COMPARISON OF TEXT SIMILARITY MEASURES USED

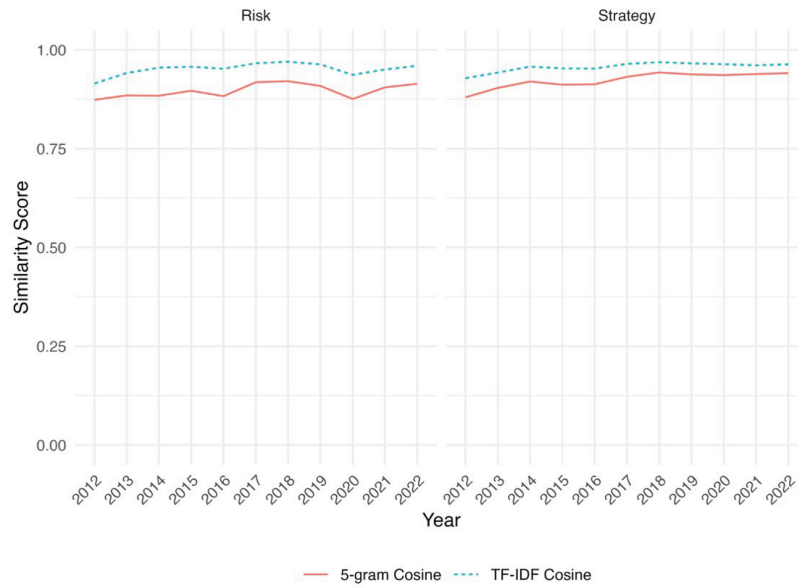
Method	*Input*	*Comparison*	*Discussion*
TF/IDF cosine	Frequency of each word (token) and weighted term frequencies	Measures the cosine distance between the texts based on the relative weights of each token	Highlights important word changes as opposed to minor syntactical changes
5-gram shingle cosine	Frequency of 5-word (token) increments	Measures the cosine distance between texts based on shingles	Accounts for word order and provides context

As a preliminary matter, we find that, on average, mutual fund disclosures have a high degree of year-to-year similarity. Table 3 displays average text similarity scores using both measures for the strategy and risk sections, with 0 indicating no common words and 1 indicating perfectly identical text. The TF-IDF/cosine similarity score—indicated by the blue dashed line—produces the most conservative average measures of similarity. This result is consistent with *how* each measure is used to conceptualize similarity. The TF-IDF/cosine approach weighs words that are unique to a disclosure more heavily, meaning distinctive features of each fund, like its name, have a greater influence on the score. Intuitively, any truly distinctive elements of the fund's investment strategy or risk profile should stay relatively similar across time, so this approach best reflects funds' stability over time.

The 5-gram/cosine measure, by contrast, shows less variability over time. This measure calculates similarity using 5-gram token frequencies, meaning it captures textual overlap across a high-dimensional space defined by every five-word sequence in each disclosure. This measure is sensitive to local phrase-level patterns rather than individual word counts. As a result, it is more robust to minor textual changes: substituting a synonym or reordering a clause will disrupt far fewer 5-grams than it would unigrams. This stability holds even when the underlying texts are relatively short, where sparser data might otherwise produce erratic similarity scores. The result is a measure that

tracks genuine structural repetition in language rather than superficial lexical overlap.[129]

FIGURE 1: COMPARISON OF TEXT SIMILARITY MEASURES ACROSS TIME (2011-2022) AND BY DISCLOSURE SECTION



Comparing funds' risk and strategy disclosures reveals that, for both score types (TF-IDF/cosine and 5-gram/cosine), risk sections exhibit *less year-to-year similarity* than strategy sections.[130] This result intuitively makes sense. The risks associated with

129. Figure 1 also illustrates that disclosure language in both the risk and strategy sections increase in similarity over time—although the growth in repetitive language does not consistently increase over the period of study.

130. We conducted a series of t-tests comparing our text similarity measures for a fund's risk section to text similarity measures for a fund's strategy section by year. Across the 44 t-tests conducted (four measures times 11 years per measure), the difference between the (lower) risk section text similarity score and the (higher) strategy section score was statistically significant in 33 (75%) of our t-tests. In the other 11, we found no statistical difference between the risk and strategy section scores. T-tests comparing scores for risk and strategy sections for all years indicate that the differences between lower average similarity scores for risk disclosures and higher similarity scores for strategy sections are also statistically significant when pooled across all years. For replication code, see *GSU Mutual Fund Research, supra* note 125.

these strategies are a function of external events beyond the funds' control. In other words, risk sections reveal changes resulting from both internal events like altered investment strategies, which are relatively rare, and external events like COVID-19 or inflation spikes. Risk sections, as a result, tend to be revised more frequently than strategy sections.

TABLE 3: SUMMARY STATISTICS FOR TEXT SIMILARITY SCORES BY DISCLOSURE SECTION FOR ALL DISCLOSURES, ALL YEARS

Panel A: Investment Strategy							
	S.D.	*Min*	*25%*	*Mean*	*Median*	*75%*	*Max*
TF-IDF/ cosine	0.1014	0	0.9605	0.9563	0.9943	1	1
5-gram/ cosine	0.1406	0	0.9137	0.9233	0.9719	1	1

Panel B: Principal Risk							
	S.D.	*Min*	*25%*	*Mean*	*Median*	*75%*	*Max*
TF-IDF/ cosine	0.0922	0.0208	0.9477	0.9518	0.9839	0.9973	1
5-gram/ cosine	0.1338	0	0.8664	0.8968	0.9347	0.9805	1

Strategy sections are more stable by design. Investment strategy descriptions—where funds tell investors *how* their money will be invested—should remain relatively constant over time. That is, mutual funds are more likely to make incremental changes in investment strategy than radical ones.

While strategy sections are more stable,[131] funds appear to revise their risk disclosures more consistently—perhaps in

131. We conducted a series of t-tests comparing our text similarity measures for a fund's risk section to text similarity measures for a fund's strategy section by year. Across the 22 t-tests conducted (two measures times 11 years, per measure), the difference between the (lower) risk section text similarity score and the (higher) strategy section score was statistically significant in 13 (59.1%) of our t-tests. In the other nine, we found no statistical difference between the risk and strategy section scores. T-tests comparing scores for risk and strategy sections for all years indicate that the differences between lower average similarity scores for risk disclosures and higher similarity scores for

response to external events. Funds' revisions to their strategy disclosures are more idiosyncratic. Take, for example, funds' responses to COVID-19. In 2020, funds revised their risk disclosures *en masse* in response to this external shock (see Figure 1).[132] But we do not see this behavior mirrored in funds' strategy disclosures. Rather, funds revised their strategies at roughly the same rate as before.

III.
Findings: Boilerplate & Change

In this Section, we use our text similarity measures to evaluate our four hypotheses. We begin with our hypotheses about boilerplate's use in promoting shared meaning and lowering transaction costs. We find that funds holding similar assets tend to contain different levels of boilerplate than other types of funds. We also find that disclosures within fund families are much more likely to share language than disclosures across different fund families. Both findings suggest that boilerplate is widespread—but that boilerplate serves a purpose in creating shared meaning and reducing transaction costs.

We then evaluate our hypotheses about language change. If boilerplate plays a role in developing shared meaning and promotes efficiency, then deviations from boilerplate should be meaningful. Funds executing emerging or evolving investment strategies, such as innovation-related or ESG strategies, are less likely to repeat language from year to year than other types of funds. We also find that changes in a fund's investment strategy disclosure are correlated with changes in the distribution of asset types held by the fund. In other words, the data suggest that changes in fund assets are also reflected in new (not repetitive) language.

A. *Stability in Disclosure Language*

1. *Boilerplate Patterns by Fund Type*

We first evaluate our *Shared Meaning* hypothesis: Funds with similar investment styles will exhibit similar levels of boilerplate.

---

strategy sections are also statistically significant when pooled across all years. For replication code, see *GSU Mutual Fund Research*, *supra* note 125.

132. *Aggregated Risks*, *supra* note 7, at 83–89.

We began with the intuition that some types of funds employ more stable strategies (i.e., bond funds) than others (i.e., domestic equity funds). We expected that funds with more stable strategies, like money markets and fixed-income funds, would have higher similarity scores compared to domestic and foreign equity funds.

We assign asset classes using the Center for Research in Securities Prices (CRSP) categories.[133] We group funds into six overarching asset classes using the CRSP codes: domestic equity ("DE"),[134] foreign equity ("FE"),[135] Fixed Income ("FI"),[136] Index ("I"), Money Market ("M"),[137] and Other ("O").[138] We then categorized funds by merging our prospectus data with the CRSP Survivor-Bias-Free U.S. Mutual Fund Database investment objective codes.[139]

Figure 2 displays the distribution of asset class types across our data. DE funds are the largest fund class in our sample, comprising 32% of our observations. FE funds, on the other hand, comprise only 14%. FI funds (including bonds) account for 10.5%, followed closely by M (8.9%) and I (8.4%) funds. O

133. The Center for Research in Securities Prices began as an academic research endeavor with the University of Chicago Booth, which was acquired by Morningstar in 2026, and provides information about public securities, enabling research on firms and funds. CTR. FOR RSCH. IN SEC. PRICES, https://www.crsp.org (last visited July 31, 2026).

134. "Domestic equity" refers to funds that primarily hold stocks from domestic (U.S.) firms. CTR. FOR RSCH. IN SEC. PRICES, *CRSP Style Code*, https://www.crsp.org/products/documentation/crsp-style-code-0 (last visited Jan. 3, 2023) [hereinafter CRSP Style Code].

135. "Foreign equity" refers to funds that primarily hold stock from foreign (non-U.S.) firms. *Id.*

136. "Fixed income" funds invest in municipal, corporate, and U.S. federal bonds. *Id.*

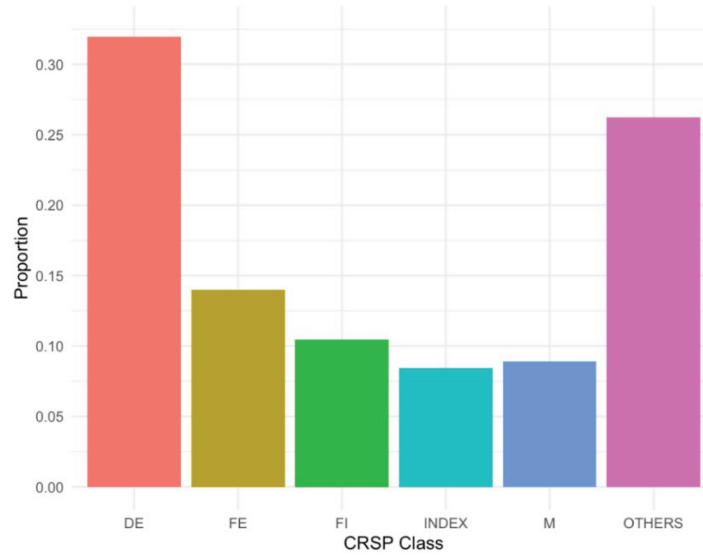
137. "Money Market" funds are "are generally defined throughout the world as regulated funds that are restricted to holding short-term, high-quality debt instruments." 2022 Investment Company Fact Book: A Review of Trends and Activities in the Investment Company Industry, INV. CO. INST., (2022), https://www.ici.org/doc-server/pdf%3A2022_factbook.pdf.

138. Other funds in the CRSP classification include mortgage-backed securities, currency funds and miscellaneous strategies like alternative credit, energy MLP funds, and multi-strategy funds. CRSP Style Code, *supra* note 134.

139. Note the CRSP database excludes closed-end funds and other registered investment vehicles that are not mutual funds, bonds, international equities, or money market funds. *See* CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND GUIDE FOR SAS, ASCII & R, CTR. FOR RSCH. IN SEC. PRICES (CRSP) (Feb. 12, 2026), https://www.crsp.org/products/documentation/survivor-bias-free-us-mutual-fund-guide-sas-ascii-r.

funds account for 26.2% of the sample. We dropped uncategorizable funds from the analysis.

FIGURE 2: DISTRIBUTION (I.E., PROPORTION) OF CRSP ASSET CLASSES IN DATA, ALL FUNDS FOR ALL YEARS (2011-2022)



Similarity scores are high across all fund types and both measures—most over an 85% match of the text between years (see Table 4). Although the similarity scores are high across year, measure, and fund type, they do exhibit year-to-year variation (see Figures 3 and 4).

FIGURE 3: INVESTMENT STRATEGY SIMILARITY SCORES BY ASSET CLASS (CRSP), BOTH MEASURES FOR ALL YEARS (2011-2022)

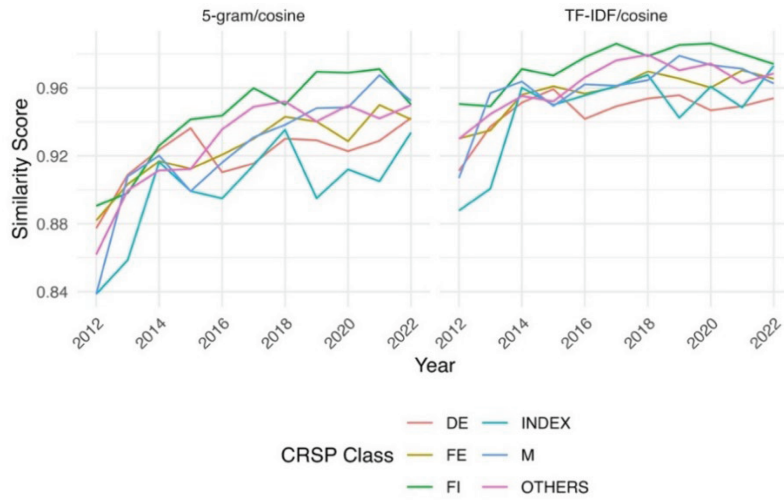


FIGURE 4: PRINCIPAL RISKS SIMILARITY SCORES BY ASSET CLASS (CRSP), BOTH MEASURES FOR ALL YEARS (2011-2022)

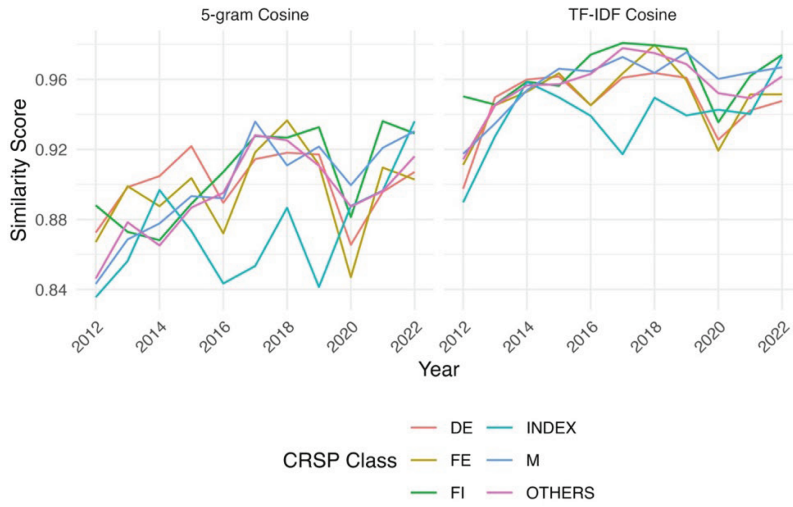


Table 4: Investment Strategy Similarity Mean Scores by Asset Class, both Measures for all Years Combined (2011-2022)

	DE	*FE*	*FI*	*Index*	*M*	*Other*
TF-IDF/cosine	0.94	0.95	0.97	0.94	0.95	0.96
Five-gram/cosine	0.92	0.92	0.94	0.90	0.92	0.92

Table 5: Principal Risks Similarity Mean Scores by Asset Class, both Measures for all Years Combined (2011-2022)

	DE	*FE*	*FI*	*Index*	*M*	*Other*
TF-IDF/cosine	0.94	0.94	0.96	0.93	0.95	0.95
Five-gram/cosine	0.90	0.89	0.90	0.87	0.89	0.89

To test the *Shared Meaning* hypothesis, we first compare the average amount of recycled language in each asset class (by disclosure section) with the average amount of recycled language in *other* asset classes. We use difference-in-means tests (t-tests) to determine whether mean text similarity scores systematically vary across asset classes. We conducted 90 t-tests in total.[140]

First, we find some support for the shared meaning hypothesis—boilerplate patterns emerge when similarity scores are viewed by fund type. Using the keyword-weighted TF-IDF/cosine measure, boilerplate in one asset class differs (in a statistically significant way) from boilerplate in the paired asset class for 60% of strategy sections and 66.7% of risk sections.[141]

For example, disclosures by fixed-income (FI) funds employ a different level of boilerplate, on average, as compared to every other type of fund.[142] FI funds, which invest in debt

140. The t-tests are done for each asset class pair and section. In total, we conducted 90 t-tests: 15 total pairs of asset classes (e.g., DE/FE, DE/FI, etc.) times two text similarity scores (TF-IDF/cosine and 5-gram/cosine) times three different disclosure subsets (strategy, risk, and combined).

141. We observed similar (though somewhat weaker) results using our 5-gram/cosine measure: 10 of 15 (66.7%) of strategy asset class pairs and 6 of 15 (40.0%) of risk asset class pairs demonstrated statistically significant differences.

142. This is true for both the TF-IDF and 5-gram cosine measures. Results from difference-in-means tests as follows. In t-tests using 5-gram/cosine, comparisons of FI with each other asset class type (producing five t-tests) produced t-statistics ranging from 2.88 to 6.54 ($p < 0.01$). Difference-in-means

and the bond markets, are broadly considered more stable as a form of guaranteed or principal-protecting assets, compared to stocks like DE funds.[143] Our finding reflects this reality: FI funds' discussions of strategy exhibit higher similarity and recycled language due to the investment approach.

Other boilerplate patterns by asset classes are not so intuitive. For example, I funds had statistically significantly *lower* similarity scores than DE funds and FE funds using 5-gram/cosine.[144] Yet we expected I funds' investment strategy disclosures—like FI strategy disclosures—to be *more* stable than DE and FE strategies.

Some results, like M funds, were just noisier.[145] Risk section results performed similarly to strategy section scores across the fund categories, but the differences were muted.[146]

---

tests using TF-IDF/cosine measures produced slightly stronger results, with t-statistics ranging from 3.47 to 5.88 ($p < 0.01$).

143. For a discussion of fixed income assets see Jeffrey T. Dinwoodie, *Fixed Income Securities and SEC Rule 15c2-11: History, Context, Uncertainties-and A Pathway Forward*, 41 YALE J. ON REG. BULL. 29, 33 (2023) (describing the differences between fixed income securities and equity securities); *Bonds - FAQ*, INVESTOR.GOV: U.S. SEC. AND EXCH. COMM'N, https://www.investor.gov/introduction-investing/investing-basics/investment-products/bonds-or-fixed-income-products/bonds (last visited July 31, 2026) (describing fixed income products and risk exposure); *Money Market Funds*, INVESTOR.GOV: U.S. SEC. AND EXCH. COMM'N, https://www.investor.gov/introduction-investing/investing-basics/investment-products/mutual-funds-and-exchange-traded-5 (last visited July 31, 2026); *Changing Market Conditions*, *supra* note 18, at 864.

144. Results from difference-in-means tests using 5-gram/cosine were as follows. A comparison of Index and DE similarity scores produced a t-statistic of 2.87 ($p < 0.01$); Index and FE, a t-statistic of 3.208 ($p < 0.01$). Difference-in-means tests using TF-IDF/cosine were not statistically significant: Index versus DE produced a t-statistic of 1.363 ($p = 0.173$); Index v. FE, a t-statistic of 1.524 ($p = 0.128$). Shadow indexing, where active funds mimic the allocations within indices can shrink the distance between DE and Index funds. Similarly, indexed products have been used by fund complexes to attract new money and millennial investors, perhaps causing more attention to be placed on distinguishing index funds with unique language. *See, e.g.*, Michal Barzuza et al., *Shareholder Value (s): Index Fund ESG Activism and The New Millennial Corporate Governance*, 93 S. CAL. L. REV. 1243, 1253 (2020).

145. Money market funds' investment strategy sections, for example, exhibited slightly less change than DE funds (TF-IDF/Cosine), more compared to Index funds (5-gram/cosine), and consistently less change than FI funds (both measures). All difference-in-means tests described here demonstrated statistically significant differences between asset class types with a t-statistic of at least 2.73 ($p < 0.01$).

146. Fixed income funds had higher similarity scores than index funds. *See* Table 5.

Different types of funds use boilerplate at different rates, and this provides initial support for our intuition about boilerplate's function to advance shared meaning. To take the inquiry one step further, we examine whether two disclosures of the same asset class are more similar to each other than a disclosure is to the disclosure of another asset class. For example, we would expect two DE disclosures to be more similar to each other than a DE disclosure and an FI disclosure are.

We test this aspect of the shared meaning hypothesis by randomly selecting 500 disclosures in each asset class-year and calculating the pairwise similarity across all possible pairs (i.e., "in-class" pairs). We repeated this exercise for a random sample of 500 disclosures from different asset classes (i.e., "out-of-class" pairs). For each year, we used t-tests to compare "in-class" and "out-of-class" text similarity scores. If the shared meaning hypothesis is correct, we should find that "in-class" pairs have higher text similarity scores than "out-of-class" pairs. In total, we assessed similarity of both investment strategy and risk sections across 12 years (2011-2022) for each of the six asset classes (DE, FE, FI, M, I, and O).[147] We also examined the difference between in-class and out-of-class pairs pooled across all years.

We find strong support for our shared meaning hypothesis when looking across all years. In-class disclosure pairs are more similar than out-of-class disclosure pairs using both TF-IDF/cosine and 5-gram shingle/cosine similarity measures. This is true for both strategy and risk sections (see Table 6).

147. For each comparison, we used pairwise comparisons of 500 randomly sampled "in-class" disclosures and 500 randomly sampled "out-of-class" disclosures. For both groups, N = (500*500-500)/2 = 124,750.

TABLE 6: IN-CLASS AND OUT-OF-CLASS TEXT SIMILARITY COMPARISONS, BY ASSET CLASS AND POOLED ACROSS 2011-2022

	Strategy		*Risk*	
	TFIDF/cosine	*5-gram shingle/cosine*	*TF-IDF/cosine*	*5-gram shingle/cosine*
DE	(0.589, 0.3479) t = 32.1712, p = 0.0000	(0.0110, 0.008) t = 23.4644 p = 0.0000	(0.6926, 0.6886) t = 13.0000 p = 0.0000	(0.0261, 0.0225) t = 18.4975 p = 0.0000
FE	(0.3521, 0.3415) t = 30.9339, p = 0.0000	(0.0101, 0.0071) t = 23.9629 p = 0.0000	(0.6807, 0.6715) t = 28.5618 p = 0.0000	(0.0255, 0.0225) t = 16.3865 p = 0.0000
FI	(0.3536, 0.3465) t = 21.2146, p = 0.0000	(0.0091, 0.0069) t = 20.3948 p = 0.0000	(0.6817, 0.6781) t = 11.6896 p = 0.0000	(0.0244, 0.0235) t = 15.8108 p = 0.0000
INDEX	(0.3496, 0.3431) t = 19.3271, p = 0.0000	(0.0108, 0.0090) t = 15.7111 p = 0.0000	(0.6720, 0.6710) t = 4.4245 p = 0.0000	(0.0255, 0.0227) t = 17.4139 p = 0.0000
M	(0.3384, 0.3296) t = 38.4811, p = 0.0000	(0.0116, 0.0090) t = 22.2003 p = 0.0000	(0.6656, 0.6578) t = 25.5259 p = 0.0000	(0.0260, 0.0227) t = 17.4139 p = 0.0000
OTHERS	(0.3408, 0.3281) t = 38.4811, p = 0.0000	(0.0101, 0.0074) t = 21.4136 p = 0.0000	(0.6673, 0.6619) t = 17.8229 p = 0.0000	(0.0253, 0.0221) t = 16.9300 p = 0.0000

Note: The first number in each cell is the *in-class* mean; the second is the *out-of-class* mean. The differences in means were calculated using a two-tailed t-test.

We see similar results when examining in-class versus out-of-class similarity by year, although our results vary slightly by asset class, disclosure section, and similarity measure.[148] We find the most support for the shared meaning hypothesis in funds' strategy disclosures. This is unsurprising. Funds are exposed

148. A full table of results is available from the authors upon request.

to some of the same external risks—the impact of COVID, for example—regardless of asset class. Strategies, on the other hand, should have elements unique to the funds' asset class. Funds are likely to include more language about investing in U.S. securities in a DE fund disclosure than in an FI fund disclosure.

In each of the 72 tests of similarity using 5-gram shingles, in-class disclosures are more similar to each other than out-of-class disclosures. Of these 72, 70 are statistically significant at $p < 0.05$ using a two-tailed t-test. Using TF-IDF, in-class disclosures show greater similarity in 61 of 72 comparisons (84.7%); all 61 are statistically significant at $p < 0.05$.

Our analysis of risk disclosures produces a similar result. As with strategy disclosures, we found a stronger relationship among in-class risk disclosures when using 5-gram shingles. The in-class similarity was higher in all 72 comparisons using the 5-gram shingle measure and statistically significant in 70 (97.2%). When using TF-IDF, the in-class similarity was greater in 53 of the 72 tests (73.6%) and statistically significant in the expected direction in 51 (70.8%).

Taken together, these results are strongly suggestive—though not definitive—that recycled language convey shared meaning. Stronger results using the 5-gram shingle measure indicate the presence of identical language across disclosures that, on average, share asset classes and year but no other common features.

2. *Fund Families & Boilerplate's House Language*

We next turn to our *Transaction Cost* hypothesis: Funds within the same fund family should share more language than funds outside of their family. The data provide strong support for the idea that within fund families, legal teams develop family-specific boilerplate, or "house language," for both strategy and risk disclosures. House language is evidenced by statistically significant higher similarity scores between funds within the fund family compared to funds in random pairings.

Our findings reflect a market reality. While each fund is a legally distinct entity,[149] most funds operate as part of a complex

149. Mutual funds are structured legally as a corporation or trust. Jill E. Fisch, *Rethinking the Regulation of Securities Intermediaries*, 158 U. PA. L. REV. 1961, 1967 (2010).

or "family" of funds managed by a single fund sponsor.[150] This is more intuitive than it may seem at first blush. Most are likely familiar with large fund families such as BlackRock, Fidelity, or State Street[151]—funds that have brand recognition and are a "household" name. Funds within these fund families often share operational functions such as administration, auditing and reporting, and legal.[152] Researchers have documented that individual funds in large fund families are managed as a part of a broader portfolio of funds rather than on an individual basis.[153]

Does coordination in management and operations translate to coordinated disclosure language? In other words, will the disclosures of BlackRock funds have similar language by virtue of being part of the same family? To examine this intuition, we grouped individual funds by fund family. To group funds, we leveraged the fund names and business addresses to match individual funds to the same sponsor.[154] In our data, we found 208 fund groupings with at least two funds.[155] The number of funds in each grouping ranged from 2 to 105, with a mean of 8.33. But the distribution is skewed by the large number of families with few funds (the left tail) and by the small number of families that have many affiliated funds.[156] To illustrate this

---

150. At year-end 2022, 812 fund sponsors operated in the U.S. market. INV. CO. INST., 2023 INVESTMENT COMPANY FACT BOOK: A REVIEW OF TRENDS AND ACTIVITIES IN THE INVESTMENT COMPANY INDUSTRY 26 (2023) [hereinafter 2023 ICI Factbook], https://www.ici.org/system/files/2023-05/2023-factbook.pdf; "The concentration of mutual fund and ETF assets managed by the largest fund complexes has increased over time. The share of assets managed by the five largest firms rose from 35 percent at year-end 2005 to 55 percent at year-end 2022." *Id.* at 27.

Common fund sponsors include advisory firms, banks, insurance companies, or brokerage companies—all known as fund sponsors. Choi, *supra* note 17, at 40.

151. Barzuza, *supra* note 144, at 1253.

152. *See, e.g.*, 2023 ICI Factbook, *supra* note 150, at 28 (describing arrangements to share operational services among funds within a fund family).

153. *See, e.g.*, Jose Miguel Gaspar, Massimo Massa & Pedro Matos, *Favoritism in Mutual Fund Families? Evidence on Strategic Cross-Fund Subsidization*, 61 J. FIN.O 73 (2006) (demonstrating that sponsors engage in cross-fund subsidization to create "high-value funds.").

154. We use the following fields to match fund families: 'zipba', 'bas1', and 'bas2' to group funds into families.

155. Note, this number includes all funds that are single actors and thus each fund represents a fund family with 1 fund.

156. Data on file with authors; *see GSU Mutual Fund Research*, *supra* note 125 (hosting replication code).

point, over 400 funds in our sample come from the top five fund families, like Oppenheimer and BlackRock.[157] Conversely, 141 fund families have between 2–5 funds.[158]

We observe less similarity in the strategy sections within fund families compared to risk sections. This finding is intuitive—strategy sections describe the investment strategy specific to each fund. This creates less room for house language to shape the disclosure. However, even funds with different investment strategies may be exposed to similar market risks or external risks such as COVID-19, inflation, and other widespread phenomena.[159]

To explore the effects of house language, we compared similarity scores of DE funds' disclosures within the same fund family to disclosures of randomly paired DE funds (i.e., funds not in the same family).[160] For example, if Fund Family A has 5 DE funds, we compared the risk and strategy statements of A1, A2, A3, A4, and A5 using the two measures and created an average similarity score for Fund Family A in a given year. We repeated this exercise for 10 fund family groupings,[161] and created average, composite scores for each measure, as depicted in Figure 5 and Figure 6.

To create the comparison with the *unrelated* pairs, we randomly generated groupings of DE funds that do not belong to the same fund family and repeated the calculation process. For example, we calculated the similarity of risk and strategy statements between Fund A and Fund B for a given year—so long as A and B are from different fund families. Averages were taken for each measure in each year.

Similarity scores are higher within fund families for both risk and strategy sections across all years for both measures compared to random pairs. All differences in text similarity between within-family and out-of-family pairs were uniformly

157. Data on file with authors. Top 5 means the largest 5 fund families based on the number of funds and assets under management (AUM).

158. Data on file with authors.

159. *Aggregated Risks*, *supra* note 7, at 70–74 (describing a case study of three different types of market risks and funds' varied responses to each).

160. We focus on domestic equity only to exclude variation introduced by different asset classes, i.e., differences between fixed income and equity securities, for example.

161. The top 5 fund family by number of funds in matched business addresses are Oppenheimer, Dreyfus, T. Rowe Price, Deutsche Bank, and BlackRock. Data on file with authors.

statistically significant. (see Tables 7 and 8). In total, we conducted 48 t-tests.[162]

Disclosures originating from the same fund family, especially in the principal risk section, are more similar than disclosures from random pairs. It would be surprising if funds within the same family described similar risks in disparate ways. From this, we can infer that using consistent language for common risks promotes efficiency and reduces transaction costs.

We also observe that within-family similarity is strongest using the TF-IDF/cosine measure. Because this measure uses weighted keywords, this indicates that disclosures from the same fund family include the same unique and/or uncommon words. This is true of both the principal risks and investment strategy disclosures.[163]

Out-of-family disclosure pairs show almost no boilerplate (see Table 8).[164] Using the 5-gram shingle measure, out-of-family pairs share little-to-no common language used in the same order.[165]

162. The 48 tests include 24 each for strategy and risk (two measures per year across 12 years).

163. Funds employing similar strategies or relying on the same subadvisors will use similar and/or identical language across disclosures. Alternatively, this measure may capture the inclusion of the fund family name, which would be weighted heavily under the TF-IDF weighting scheme.

164. This conclusion is supported by an analysis of the variance of each of these out-of-family distributions, which is approximately 0 for all score-year subsets. This means, practically speaking, that the distribution of similarity scores between out-of-family pairs is clustered very tightly around zero.

165. Recall that the 5-gram cosine takes context into account, so formatting and order of text along with the language used are reflected in the near-zero scores.

FIGURE 5: COMPARISON OF TEXT SIMILARITY MEAN SCORES OF PRINCIPAL RISK SECTIONS BETWEEN WITHIN-FUND FAMILY PAIRS AND OUT-OF-FUND FAMILY PAIRS, BY YEAR

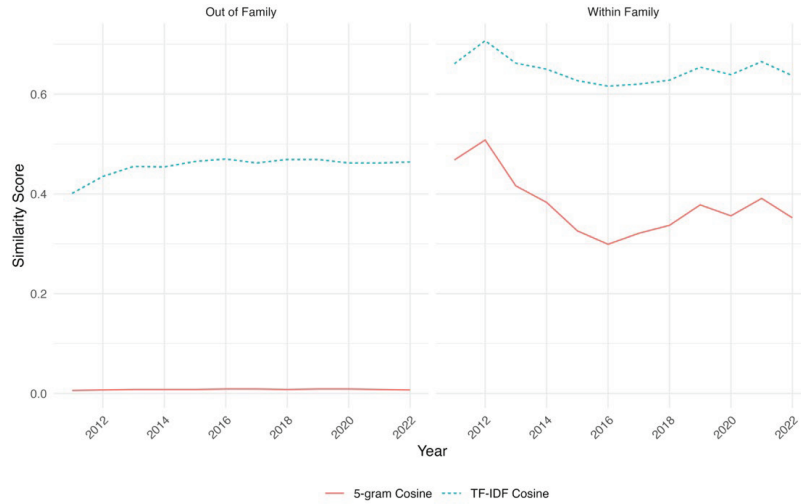


TABLE 7: COMPARISON OF MEAN SIMILARITY SCORES BETWEEN WITHIN-FUND FAMILY AND OUT-OF-FUND FAMILY PAIRS FOR PRINCIPAL RISK DISCLOSURES, BY YEAR

	TF-IDF/cosine	*Five-gram/cosine*
2011	(0.661, 0.401)	(0.468, 0.006)
2012	(0.707, 0.435)	(0.508, 0.007)
2013	(0.662, 0.455)	(0.416, 0.008)
2014	(0.65, 0.454)	(0.383, 0.008)
2015	(0.627, 0.465)	(0.326, 0.008)
2016	(0.616, 0.47)	(0.299, 0.009)
2017	(0.62, 0.462)	(0.321, 0.009)
2018	(0.628, 0.469)	(0.337, 0.008)
2019	(0.654, 0.469)	(0.378, 0.009)
2020	(0.639, 0.462)	(0.356, 0.009)
2021	(0.665, 0.462)	(0.391, 0.008)
2022	(0.637, 0.464)	(0.352, 0.007)

Note: The first number in each cell is the *within-fund family* mean; the second is the *out-of-fund family* mean. The differences in means were calculated using a two-tailed t-test. Every comparison across the period of study and for all scores is statistically significant at $p < 0.001$.

FIGURE 6: COMPARISON OF TEXT SIMILARITY MEAN SCORES OF INVESTMENT STRATEGY SECTIONS BETWEEN WITHIN-FUND FAMILY AND OUT-OF-FUND FAMILY PAIRS, BY YEAR

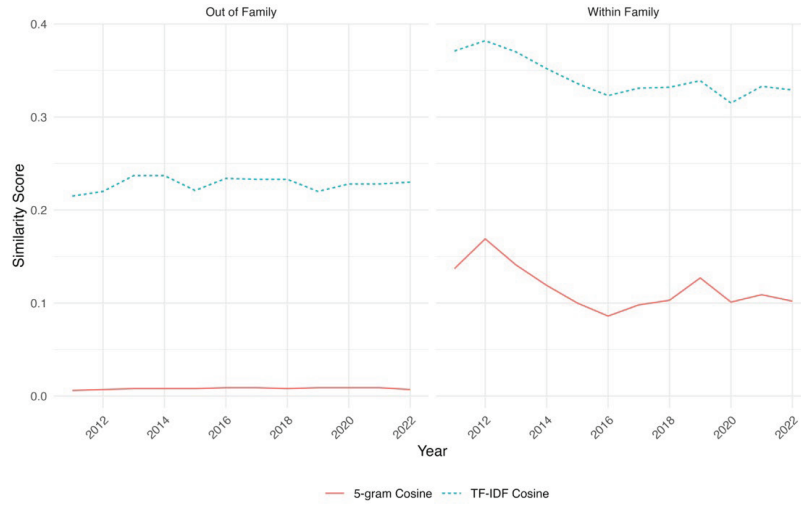


TABLE 8: COMPARISON OF MEAN SIMILARITY SCORES BETWEEN WITHIN-FUND FAMILY AND OUT-OF-FUND FAMILY PAIRS FOR INVESTMENT STRATEGY DISCLOSURES, BY YEAR

	TF-IDF/Cosine	*Five-gram/Cosine*
2011	(0.371, 0.215)	(0.137, 0.006)
2012	(0.382, 0.22)	(0.169, 0.007)
2013	(0.37, 0.237)	(0.141, 0.008)
2014	(0.352, 0.237)	(0.119, 0.008)
2015	(0.336, 0.221)	(0.1, 0.008)
2016	(0.323, 0.234)	(0.086, 0.009)
2017	(0.331, 0.233)	(0.098, 0.009)
2018	(0.332, 0.233)	(0.103, 0.008)
2019	(0.339, 0.22)	(0.127, 0.009)
2020	(0.315, 0.228)	(0.101, 0.009)
2021	(0.333, 0.228)	(0.109, 0.009)
2022	(0.329, 0.23)	(0.102, 0.007)

Note: The first number in each cell is the *within-fund family* mean; the second is the *out-of-fund family* mean. The differences in means were calculated using

a two-tailed t-test. Every comparison across the period of study and for all scores is statistically significant at $p < 0.001$.

B. *Boilerplate's Opposite: Change*

Funds repeat disclosure language from year to year and within fund families. But *some* fund language changes over time. In this Subsection, we evaluate whether changes are systematic (i.e., meaningful) or random.

First, we evaluate our *Emerging/Evolving Strategy* hypothesis*:* strategy disclosures for funds pursuing emerging or evolving investment strategies will contain less boilerplate than disclosures with traditional strategies. We find that funds adopting innovation-related or ESG strategies contain significantly less boilerplate than funds not pursuing such strategies.

We then turn to our *Portfolio Change* hypothesis: Strategy disclosures will change more following portfolio changes compared to funds maintaining a stable distribution of portfolio assets. We isolate this effect by looking at funds with high levels of language change (i.e., lower levels of boilerplate). We then evaluate absolute changes in those funds' mix of equity securities, debt securities, and cash compared to other funds without a high level of language change. We find that, across all four asset types, funds with high levels of language change are significantly more likely to also change their portfolio composition.

1. *Emerging Strategy Funds Update Disclosure Language More Frequently Than Stable Strategy Funds*

We first look at funds with emerging or evolving strategies, like ESG and crypto assets. We expect emerging strategy funds will use less boilerplate than funds using established investment strategies—and we find empirical support for this hypothesis.

Language used to describe new strategies and attendant risks will be less settled—the terms of art for that type of disclosure have not yet crystallized. When the field is less settled, funds face higher rates of learning that require more rapid disclosure additions to stem growing liability threats.[166] Researchers have

166. *Aggregated Risks*, *supra* note 7, at 64–66 (describing how funds learn from the crowd and update risk perceptions and disclosures based on experiences and feedback from other funds' disclosures); *see also, Staff Statement on Importance of Delivering Timely and Material Information to Investment Company*

found that disclosure language converges into common themes over time.[167] The convergence suggests that funds learn about risk and ideal strategy expressions from other disclosures—evidencing shared meaning in practice.[168] However, convergence can only occur after sufficient time has passed for funds to learn and change the language to reflect the common, accepted phrasing.[169] With emerging investment strategies, a fund's time to learn is shorter, and the new risks are less known and understood. As a result, funds' disclosure language should be more variable (and less recycled) when using emerging investment strategies versus established strategies.

To test this theory, we identified funds investing in cryptocurrencies and digital assets[170] as one subset of innovation, and funds pursuing an ESG strategy[171] as another subset. We

---

Investors DIV. OF INV. MGMT: SEC. & EXCH. COMM'N (Apr. 14, 2020), https://www.sec.gov/investment/delivering-timely-material-information.

167. *Aggregated Risks*, *supra* note 7, at 85 (finding convergence in language).

168. *Id.* (finding empirical evidence of funds learning and updating risk perceptions and disclosure language).

169. "Funds' abilities to observe and learn are enhanced and, perhaps, encouraged by the regulatory environment governing disclosures." *Id.* at 65; *see also* Bryan D. Jones & Frank Baumgartner, THE POLITICS OF ATTENTION: HOW GOVERNMENT PRIORITIZES PROBLEMS (1999) (describing learning and convergence in a policy-making setting); Jens O. Zinn, SOCIAL THEORIES OF RISK AND UNCERTAINTY: AN INTRODUCTION 6 (2008) (finding that social processes influence individual risk assessments); Brian T. McCann, *Using Bayesian Updating to Improve Decisions Under Uncertainty*, 63 CAL. MGMT. REV. 26 (2020) (describing the general principles and applications of Bayesian updating which can be applied to mutual funds' learning from other funds' disclosures).

170. To identify innovation funds using cryptocurrencies, we manually reviewed disclosures to find discussions of crypto assets and risks, finding language variation and creating a final set of key words. If a fund included any of the key words in the strategy discussion, then the fund was labeled an innovation fund. The keywords are: 'Virtual currency', 'Digital currency', 'Digital wallet', 'Ethereum', 'Analogies to fiat currency', 'Convertible virtual currency', 'Digital commodity', 'Alternative crypto', 'Derivative in conjunction', 'Altcoins', 'Block reward', 'NFT', 'crypto', 'bitcoin', 'cryptocurrency', Blockchain', 'crypto', 'bit coin', 'digital assets', 'alternative currency', 'Altcoin', 'Blocks', 'Coin', 'Cryptoasset'.

171. To identify ESG funds, we began with a list of sustainable funds provided by Morningstar. MORNINGSTAR, U.S. SUSTAINABLE FUNDS LANDSCAPE 2024 IN REVIEW (2024), https://www.morningstar.com/lp/sustainable-funds-landscape-report. We then added any additional funds with that have "Sustainable", "ESG", "Social", or "Clean Energy" in their name. We also include all funds that belong to any of five fund families that are self-identified as ESG strategies: Calvert, Parnassus, Pax, Praxis, and Trillium. Our approach follows that used by Professor Jonathan Zytnick.*See* Jonathan Zytnick, *Do Mutual Funds Represent Individual Investors?*, 9 (Oct. 18, 2024) (unpublished paper) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3803690.

compare the similarity scores of the innovation fund subgroups against the similarity scores of all other funds (non-innovation) using a difference in means test.[172]

We chose cryptocurrencies and ESG as our two emerging strategies using a most-likely case design.[173] Both cryptocurrencies and ESG investing are subject to changing regulatory environments through new rules[174] and enhanced SEC scrutiny.[175] Both strategies also grew rapidly during our study period.[176] Both

---

172. Our sample contains 607 ESG funds and 1498 "innovation" funds. Data on file with authors.

173. Most likely case studies focus on crucial cases that are "most likely or least likely to exhibit a certain outcome." Chiara Ruffa, *Case Study Methods: Case Selection and Case Analysis*, *in* THE SAGE HANDBOOK OF RESEARCH METHODS IN POLITICAL SCIENCE AND INTERNATIONAL RELATIONS 1133, 1139 (citing JOHN GERRING, CASE STUDY RESEARCH: PRINCIPLES AND PRACTICES 243 (2007)).

174. In March 2022, the SEC proposed rule around ESG investing that would restrict the use of ESG in funds' names and require ESG funds to provide additional disclosures around their ESG integration, focus, or impact strategies. Press Release, SEC, SEC Proposes to Enhance Disclosures by Certain Investment Advisers and Investment Companies About ESG Investment Practices (May 25, 2022), https://www.sec.gov/news/press-release/2022-92.

175. The SEC has brought several enforcement actions against those selling securities and investments in crypto assets and has a page cataloging such actions available here, https://www.sec.gov/spotlight/cybersecurity-enforcement-actions. Perhaps the highest profile of the enforcement actions, which implications for registered funds included in this study, is the enforcement action against FTX in 2022. Michael Thasher, *Institutions Weigh Exposure to FTC Crypto Exchange*, PENSIONS & INVESTMENTS (Nov. 10, 2022), https://www.pionline.com/cryptocurrency/institutional-investors-measure-damage-ftx-collapse-portfolios (noting institutional investors with direct or indirect exposure to FTX including BlackRock).

In 2021, the SEC created an Enforcement Task Force Focused on Climate and ESG Issues. Press Release, SEC, SEC Announces Enforcement Task Force Focused on Climate and ESG Issues (Mar. 4, 2021), https://www.sec.gov/news/press-release/2021-42. For an example of SEC enforcement actions based on ESG strategies in registered funds, see the 2022 action against BNY Mellon for "misstatements and omissions about Environmental, Social, and Governance considerations in making investment decisions for certain mutual funds" for which it agreed to pay $1.5 million in fines. In the Matter of BNY Melon Investment Advisor, Inc., Administrative Proceeding, File No. 3-20867 (2022); *see also* Press Release, SEC, SEC Charges BNY Mellon Investment Adviser for Misstatements and Omissions Concerning ESG Considerations (May 23, 2022), https://www.sec.gov/news/press-release/2022-86.

176. In 2021, Forbes estimated that $17 billion dollars of institutional investor capital entered crypto markets. Lawrence Wintermeyer, *Institutional Money Is Pouring Into The Crypto Market And It's Only Going To Grow*, FORBES (Aug. 12, 2021), https://www.forbes.com/sites/lawrencewintermeyer/2021/08/12/institutional-money-is-pouring-into-the-crypto-market-and-its-only-going-to-grow/?sh=5a7cf3081459. PWC reported that despite "turbulences, crypto /

strategies introduce new risks and uncertainties.[177] In other words, we expect that if emerging investment strategies have a relationship with boilerplate language, we will observe it with ESG and cryptocurrencies.

Table 9 and Table 10 show that both cryptocurrency and ESG funds have less similar strategy and risk section disclosures compared to other (non-innovation) funds, and the differences are statistically significant at $p < 0.001$ for both measures. Funds investing in innovation strategies use less recycled language in the strategy and risk discussions compared to non-innovation funds.

---

blockchain has established itself as an essential part of professional and institutional investing." *How Banks and Institutional Investors Are Embracing the Crypto / Blockchain Opportunity*, STRATEGY &: PWC, https://www.strategyand.pwc.com/ch/en/industries/financial-services/crypto-investor-survey.html, (last visited Aug. 4, 2023). After FTX, however, the numbers dropped sharply, some estimating as much as 95% from 2018. Stacey Elliot, *Institutional Crypto Investments Dropped 95% to $433 Million in 2022* DECRYPT (Jan. 4, 2023), https://decrypt.co/118425/crypto-investment-falls-95-percent-to-433-million-2022.

In 2022, the ESG investment market in the US was estimated at $330 billion in assets under management (AUM) with expected strong growth. Greg Norman & Simon Toms, *ESG: 2021 Trends and Expectations for 2022*, HARV. L. SCH. FORUM ON CORP. GOV. (Feb. 25, 2022), https://corpgov.law.harvard.edu/2022/02/25/esg-2021-trends-and-expectations-for-2022; *see also* Zoe Q. Cooper Sutton et al., *ESG in 2021 So Far: An Update*, SKADDEN (Sept. 1. 2021), https://www.skadden.com/insights/publications/2021/09/esg-in-2021-so-far-an-update. In 2022, the Dow Jones issued a press release stating that ESG investing was the "number one growth opportunity" for investment professionals. Press Release, Dow Jones, ESG Investment Expected to More Than Double in the Next Three Years New Research from Dow Jones Shows (Sept. 7, 2022), https://www.dowjones.com/press-room/esg-investment-expected-to-more-than-double-in-the-next-three-years-new-research-from-dow-jones-shows/.

177. For an overview of crypto asset risks and the potential for crypto risks to spillover to financial markets and the real economy, see Maria Teresa Chimienti, Urszula Kochanska & Andrea Pinna, *Understanding the Crypto-Asset Phenomenon, Its Risks and Measurement Issues*, EUROPEAN CENTRAL BANK 19 Econ.Bull. (2019), https://www.ecb.europa.eu/pub/economic-bulletin/html/eb201905.en.html. For a discussion of ESG investment risks and uncertainty, see, e.g., Gil Cohen, *The Impact of ESG Risks on Corporate Value*, 60 REV. QUANT. FIN. ACC. 1451 (2023), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9888342/ (finding that traditional risk models such as CAPM fail to account for social risks and current models do not accurately capture it); Doron Avramov et al., *Sustainable Investing with ESG Rating Uncertainty*, 145 J. FIN. ECON. 642, 663–64 (2022) (finding that ESG rating uncertainty "leads to higher perceived market risk").

TABLE 9: T-TEST COMPARISON OF SIMILARITY SCORE MEANS ON INVESTMENT STRATEGY AND PRINCIPAL RISK SECTIONS FOR ESG AND NON-ESG FUNDS' DISCLOSURES, BOTH MEASURES FOR ALL YEARS COMBINED

	Investment Strategy (ESG, Non-ESG)	*Principal Risks (ESG, Non-ESG)*
TF-IDF/Cosine	(0.890, 0.957)***	(0.909, 0.952)***
Five-gram/Cosine	(0.846, 0.925)***	(0.836, 0.898)***

Note: The first number represents the ESG mean, and the second number represents the non-ESG mean. ***: $p < 0.001$.

TABLE 10: T-TEST COMPARISON OF SIMILARITY SCORE MEANS ON INVESTMENT STRATEGY AND PRINCIPAL RISK SECTIONS FOR INNOVATION AND NON-INNOVATION FUNDS' DISCLOSURES, BOTH MEASURES FOR ALL YEARS COMBINED

	Investment Strategy (Innovation, Non-Innovation)	*Principal Risks (Innovation, Non-Innovation)*
TF-IDF/Cosine	(0.946, 0.957)***	(0.940, 0.952)***
Five-gram/Cosine	(0.910, 0.924)***	(0.883, 0.898)***

Note: The first number represents the Innovation mean, and the second number represents the non-Innovation mean. ***: $p < 0.001$.

FIGURE 7: COMPARISON OF SIMILARITY SCORE MEANS OF INVESTMENT STRATEGY AND PRINCIPAL RISK SECTIONS FOR ESG AND NON-ESG FUNDS' DISCLOSURES, BOTH MEASURES FOR ALL YEARS COMBINED

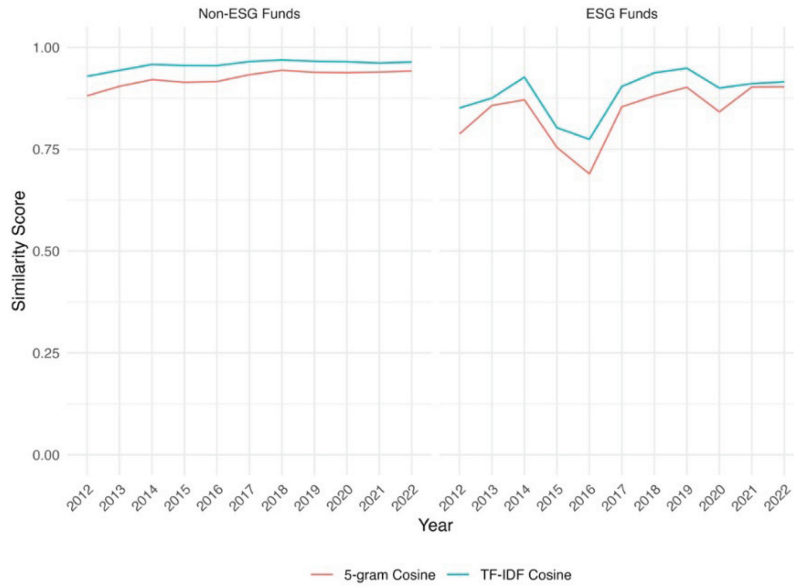
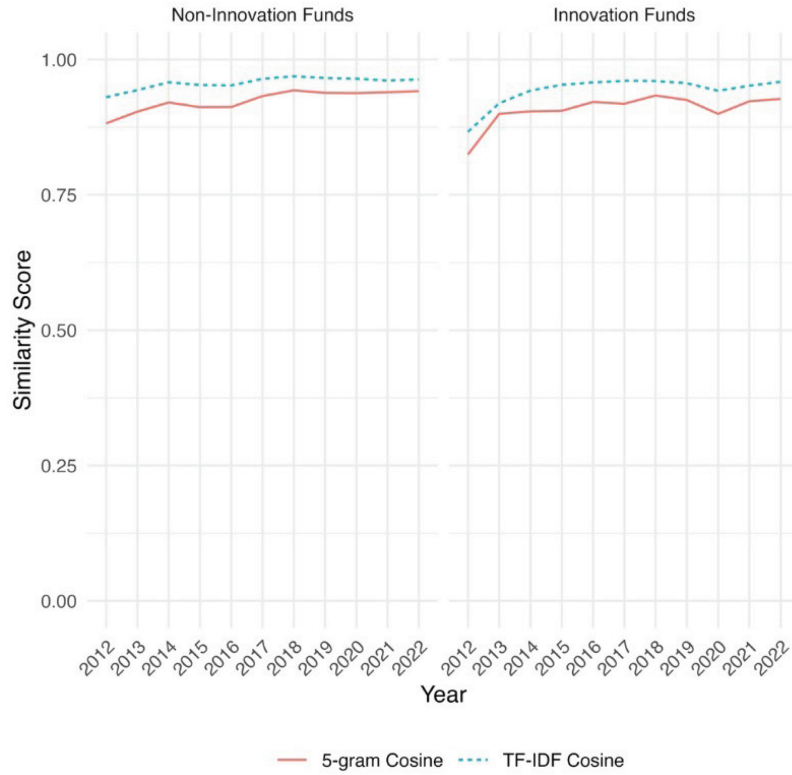


FIGURE 8: COMPARISON OF SIMILARITY SCORE MEANS ON INVESTMENT STRATEGY AND PRINCIPAL RISK SECTIONS FOR INNOVATION AND NON-INNOVATION FUNDS' DISCLOSURES, BOTH MEASURES FOR ALL YEARS COMBINED



Differences in means are most pronounced for the ESG subgroup. The year-by-year average similarity scores also show more movement, with ESG funds exhibiting sharp changes in the risk sections.

To illustrate the differences, consider two disclosures from BNY Mellon DE funds in 2021. The ESG fund, BNY Mellon Sustainable US Equity Fund, Inc., with a similarity score of .817[178] (just below the mean for ESG funds), increased its word count

178. Reporting the 5-gram shingle cosine score.

by 57% and added a new closing paragraph.[179] The non-ESG fund, BNY Mellon Large Cap Equity Fund, with a similarity score of .95[180] (slightly above the mean for non-innovation funds), changed its word count by 10%, dropping a single sentence at the end.[181]

2. *Significant Changes to Investment Strategy Disclosures & Portfolio Changes*

Next, we turn to the relationship between boilerplate language and portfolio change. We know that boilerplate predominates in fund disclosures from year to year and within fund families. As indicated by our *Portfolio Change* hypothesis, we expect that funds with higher levels of language change (i.e., lower levels of boilerplate) will have higher levels of portfolio change. That is, when funds change language, we expect it to be related to changes in the types of assets that they hold. This is exactly what we find.

We begin this assessment by looking at a subset of investment strategy disclosures that exhibited the most change among all disclosure-year pairs. We do not focus on risk statement changes because variability in risk statements may reflect external events, such as COVID-19, and may not signal a true change in the portfolio. In particular, we examine textual changes in funds in the bottom 5% of all five-gram cosine

---

179. BNY Mellon Sustainable US Equity Fund, Inc., Summary Prospectus, Form 497K (May 31, 2021). The fund changed portfolio manager to subadvisor and added the following paragraph:

> The fund may also invest in low- or non-dividend paying companies that may not have been identified by the model discussed above, if the fund's portfolio manager believes such investment will manage portfolio risk or have the potential for dividend payments in the future. The fund's portfolio manager monitors the holdings in the fund's portfolio and considers selling a security if the company's relative attractiveness deteriorates or if valuation becomes excessive. The fund also may sell a security if an event occurs that contradicts the portfolio manager's rationale for owning it, such as deterioration in the company's fundamentals. In addition, the fund may sell a security if the portfolio manager believes better investment opportunities emerge elsewhere.

Id.

180. Reporting the 5-gram shingle cosine score.

181. BNY Mellon, Summary Prospectus, Form 497K (Dec. 31, 2021). The disclosure removed this language: "When the fund enters into derivatives transactions, it may be required to segregate liquid assets or enter into offsetting positions, in accordance with applicable regulations."

similarity scores.[182] This qualitative assessment lets us peer into how a reader may perceive the changed text. We then turn to a quantitative assessment of portfolio change associated with the most revised disclosures, examining the correlation between investment strategy disclosure change and portfolio holdings.

Our qualitative assessment suggests that disclosure changes do reflect modifications to portfolio holdings—or at least the ranges of categories in which the fund is planning to invest. Changes also reflect an effort to better explain the fund's investment philosophy, provide more detail about the fund's investment or valuation strategies, or expand on the role of affiliate advisors.

Consider, for example, changes in Brown Capital Management's Mid-Cap Fund strategy disclosure from 2013 to 2014.[183] In 2013, Brown Capital described the "two specific criteria" required for an asset's inclusion in the fund—the Advisor must "believe [the company] can generate a prospective earnings growth rate in excess of the overall market's earning growth rate" and that the asset is "trading at a reasonable valuation." The 2013 investment strategy disclosure focused on describing each of these two steps, providing details about the sources used to construct models to assess a company's projected growth and determine the asset's value.

The 2014 disclosure reveals no fundamental change in these two steps but is less technical in its approach. Whereas the 2013 disclosure discusses how the Advisor constructs "company-specific risk premiums" using the 5-year Treasury yield, the 2014 disclosure focuses on less tangible measures of a company's potential for growth. For example, the 2014 disclosure professes the Advisor's belief that "a sustained commitment to a portfolio of exceptional companies will, over time, generate attractive long-term returns" and that such "exceptional companies save time, lives, money, and headaches and provide an exceptional value proposition to consumers."[184]

---

182. A fund in the bottom 5% of the *similarity* scores means that the disclosure changes more than 95% of all other fund disclosures.

183. The five-gram cosine score comparing the 2014 investment strategy disclosure to the 2013 disclosure is 0.217. The five-gram Jaccard and BOW scores (0.114 and 0.024, respectively) suggest even less similarity. The TF-IDF cosine score (0.59) suggests that many of the keywords substantially changed—which is exactly what we observe with our qualitative analysis.

184. Brown Capital Management Mutual Funds, Mid-Cap Funds. Comparison of 0001209286-13-000311 with 0001209286-14-000382.

Whether these changes represent an actual difference in portfolio holdings is unclear from a qualitative analysis—but they certainly represent a notable deviation from 2013's dry and technical approach. In short, the extensive changes to the fund's strategy disclosure make it more accessible but also less precise.

Other funds' disclosures reflect the addition of new subadvisors and a description of their strategies. This, in turn, suggests that the changes represent a portfolio shift. For example, Wilshire changed subadvisors for its Large Cap Core Plus fund between 2011 and 2012, dropping UBS Global Asset Management Americas and adding Santa Barbara Asset Management. The 2011 disclosure described UBS as employing a "primarily fundamental analysis and, to a lesser extent, quantitative analysis to identify securities that are underpriced and overpriced relative to their fundamental value."[185] Wilshire describes Santa Barbara, alternatively, as adopting a "Dividend Growth strategy" and focusing on "higher dividend yields."[186]

Next, we turn to a quantitative analysis of portfolio change to assess whether large changes in strategy disclosures systematically portend portfolio change. We examine portfolio change by looking at the distribution of equity securities, debt securities, cash, and other holdings in each fund's portfolio of assets.[187] We use this measure of asset distribution to see how overall risk profiles change for a given fund. Consider the example of a personal savings profile where a person holds a mixture of cash savings, guaranteed income products like bonds and investments in stocks. The individual bonds or stocks may change without the overall risk profile changing. But the risk profile for an individual investor changes as an investor allocates more money to cash or guaranteed assets (less risk) or more money to equities (more risk).

---

185. Wilshire Large Cap Core Plus Fund, Summary Prospectus (497K), May 1, 2011, https://www.sec.gov/Archives/edgar/data/890453/000139834411001008/fp0002882_487k.htm.

186. Wilshire Large Cap Core Plus Fund, Summary Prospectus (497K), May 1, 2012, https://www.sec.gov/Archives/edgar/data/890453/000139834412001573/fp0004713_497k.htm.

187. Fund portfolio holdings are assigned an asset type using CRSP data codes. Debt securities include bonds and debt instruments. The "other" category reflects portfolio assets with a negative market_val, which could mean a short, swap, or derivative position, or indicate a missing value. *See Survivor-Bias-Free U.S. Mutual Fund Guide*, CTR. FOR RSCH. IN. SEC. PRICES (CRSP) (2014), https://wrds-www.wharton.upenn.edu/documents/410/CRSP_MFDB_Guide.pdf.

We prefer assessing risk profiles rather than simply counting how much a given fund buys or sells particular assets. Individual sales occur with considerable frequency due to fund flows and economic factors.[188] Changes to specific company holdings of a fund may not signal a change in the overall strategy as strongly as a change to the risk profile. But if a fund changes its distribution of cash to debt or equity securities, its risk profile shifts.

We measure the percentage of equity securities, debt securities, cash, and other holdings comprising a fund's portfolio each year. We then calculate the difference in the portfolio's distribution across these four assets by taking the absolute value of the difference between the prior year's and the current year's portfolio distributions.[189] For example, if a fund held 90% debt securities in 2012 and 85% debt securities in 2013, we would record the difference as five percentage points.

TABLE 11: DISTRIBUTION OF ABSOLUTE CHANGES IN PORTFOLIO (PERCENTAGE POINT CHANGE) IN NON-HIGH AND HIGH CHANGE DE DISCLOSURES (5-GRAM SHINGLE COSINE)

	Mean	*S.D.*	*Min*	*25%*	*Median*	*75%*	*Max*
Panel A: Non-High Change Disclosures							
Debt Sec.	1.86	5.42	0	0.04	0.38	1.53	100
Cash	1.65	5.24	0	0.04	0.28	1.30	98.04
Equity Sec.	1.90	5.05	0	0.01	0.26	1.72	85.71
Other	2.11	4.91	0	0.06	0.41	1.81	57.23
Panel B: High Change Disclosures							
Debt Sec.	4.68	11.72	0	0.11	0.71	2.95	88.29
Cash	3.2	9.65	0	0.06	0.43	2.13	98.04
Equity Sec.	3.27	9.13	0	0.00	0.44	2.70	84.23
Other	3.96	9.62	0	0.60	0.62	2.54	85.71

Note: Portfolio changes are calculated as the absolute difference between the prior year's portfolio and the current year's portfolio. "High Change" reflects whether the five-gram cosine similarity measure for a fund's disclosure of its investment strategy is in the lowest 5% of the distribution for that particular year.

188. Anne Tucker, *The Long & The Short: Portfolio Turnover Ratios & Mutual Fund Investment Time Horizons*, 43 J. OF CORP. L. 581, 588 (2018).

189. We also analyze the data using a measure of squared difference. Our results are robust across measures.

We assess similarity in this Subsection using the five-gram cosine metric. As discussed above, this measure allows us to retain the context of the text. Our findings, however, are consistent across both measures (TF-IDF/cosine and five-gram/cosine).

To do this, we compare portfolio changes in funds with a great deal of change in their disclosure text to portfolio changes in funds with less change in disclosure text. A disclosure is categorized as "high change" if its text similarity score is in the lowest 5% of scores for a particular year.[190] We limited our analysis to funds' investment strategy disclosures. Changes in these disclosures are the most likely predictors of actual portfolio change.

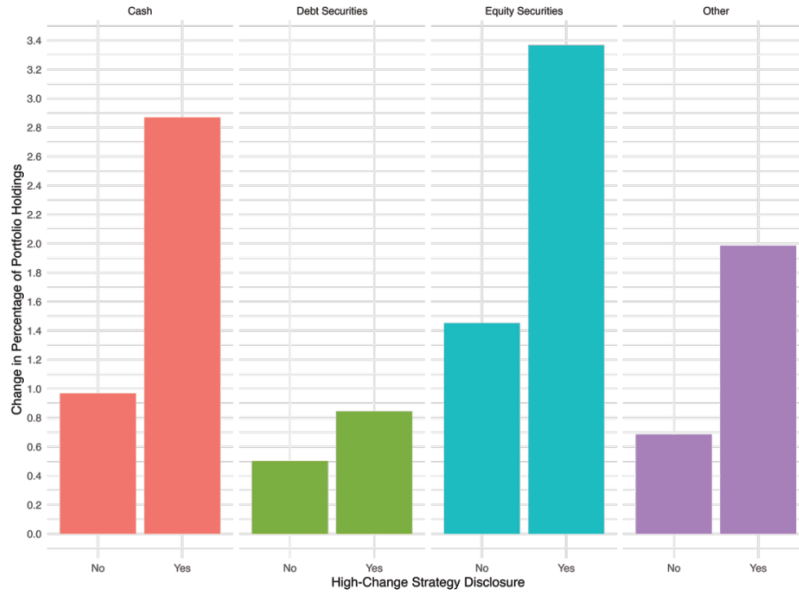
The summary statistics in Table 11 indicate that portfolio changes are more pronounced in high-change disclosures across all four asset types. In other words, portfolio changes are, on average, larger and more variable when funds' disclosures exhibit the most change for a particular year.[191] Funds with comparatively stable investment strategy language, on the other hand, experience far less change and variability in assets. These differences in average portfolio change are also displayed in Figure 9. [192]

190. We determine change by year to control for external shocks to the market, such as the Puerto Rican debt crisis or COVID-19. Pooling the disclosures and then determining whether a disclosure fell into the lowest 5% of text similarity measures would over-weight years where the market experiences a great deal of turmoil and would bias our results.

191. For clarity, most change is capturing the top 5% of disclosure changes for a given year, which is measured by being in the bottom 5% of similarity scores for that year.

192. We conducted additional sensitivity tests. by repeating the tests above and adding a comparison for large fund families versus small fund families. The analysis is available in our replication file and on file with the authors. Adding fund family size allows us to observe if the portfolio effects are driven by fund size or are a phenomenon that exists across different segments of the market. We find that the results of the large fund family group are similar to the results of the small fund family group and mirror the main findings. In other words, when analyzed by fund family size, our results hold for large and small fund families.

FIGURE 9: COMPARISON OF MEAN PORTFOLIO CHANGE FOR HIGH-CHANGE AND NON-HIGH-CHANGE DE DISCLOSURES
(5-GRAM SHINGLE/COSINE)

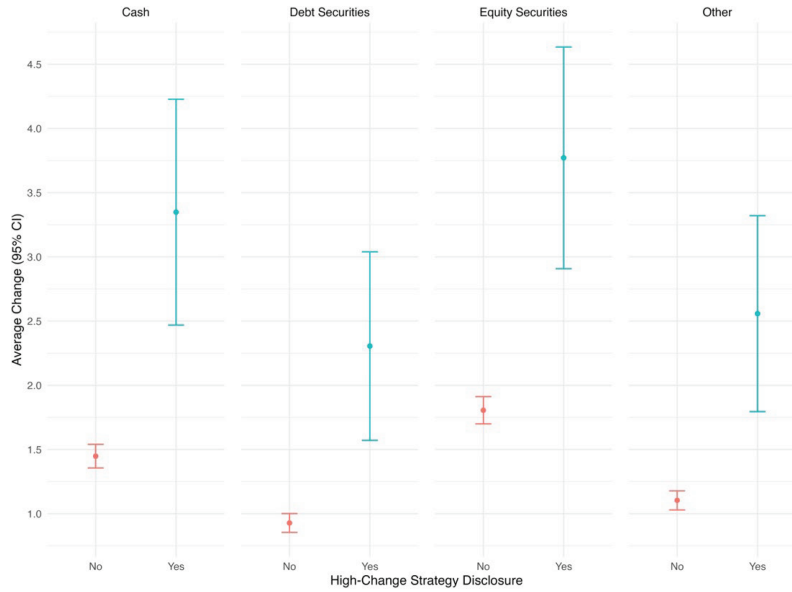


Note: A disclosure is categorized as "high-change" if it scores in the lowest 5% of 5-gram shingle similarity scores for a particular year for all DE funds.

These differences in means are also statistically significant. We used a series of t-tests to examine the difference in average portfolio changes between high-change and non-high-change investment strategy sections. In each case, we were able to statistically distinguish the high-change and non-high-change disclosures at the $p < 0.001$ level. We display the means and our 95% confidence interval estimates in *Figure 10*.[193]

193. Our findings are robust when we calculate the same t-tests using TF-IDF/cosine. These sensitivity results are available from the author upon request.

FIGURE 10: MEAN ASSET CHANGE IN FUND PORTFOLIO IN NON-HIGH CHANGE AND HIGH CHANGE DE DISCLOSURES (5-GRAM SHINGLE COSINE)



Note: This figure shows high-change DE disclosures with the blue line labeled "yes"; non-high change disclosures with the red line labeled "no"; scores reported in the 95% confidence interval; all scores are calculated using the 5-gram shingle cosine similarity measure.

Designating a disclosure as "high change" is a rather blunt instrument. Yet, we find that the relationship between disclosure similarity and portfolio change is quite robust to more granular analysis. We estimated the correlation coefficient (Pearson's *r*) and the concomitant t-statistic between each asset type change (debt securities, cash, equity securities, and other) and text similarity measure (TF-IDF/cosine and five-gram/cosine). The correlation coefficient for each is statistically significant and in the expected direction. Taken together, these findings suggest that as text similarity increases, portfolio change decreases.[194] A summary of these correlation tests is found in Table 12 below.

194. While these findings suggest a strong relationship between language changes and portfolio holdings, we see some reason for caution. It is possible that large funds (i.e., funds with high total net assets (TNA)) both change

Table 12: Correlation Between DE 5-Gram Shingle Cosine Text Similarity Score and Portfolio Change

	TF-IDF/ Cosine	*Five-Gram/ Cosine*
Debt Sec.	-0.051	-0.076
Cash	-0.069	-0.086
Equity Sec.	-0.076	-0.089
Other	-0.047	-0.082

Note: All correlation coefficients between similarity score and portfolio change are calculated as Pearson's *r* and are significant at $p < 0.001$. Additional information is available with the replication code.

IV.
Discussion & Conclusion

Our empirical approach to boilerplate identifies both stability and change. The findings, combined with the regulatory game model, demonstrate funds' incentives to use boilerplate to promote shared meaning and reduce transaction costs. When conditions on the ground—in the market or at the fund—change, however, funds are more willing to update disclosure language. Collectively, our work provides a theoretically grounded and empirically supported account of boilerplate's role in mutual fund disclosures.

---

their disclosures more frequently due to greater administrative capacity and are more likely to change their mix of assets. Our current data, however, lacks TNA information for many smaller stocks, which would bias any sensitivity test using this measure. Instead, we use fund family size as a proxy. Our sensitivity tests, which compare the results of large fund families to small fund families softens, suggest that statistical differences in portfolio changes are not driven by fund size. But we continue to explore ways to gauge the robustness of these results. For example, we are planning an additional sensitivity text that would smooth over some missing data across the portfolio by creating average total net asset (TNA) for a fund for purposes of comparing large and small fund results in this section. For an overview of data quality and completeness problems documented in the CRSP Survivor-Bias-Free U.S. Mutual Funds database, see Grace Liu, *Data Quality Problems Troubling Business and Financial Researchers: A Literature Review and Synthetic Analysis*, 25 J. of Bus. & Fin. Lib. 315, 321–25 (surveying the business literature's critical analysis of data quality and completeness mutual fund data and other financial databases).

Next, we present three practical applications for our theory and empirical approach, focusing on the SEC, internal fund operations, and individual investors and/or their intermediaries.

A. *Boilerplate & Regulators*

The SEC can monitor market movements, identify noncompliant disclosures, and encourage updated language when appropriate, using aggregated boilerplate analysis.

1. *Boilerplate Measures as an SEC Enforcement Tool*

Using a similar methodology to the one outlined in this Article, the SEC can identify problematic[195] boilerplate in mutual fund disclosures, such as:

(1) **Stale Disclosures—**those disclosures where no language has significantly changed for 5 or more years.[196]

(2) **Out-of-Proportion Boilerplate—**those disclosures that display the most boilerplate compared to the funds within the same asset class may trigger an additional review. Boilerplate averages can help identify outliers—the funds or fund families that change comparatively too much, or too little.

(3) **Legacy Language—**those disclosures where a fund significantly changes its investment strategy statement without updating the resulting risk narrative, or vice versa. A significant change to one section should generate changes to the other section, and if not, it should flag the disclosure for additional review.

---

195. The SEC can use boilerplate to assess and enforce compliance with disclosure regulations and advance consumer protection goals. *See, e.g.*, Michael D. Guttentag, *Protection from What? Investor Protection and the Jobs Act,* 13 U.C. DAVIS BUS. L.J. 207, 217 (2013) (describing the general "[a]cceptance of the centrality of investor protection to securities regulation in the United States").

196. The review window could be refined with more quantitative analysis of optimal review windows. We suggest a 5-year window as a starting point, in part, based on the study of 4 years of disclosure language discussed above.

Incorporating empirical analysis in the assessment adds transparency to enforcement decisions and consistency in application.[197] Empirical tools lessen the burden of individual assessment and allow limited SEC staff resources to be devoted to the most suspicious cases.[198] Further, as SEC agency power comes under increased scrutiny, including empirical data in enforcement actions and comments to quantify regulatory decisions protects the agency's administrative authority.[199]

2. *Boilerplate Measures as an SEC Market Monitoring Tool*

Collecting longitudinal data on boilerplate averages for fund types, individual funds, and fund families can also inform the SEC about disclosure patterns. Widespread and significant disclosure changes, absent regulatory interventions, suggest market changes. While it is unlikely that the SEC would use boilerplate data to confirm, for example, COVID-19's effects on securities markets, tracking changes to climate change risk disclosures may be one way to monitor the evolution of the risk.[200]

Other agencies relying on narrative disclosures as a regulatory and enforcement tool may similarly use our methodology, as could scholars in those fields. Our methodology can be used to quantify recycled language, build a descriptive account of how it is used, and theorize the uses, pitfalls, and problems of boilerplate in other disclosure regimes. What changes in credit card disclosures year to year, or medical billing disclosures between practices? Federal agencies overseeing narrative

---

197. A counter argument against empirical analysis is the bright line nature of such an approach invites clever evasion by tailoring behavior to be just outside or inside of a regulatory boundary. These types of regulatory interventions can be easily gamed to avoid meaningful oversight. In the context of boilerplate, a fund may change the minimum number of words to avoid a "stale" flag or stay one work under the "meaningful change" flag. For a discussion on the tension of principles based versus rules based regulations see, e.g., Dan Awrey, William Blair, & David Kershaw, *Between Law and Markets: Is There A Role for Culture and Ethics in Financial Regulation?*, 38 DEL. J. CORP. L. 191, 219–20 (2013) (describing rules-based financial regulations); Cristie L. Ford, *Financial Innovation and Flexible Regulation: Destabilizing the Regulatory State*, 18 N.C. BANKING INST. 27, 29 (2013) (describing dynamic financial regulation).

198. *See supra* Section I.B.

199. Loper Bright Enters. v. Raimondo, 603 U.S. 369, 412–13 (2024).

200. *Aggregated Risks*, *supra* note 7, at 93–95 (discussing the rise in climate change disclosures beginning in 2020 and the evolution of risk disclosures); *see also id.* at 97 (describing the different types of climate risks disclosed by funds).

disclosures and compliance regimes, like the Environmental Protection Agency,[201] Department of Education,[202] Commodity Futures Trading Commission,[203] and Federal Deposit Insurance Corporation,[204] may also find similar applications for boilerplate analysis. As disclosure continues to be a preferred regulatory solution, a better understanding of the limitations and potential of disclosures can help serve the underlying policy goals.[205]

B. *Boilerplate & Internal Audiences*

Funds are already the beneficiaries of boilerplate—reaping the transaction-cost and operational efficiencies of recycled language.

As the data show, legal teams in fund complexes develop family-specific boilerplate, or "house language," for disclosures. Such practices likely reduce the costs of implementing and monitoring disclosure language once the fund is operating.[206] House language also advances shared meaning among the internal review teams and those who operationalize compliance controls within a fund. Evidence of house language supports our argument that disclosure boilerplate functions as legal terms of art that balance regulatory requirements while providing fund managers with necessary operational flexibility.

---

201. For an example of EPA disclosures, see https://www.epa.gov/compliance/epas-edisclosure.

202. For an example of Department of Education disclosure, see https://fsapartners.ed.gov/knowledge-center/fsa-handbook/2022-2023/appendices/appx-e-institutional-reporting-and-disclosure-requirements.

203. For an example of CFTC registration and disclosure, see https://www.nfa.futures.org/registration-membership/who-has-to-register/fcm.html.

204. For an example of an FDIC-required disclosure, see https://www.fdic.gov/resources/regulations/federal-register-publications/dfoi.html.

205. "The American polity has for the better part of this century carried on a vigorous romance with the idea that the free flow of information is a potent remedy for social and political ills." Russell B. Stevenson, *The SEC and the New Disclosure*, 62 CORNELL L. REV. 50, 50 (1976).

206. House language resulting from shared resources bolsters claims about the internal governance function of disclosures made by other scholars. *See e.g.*, *supra* notes 163–169 and accompanying text (describing fund family house language effects). Shared resources result in shared language, reflect internal compliance mechanisms within fund families, and signal the role of in-house legal departments in crafting and implementing consistent language across a fund complex.

Stylized statements of strategy and risk also manage the complexity of certain strategies and risks.[207]

Beyond the drafting costs, a standardized approach to disclosures is logical under the securities regime, where funds must say what they do and do what they say. Disclosures create self-policing regulations. Funds are legally obligated to do what they say in the strategy section.[208] Funds must also provide a current assessment of investment risks. Fail to do either, and a fund risks regulatory enforcement or, worse, investor litigation. When investors lose confidence that a fund makes honest disclosures and hews to its own stated limits, they have new incentives to exit the fund and invest elsewhere. Funds have internal operational incentives to recycle language to make review, implementation, and adherence easier for the legal, compliance, and other internal fund shops. When funds set their own rules, they are held to a high standard if they fail to comply. Repeating familiar, well-known language helps those inside the fund tasked with monitoring and ensuring compliance with the disclosures.

C. *Boilerplate, Investors, & Intermediaries*

Given the incentives for funds to continuously recycle language, *what changes* becomes all the more important to intermediaries who review disclosures on behalf of investors (and the most diligent of individual investors). Isolating boilerplate reveals what changes—creating a red line of year-to-year changes. This goes to the heart of our findings. The incentives motivating language changes (and restricting them) suggest that most language changes are more than mere wordsmithing. Incentives to stay the same, coupled with the narrow scope of what changes, ratchet up the value of updated language. Our analysis of portfolio changes correlated with high disclosure changes confirms this. Funds incur transaction costs and regulatory risks to change language only when it is worth it, like when the countervailing threat of noncompliance—enforcement

207. *See, e.g.*, *supra* Section I.B.2. (discussing boilerplate as legal terms of art that promote shared meaning and reduces transaction costs).

208. *See, e.g.*, *Investment Company Registration and Regulation Package*, SEC, https://www.sec.gov/investment/fast-answers/divisionsinvestmentinvcoreg121504.

or litigation—is greater.[209] Investors (or more accurately, their intermediaries) should care about what changes.

Even if most individual investors are not sufficiently motivated or resourced to monitor their own investments, investment professionals, the sophisticated parties that read disclosures on their behalf are.[210] Investment intermediaries and professionals include broker-dealers, plan administrators, and industry analysts like those at Morningstar.[211]

For example, defined contribution plan administrators who oversee 401(k) and 403(b) plans care about what changes in funds because they must review and file on behalf of the plan annually,[212] and owe fiduciary duties to participants for these and other actions.[213] Recent litigation against university-sponsored defined contribution plans has brought renewed attention to plan administrators who facilitate many investors' market investments through employer-sponsored plans.[214]

---

209. Proving this last point is beyond the scope of this Article. We set out to primarily provide a descriptive account of boilerplate and the tools to do so. Our preliminary results find a relationship between those funds that change their disclosure the most and changes to their portfolio. Further, our qualitative assessment of what changes between disclosures bolsters the assertion that changes are informative by describing new investment advisors, new risks, and new strategies.

210. McClane, *supra* note 9, at 219–20 (expressing skepticism that individual investors read disclosures and discussing how disclosed information reaches investors, through analytics and other intermediaries).

211. *Aggregated Risks, supra* note 7, at 63–64.

212. *See, e.g.*, VANGUARD, *supra* note 7 ("Plan fiduciaries should review the prospectus and the fund's performance as they select and monitor plan investments."). *Id.* at 22–24 (describing fund selection for plan menus and the link to employee selection); *Hughes*, 142 S. Ct. at 738 ("[P]lans are defined-contribution plans governed by the Employee Retirement Income Security Act of 1974 (ERISA), under which each participant chooses an individual investment mix from a menu of options selected by the plan administrators.").

213. For a discussion of defined contribution plans and fiduciaries, see Anne Tucker, *Retirement Revolution: Unmitigated Risks in the Defined Contribution Society*, 51 HOUSTON L. REV. 153, 195–200 (2013). *See also* Employee Retirement Income Security Act of 1974 (ERISA), Pub. L. 93-406, 88 Stat. 829, (codified as amended in part at 29 U.S.C. Ch. 18); Hughes v. Nw. Univ., 142 S. Ct. 737, 742 (2022) (reversing the Seventh Circuit's grant of an ERISA plan administrator breach of fiduciary duty claim motion to dismiss).

214. *See e.g.*, Complaint, Garnick v. Wake Forest Univ. Baptist. Med. Ctr., 629 F. Supp. 3d 352 (M.D.N.C 2022); Tibble v. Edison Int'l, 575 U.S. 523, 531 (2015) (remanding a case on the issue of the statute of limitations application to a breach of fiduciary duty claim); Hughes, 142 S. Ct. at 742 (remanding the breach of fiduciary duty case on the merits); *see also* Duke 403(b) Plan Settlement, DUKE UNIV., https://duke403bsettlement.com [https://perma.cc/NX58-GSGM]. For a recently approved settlement, see generally Complaint,

Focusing administrator reviews on what has changed would be an efficient and effective aid for plan fiduciaries.[215]

When investors enter funds for long-term investments with goals of saving for retirement or college education,[216] knowing what changes tells an investor where to focus, especially over the life of an investment. Most investors know they should reassess investment selections, but few do, taking a "set it and forget it" approach.[217] Intermediaries can focus investors' attention on significant language changes to confirm the portfolio allocation or identify reasons to change investments.

Relatedly, in 2022, the SEC finalized rules for open-ended mutual funds and ETFs to summarize material changes in a fund's annual shareholder report.[218] These rules took effect in the summer of 2024 and can advance the investor and plan administrator uses identified above. Using our boilerplate assessment tool, the SEC can also confirm material change descriptions and flag potentially noncompliant filings.

Conclusion

Our comprehensive analysis of mutual fund disclosures over an 12-year period reveals that prevalent but meaningful changes still occur when investment strategies or market conditions shift. The balance between stability and change highlights

---

Brookins v. Ne. Univ., 731 F.Supp.3d 112 (D. Mass. 2024); Vellali v. Yale Univ., No. 16-cv-1345, 2022 WL 13684612.

215. Using the methodology described herein and code available at *https://sites.google.com/view/gsu-mutual-fund-research*, parties can generate line-by-line comparisons to see exactly what language changes.

216. 2023 ICI Factbook, *supra* note 150, at 34.

217. Laurent E. Calvert et al., *Fight or Flight? Portfolio Rebalancing by Individual Investors*. 124 *The Quarterly J. of Econ.* 124, 301–48 (2009), http://www.jstor.org/stable/40506230 (finding that few individual investors rebalance investment portfolios); *see also* Anne Tergesen, The Wall Street Journal, *Portfolio Rebalancing Is a Good Retirement Habit*, June 4, 2021, https://www.wsj.com/articles/portfolio-rebalancing-is-a-good-retirement-habit-11622815200.

218. Tailored Shareholder Reports for Mutual Funds and Exchange-Traded Funds; Fee Information in Investment Company Advertisements, 87 FR 72783-84, https://www.federalregister.gov/d/2022-23756. The rules require funds to disclose material changes to the fund's name, investment objectives, fees and expenses, principal investment strategies, principal risks, and investment advisers or sub-advisers. *Id. See also,* SEC, *Tailored Shareholder Reports for Mutual Funds and Exchange-Traded Funds*; fee information in investment company advertisement, Dec. 22, 2022, available at https://www.sec.gov/investment/tailored-shareholder-reports-mutual-funds-etfs.

the SEC's rational regulatory framework, where standardized language mitigates compliance risks, increases shared meaning, and reduces transaction costs. The consistent use of stable language in investment strategy sections and the development of "house" language within fund complexes underscore the practical benefits of reusing language, especially language that acts as disclosure terms of art. Funds often recycle previous language or adopt generic phrasing from competitors to efficiently navigate SEC mandates, meet investor demands, and maintain operational flexibility in dynamic markets.

Boilerplate, at first blush, may appear to be a bug in the SEC's regulatory design and a signal of a weak oversight system. Our regulatory model and empirical findings paint a more nuanced portrait of disclosure boilerplate, articulating the theory and finding evidence that boilerplate can be a feature of the system.

Our findings emphasize the importance of updated language, particularly when conditions on the ground change. Funds with evolving strategies, like ESG funds, and funds with observable shifts in portfolio composition are two examples.

Our theory and empirical approach offer practical insights for the SEC in enforcing disclosure requirements, identifying compliance concerns, and monitoring market trends. It also benefits investors and their intermediaries by making reviews easier and highlighting crucial changes. Funds, as both creators and beneficiaries of boilerplate, reduce transaction costs and regulatory uncertainty while providing enhanced clarity for investors and the SEC when they recycle road-tested language.