MUTUAL FUND DISCLOSURES:

reg_tech case studies

Dec. 2, 2020

AI & the future of law

Professor Anne Tucker

amtucker@gsu.edu

Georgia State University | COLLEGE OF LAW

Our Vision

We unlock legal data to advance legal knowledge and access.

Legal Analytics Lab at GSU

Partnership between the Institute for Insight at the Robinson College of Business & the College of Law

Georgia State University J. Mack Robinson College of Business

Institute for Insight

Expertise in:

- Text analytics
- Image processing
- Machine learning/AI
- Business disciplines – e.g. finance, insurance, marketing, accountancy

Georgia State University College of Law

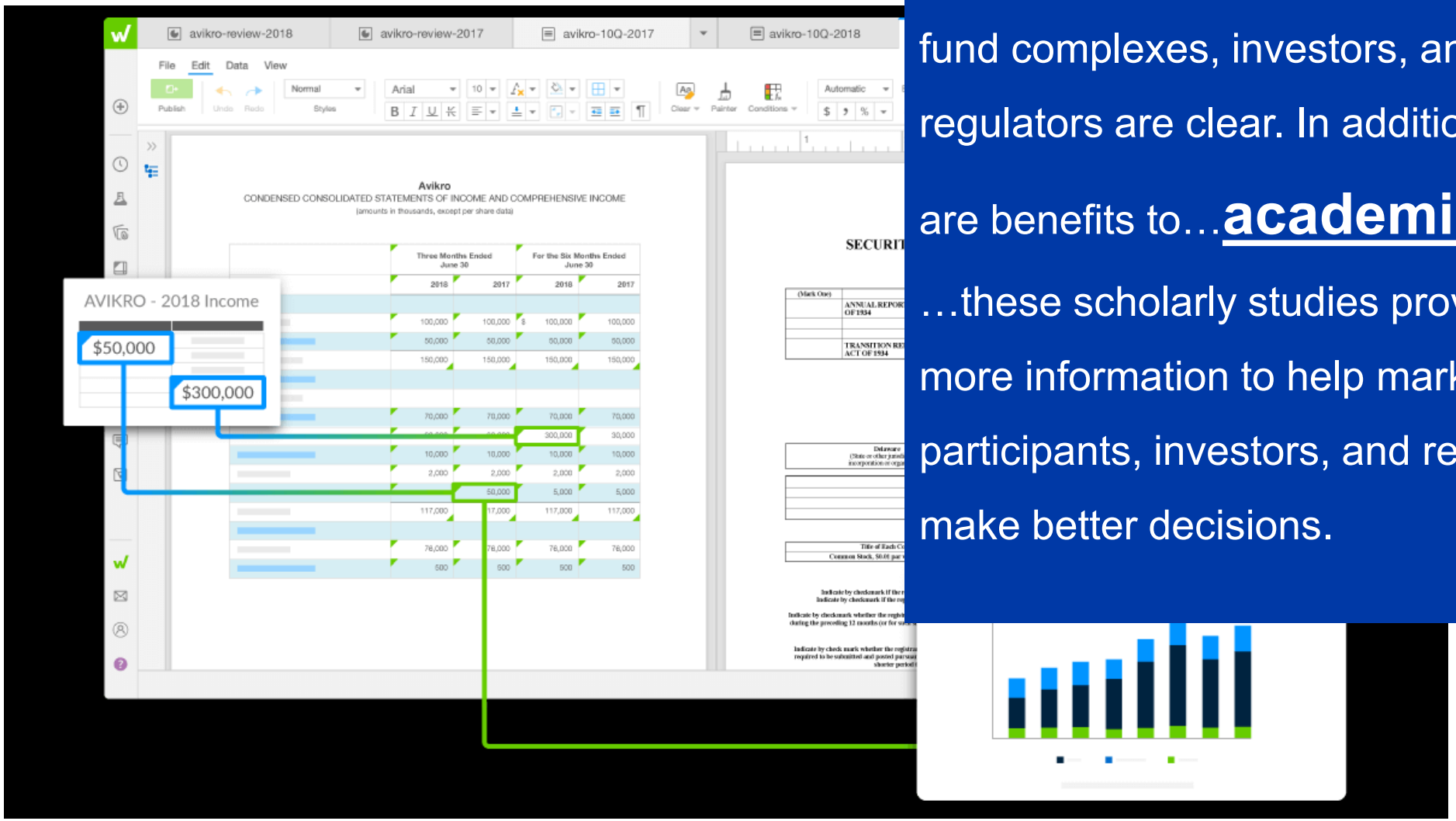
Expertise in:

- Employment law
- Corporate law
- Securities law
- Civil litigation and ADR
- Tort law
- Health law
- Tax law

Reg tech

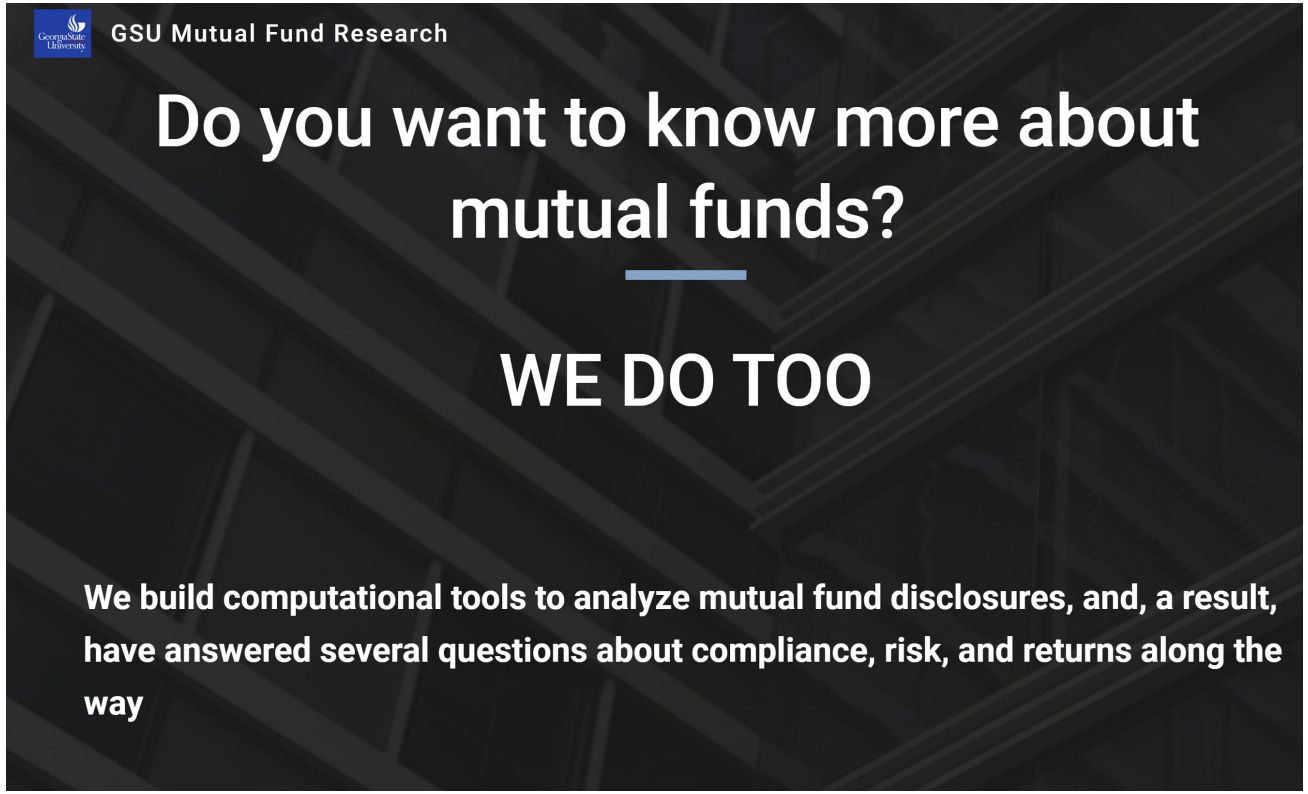
- *Regulators* use technology for efficient, transparent, and consistent oversight
- *Regulated entities* use technology for compliance efforts and to reduce legal and regulatory costs
- *Collaboration between private and public actors* to create a feedback loop for efficiency, transparency, & knowledge

XBRL SEC Edgar data



The direct benefits of XBRL to issuers, fund complexes, investors, and regulators are clear. In addition, there are benefits to…**academics**, …these scholarly studies provide more information to help market participants, investors, and regulators make better decisions.

Mutual Fund Disclosure Research Group



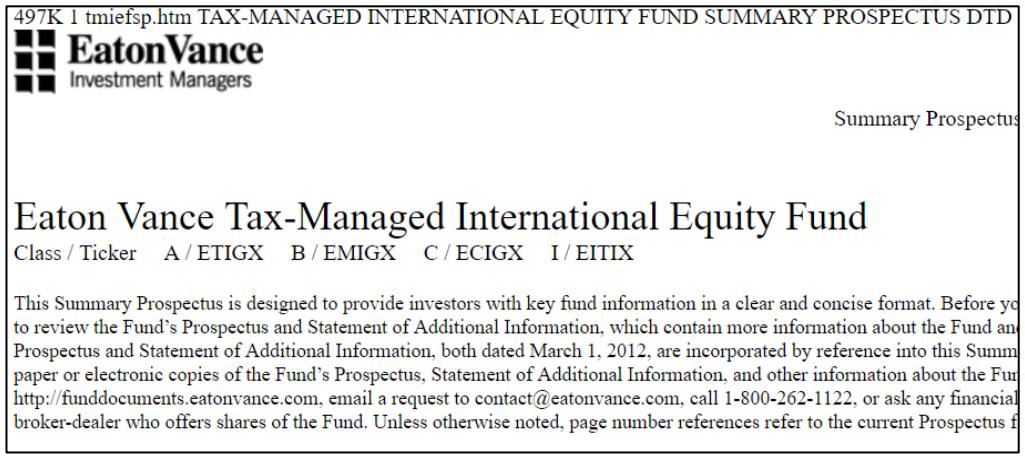
https://sites.google.com/view/gsu-mutual-fund-research

Mutual Fund Reg Tech

3 case studies

Emerging Risks	Plain English	Tone & Sentiment
Mapping the evolving risk landscape with topic modeling	Disclosures have gotten longer and more complex	Reading between the lines to discover sentiment

Corpus


497K 1 tmiefsp.htm TAX-MANAGED INTERNATIONAL EQUITY FUND SUMMARY PROSPECTUS DTD

EatonVance
Investment Managers

Summary Prospectus

Eaton Vance Tax-Managed International Equity Fund

Class / Ticker A / ETIGX B / EMIGX C / ECIGX I / EITIX

This Summary Prospectus is designed to provide investors with key fund information in a clear and concise format. Before yo
to review the Fund's Prospectus and Statement of Additional Information, which contain more information about the Fund an
Prospectus and Statement of Additional Information, both dated March 1, 2012, are incorporated by reference into this Summ
paper or electronic copies of the Fund's Prospectus, Statement of Additional Information, and other information about the Fur
http://funddocuments.eatonvance.com, email a request to contact@eatonvance.com, call 1-800-262-1122, or ask any financial
broker-dealer who offers shares of the Fund. Unless otherwise noted, page number references refer to the current Prospectus f

Data source: Prospectus 497K

- Principal Risk
- Investment Strategy

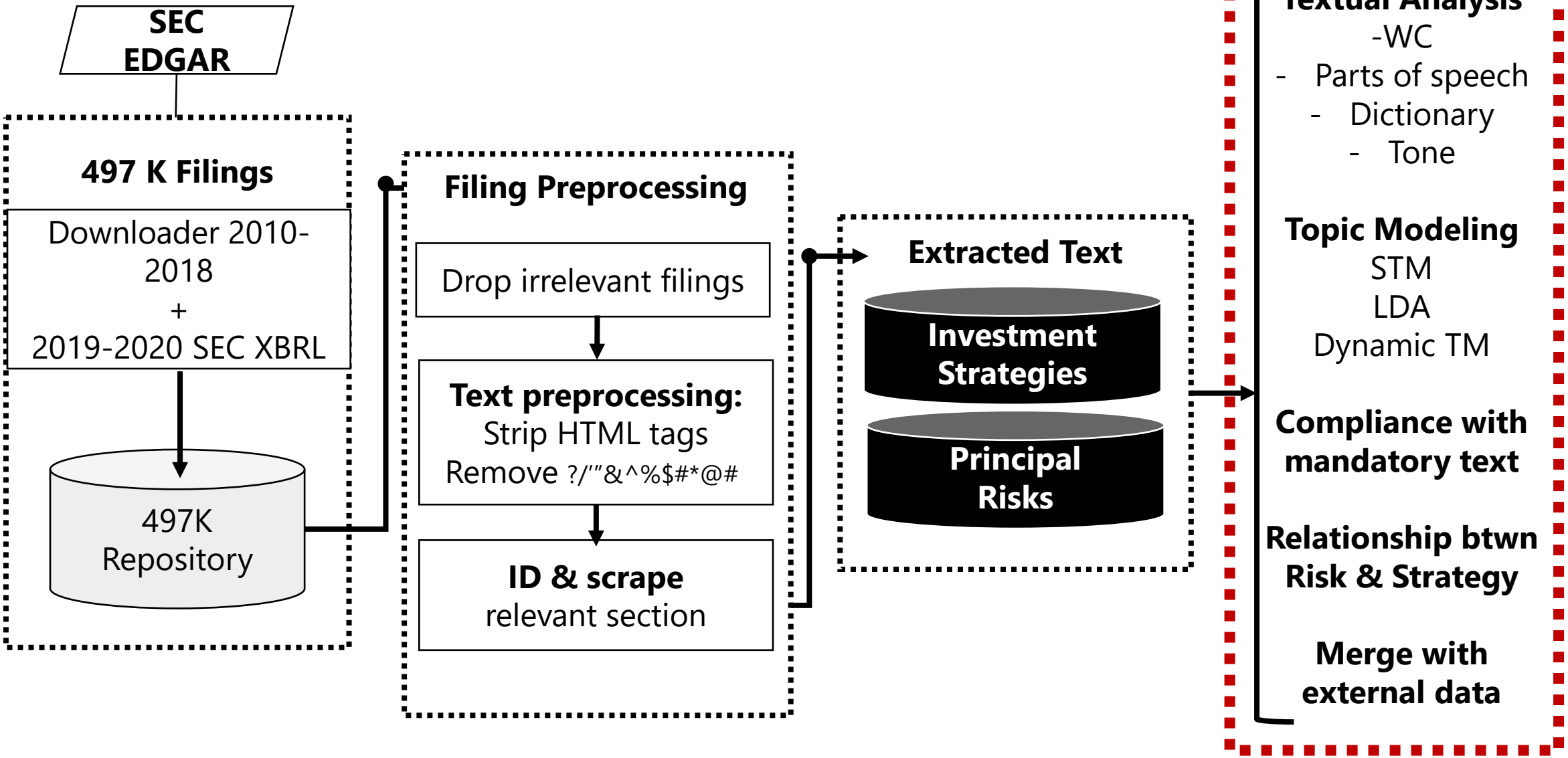
~200K disclosures

- 2010-2018 text scrape
- 2019-2020 SEC download (XBRL)

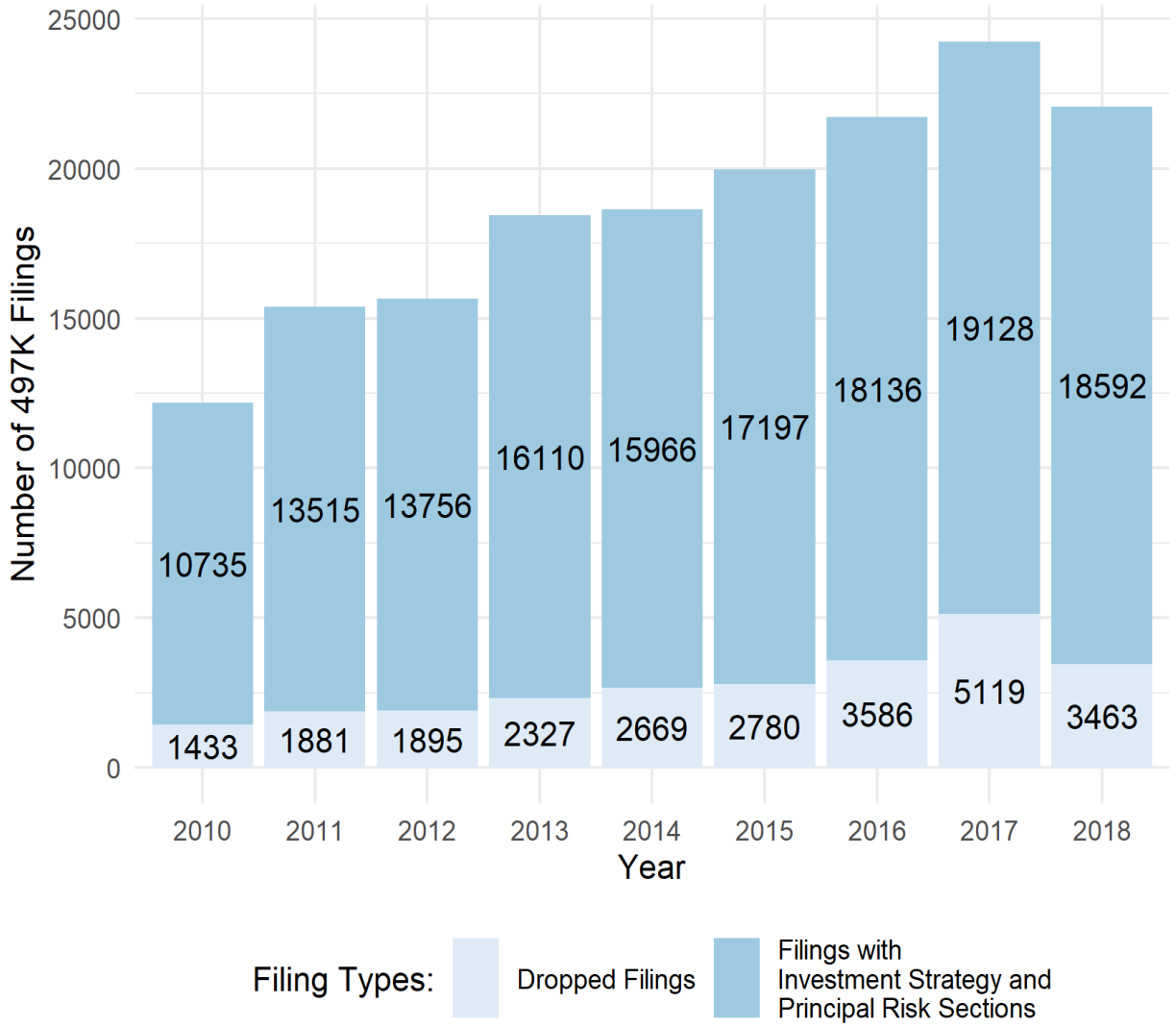
+ External Data

- CRSP
- WRDS returns, risk, etc.
- Thomson portfolio data

Process


SEC EDGAR
497 K Filings
Downloader 2010-2018 + 2019-2020 SEC XBRL
497K Repository
Filing Preprocessing
Drop irrelevant filings
Text preprocessing: Strip HTML tags Remove ?/'"&^%$#*@#
ID & scrape relevant section
Extracted Text
Investment Strategies
Principal Risks
Textual Analysis
-WC
- Parts of speech
- Dictionary
- Tone
Topic Modeling
STM
LDA
Dynamic TM
Compliance with mandatory text
Relationship btwn Risk & Strategy
Merge with external data

Corpus



CRSP Category	% of filings
Domestic Equity	23.5%
Foreign Equity	11.7%
Fixed Income	8.9%
Index	8.7%
M	23.2%
Other	23.8%

Duplicates dropped

Mutual Fund Reg Tech

3 case studies

1. Emerging Risks

Mapping the evolving risk landscape with topic modeling

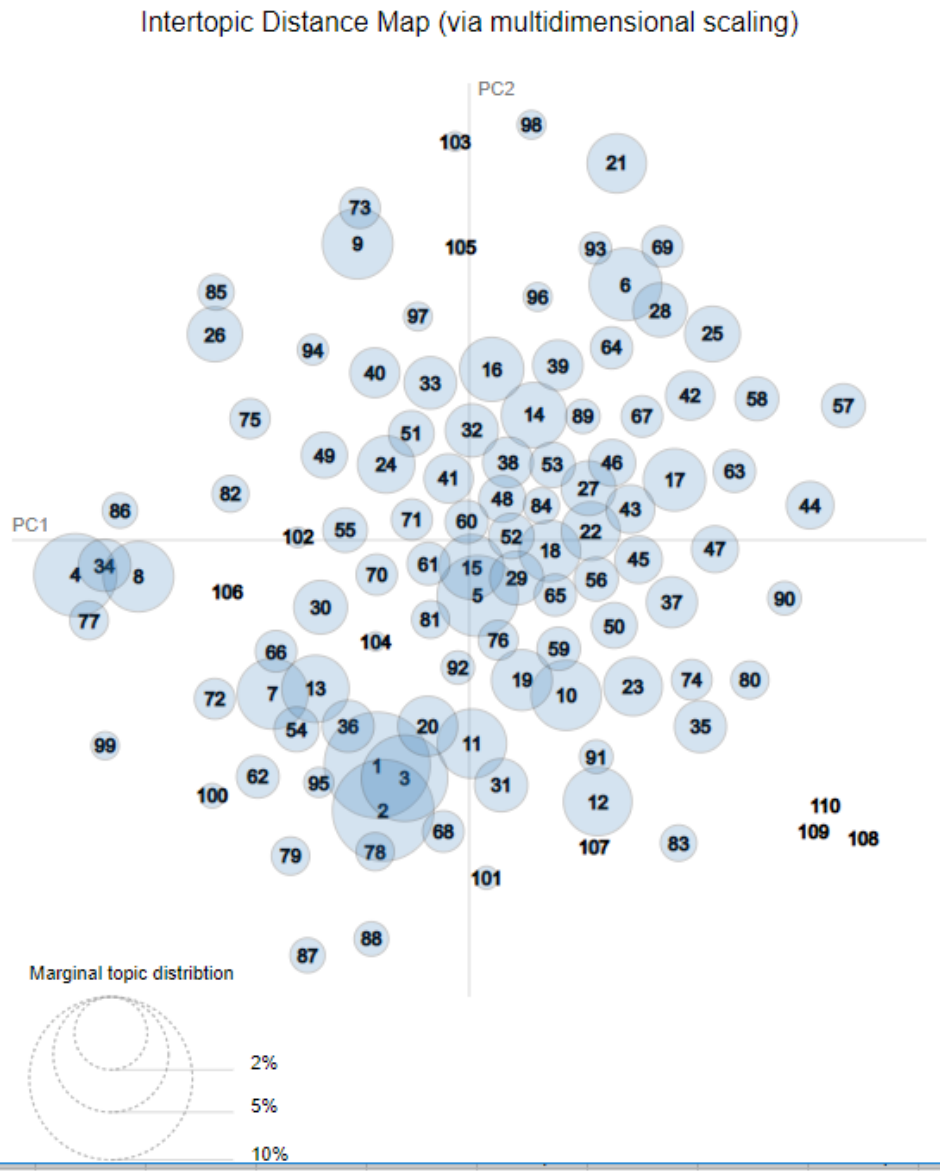
Plain English

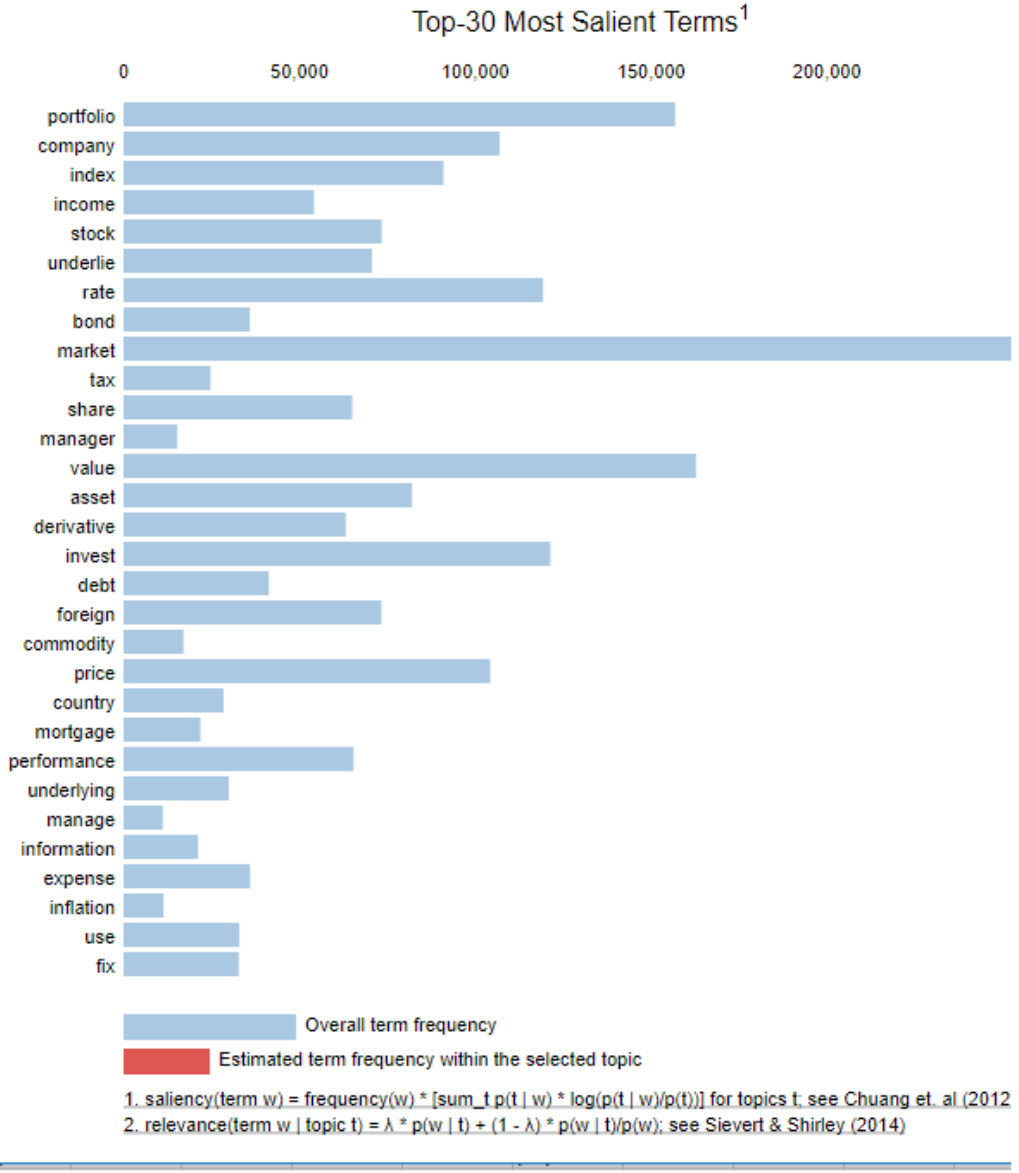
Disclosures have gotten longer and more complex

Tone & Sentiment

Reading between the lines to discover sentiment

Identifying Emerging Risks



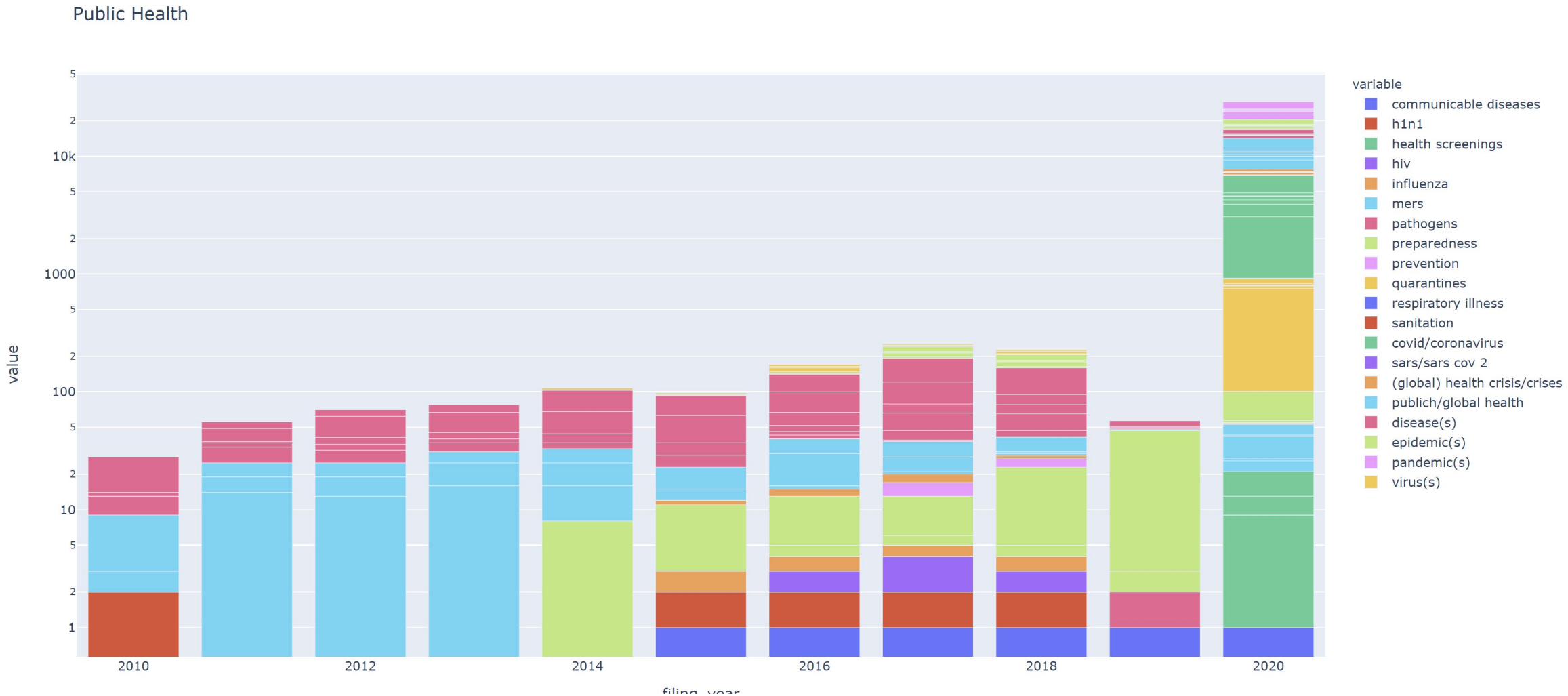


Monitoring Emerging Risks

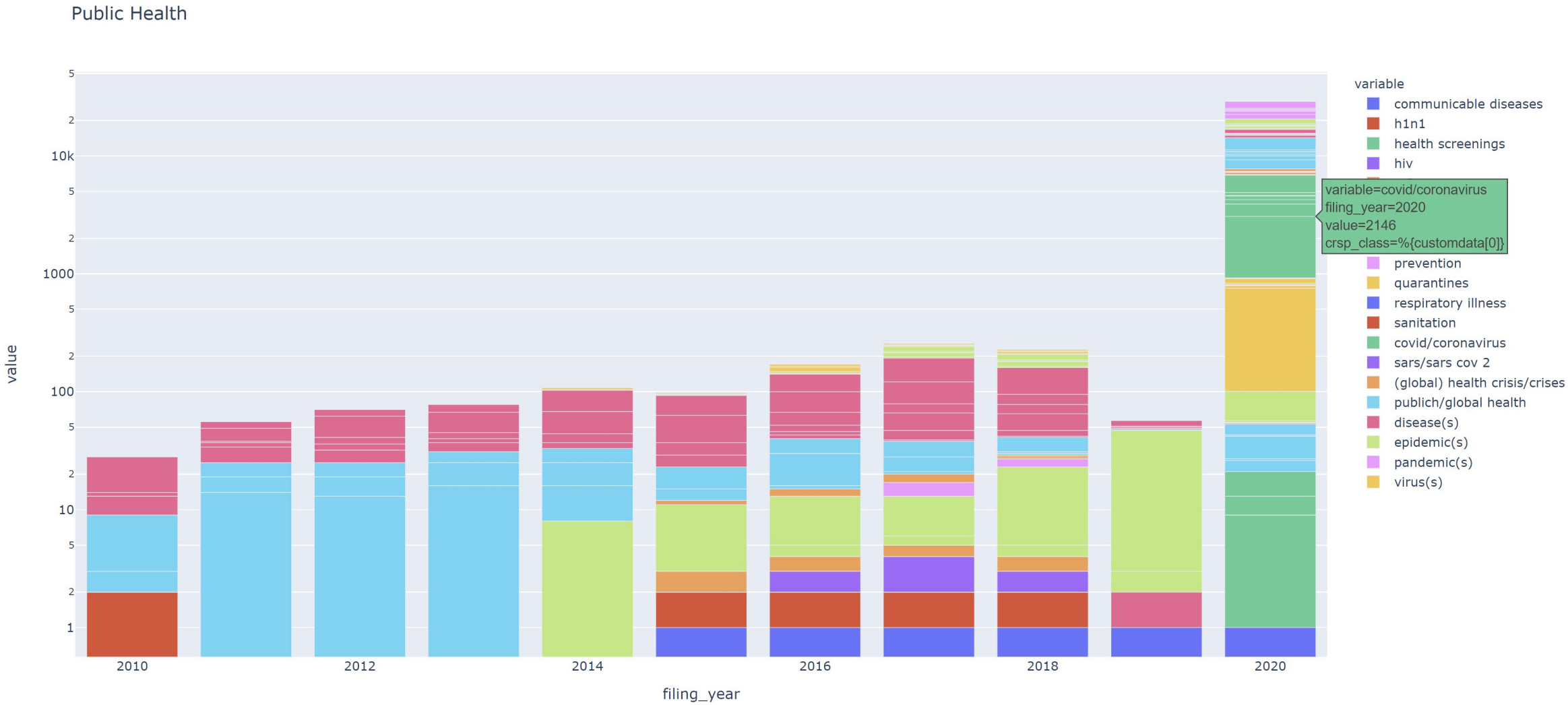
Effective immediately, the following is added to the Prospectus section Summary Prospectus-Principal Investment Risks of each XX FUND: Market Turbulence Resulting from **COVID-19** An outbreak of **COVID-19**, a respiratory disease caused by a novel coronavirus, has negatively affected the worldwide economy, as well as the economies of individual countries, the financial health of individual companies and the market in general in significant and unforeseen ways. The future impact of **COVID-19** is currently unknown, and it may exacerbate other risks that apply to XX FUND. Any such impact could adversely affect a XX FUND performance, the performance of the securities in which the XX FUND invests and may lead to losses on your investment.

- Few public health disclosures 2010-2018
- No pandemic disclosures pre-2020
  - Covid-19 additions to risk disclosures starting in 2020

Public Health Keyword Counts *(Log*)*



Public Health Keyword Counts *(Log*)*



Public Health Keywords

- Identified by manual & computational review
- 31 keywords (includes singular and plural)

communicable diseases	disease	quarantine	epidemics	health crisis	hiv
influenza h1n1	preparedness	public health	sanitation	virus	health screenings
covid 19 covid coronavirus	respiratory illness	pandemic	prevention	global health global health crisis	epidemic
mers	sars sars cov 2	pathogens			

Current Market Conditions

IM Guidance Update

MARCH 2016 | No. 2016-02

FUND DISCLOSURE REFLECTING RISKS RELATED TO CURRENT MARKET CONDITIONS

The staff is issuing this Guidance Update in order to foster investor protection by reminding mutual funds, exchange traded funds, and other registered investment companies of the importance to investors of full and accurate information about fund risks, including risks that arise as a result of changing market conditions. In particular, the staff believes that funds should review their risk disclosures on an ongoing basis and consider whether these disclosures remain adequate in light of current conditions.

Background: Importance of Risk Disclosure; Previous Commission and Staff Statements

Clear and accurate disclosure of the risks of investing in funds is important to informed investment decisions and, therefore, to investor protection. The Commission has long recognized the importance to investors of fund risk disclosure,[1] and the Commission's disclosure requirements reflect that recognition. A mutual fund, for example, is required to summarize the principal risks of investing in the fund, including the risks to which the fund's portfolio as a whole is subject and the circumstances reasonably likely to affect adversely the fund's net asset value, yield, and total return, in both its summary prospectus and at the front of its statutory prospectus.[2] In addition to this summary

Division of Investment Management Coronavirus (COVID-19) Response FAQs

The staff of the Division of Investment Management has prepared the following responses to questions about funds, advisers, and certain institutional investment managers that file Form 13F affected by COVID-19. These responses represent the views of the staff of the Division. They are not a rule, regulation, or statement of the Securities and Exchange Commission ("SEC"). The SEC has neither approved nor disapproved this content. These responses, like all staff guidance, have no legal force or effect: they do not alter or amend applicable law, and they create no new or additional obligations for any person.

For additional information concerning the SEC's response to COVID-19, please see the information available here (which provides an update on the SEC's targeted regulatory relief) and here (which describes the SEC's COVID-19 response generally).

I. Contacting the Division of Investment Management

Question I.1.

Q. How should a fund, SEC-registered investment adviser or exempt reporting adviser contact the staff if it has questions or concerns related to impacts of COVID-19 on its operations or compliance?

A. For questions regarding Form ADV, including related questions about temporary relief the SEC has provided, please email IARDLive@sec.gov.

Mutual Fund Reg Tech

3 case studies

Emerging Risks

Mapping the evolving risk landscape with topic modeling

2. Plain English

Disclosures have gotten longer and more complex

Tone & Sentiment

Reading between the lines to discover sentiment

Descriptive Findings: Summary

- **Funds are good at "check the box compliance"**
  - SAI & risk of loss statements
  - Mirrored derivative disclosures in strategy & risk sections

- **"Spirit" of disclosure is harder to meet**
  - Growing disclosure length driven by risk section
  - High readability scores ~ college or higher to read
  - Other indicators of complexity– words, tense, passive verbs

- **Variations in disclosure content & characteristics**

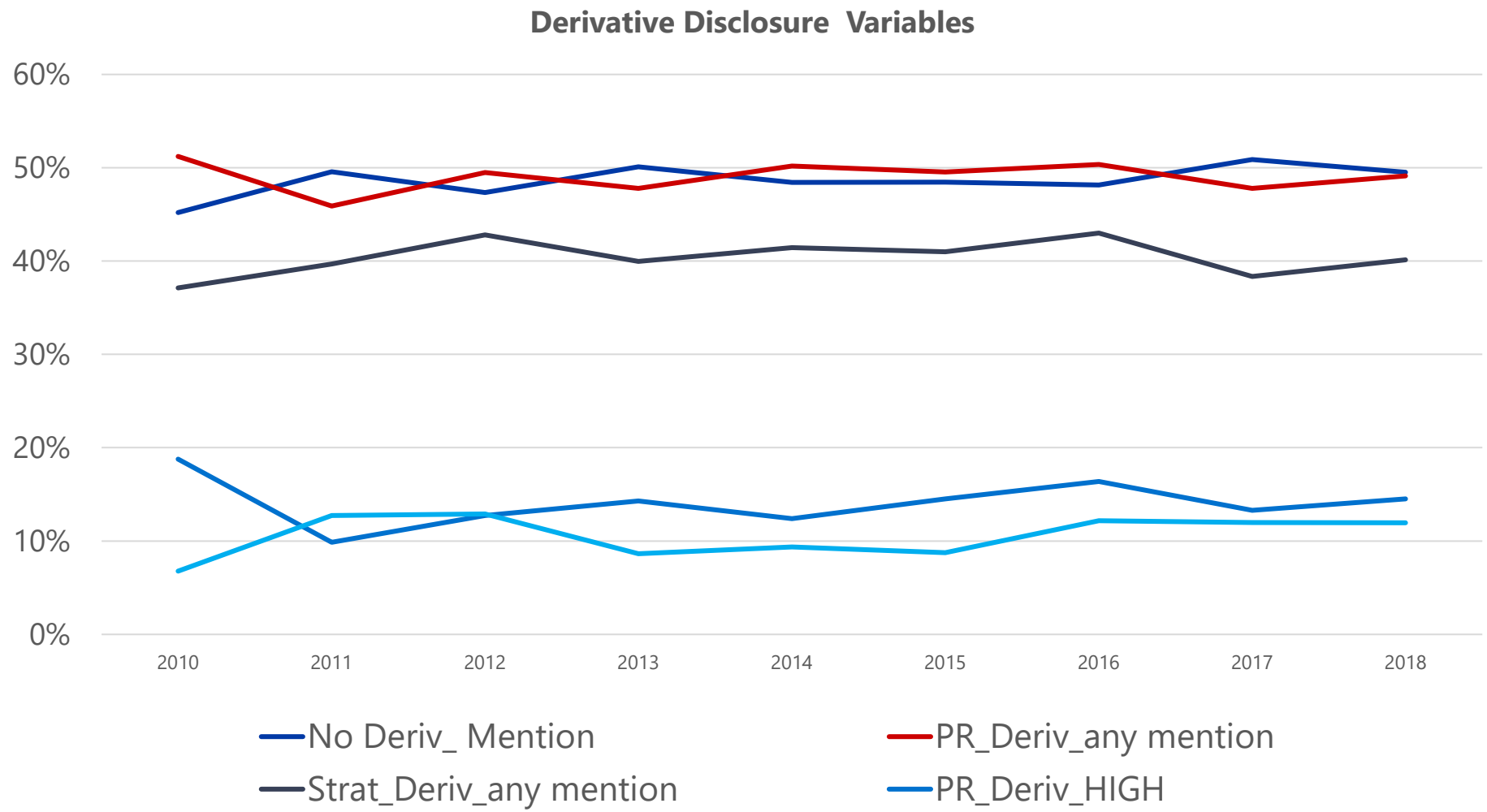
Check the Box Regulations

Total Rows	143,164
Compliance Under Money	89.53% 128,176
Compliance Under Investment	65.32% 93,514
TOTAL COMPLIANCE	90.82%

SAI Freq.	# PR	# IS
0	126,911	128,868
1	7,651	3,050
2	2,713	276
3	356	48
4	379	85

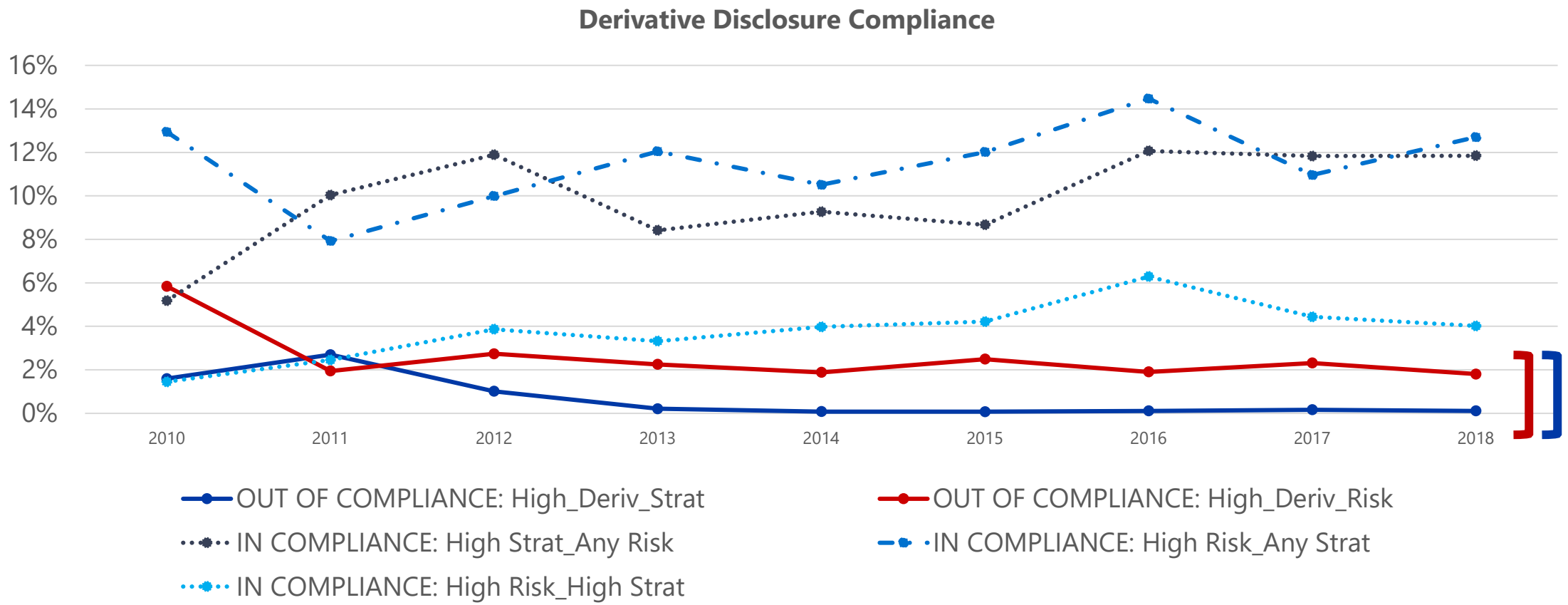
- Funds comply with check-the-box disclosures
  - Required statements; prohibitions on cross references; mirrored derivative disclosures

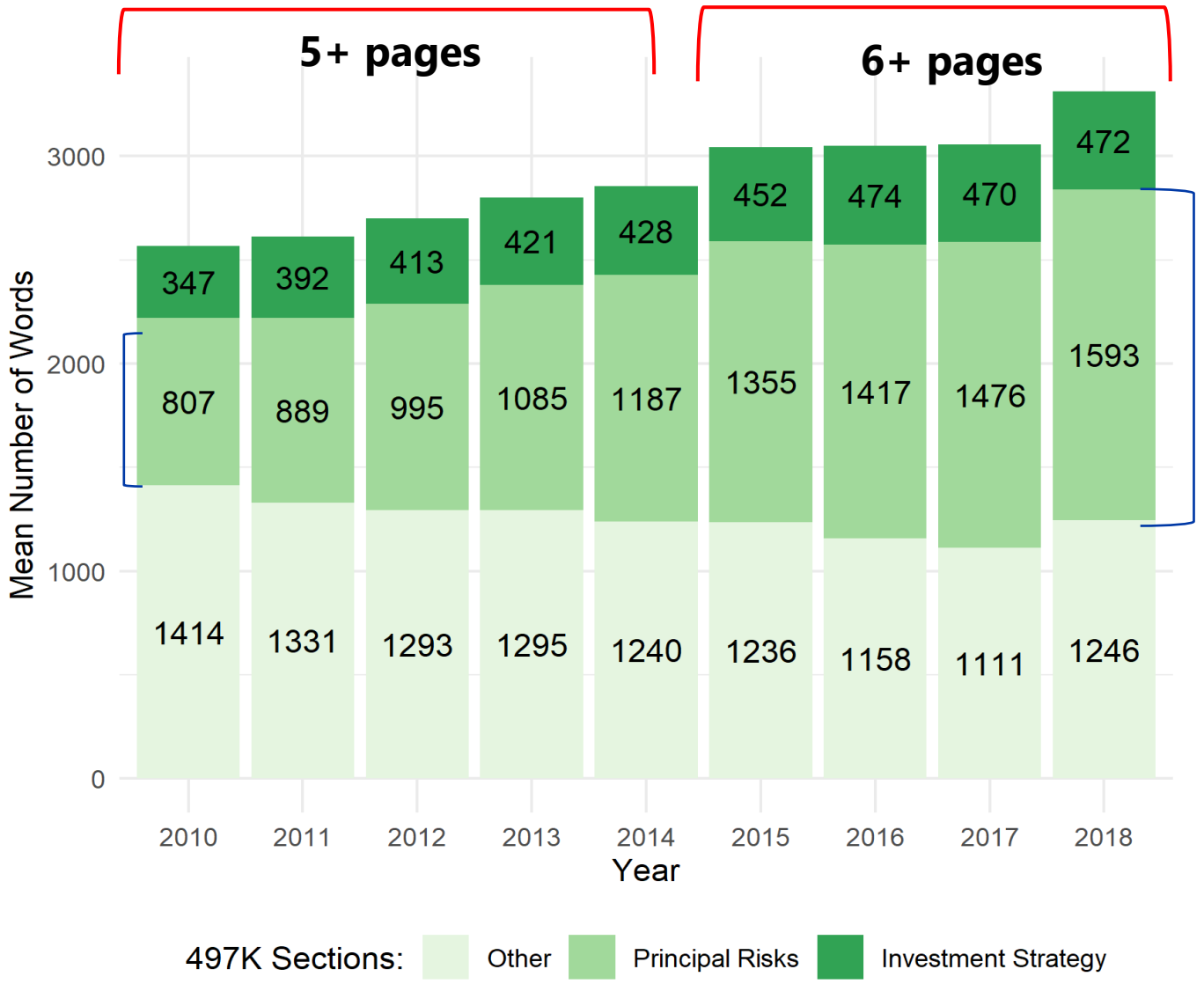
“Mirrored” Derivative Disclosures



“Mirrored” Derivative Disclosures

- > 96% funds comply with mirrored disclosures (out of compliance funds from high derivative risk sections with no derivative strategy discussion.)





Spirit of disclosures is harder to meet.

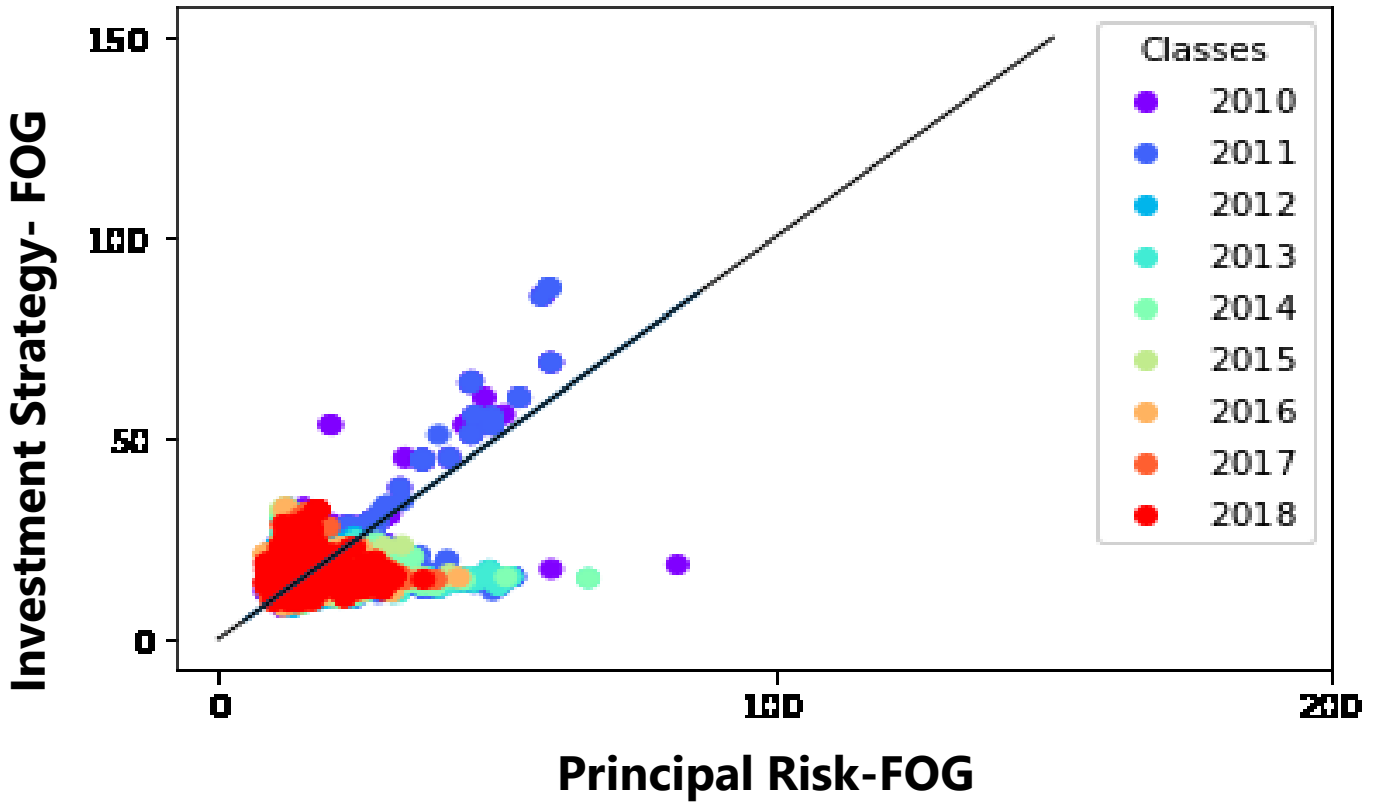
Readability

Disclosures written at college level reading

PR ave. grade level 14.1; IS ave. grade level 15.6 sections

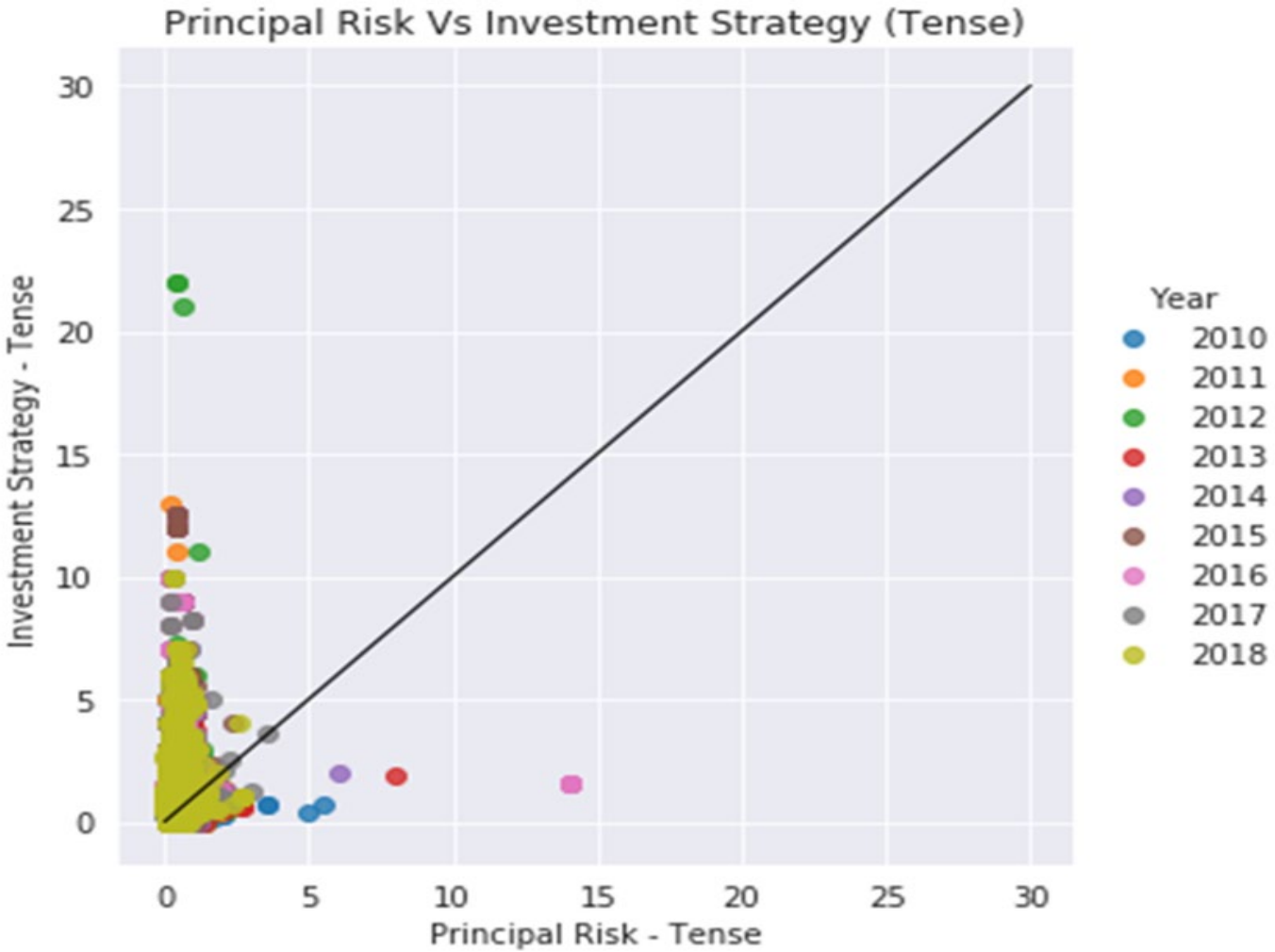
filing_date	fund_fam_name	stock_series#	fund_name	principal_strategies	principal_risks	year	strategy_fog	risk_fog	strategy_ari	risk_ari	strat_tense	s
2011-02-01	\t\t\ttiaa-cref funds\r\r\r\r\n	s000026505	Lifecycle Index 2035 Fund\r\r\r\r\n	the fund is a fund of funds that invests in in...	you could lose money over short or long period...	2011	20	22	23	24	0.0857143	
2011-12-20	\t\t\trussell investment co\r\r\r\r\n	s000001573	Russell U.S. Growth Fund\r\r\r\r\n	of the fund the fund invests primarily in co...	an investment in the fund, like any investment...	2011	15	16	15	15	0	
2011-11-01	\t\t\trussell investment co\r\r\r\r\n	s000001579	Russell Global Real Estate Securities Fund\r\r...	the fund seeks to achieve its objective by con...	an investment in the fund, like any investment...	2011	18	17	17	16	0	
	\t\t\tguinness		Guinness Atkinson Asia	the asia pacific	investing in this							

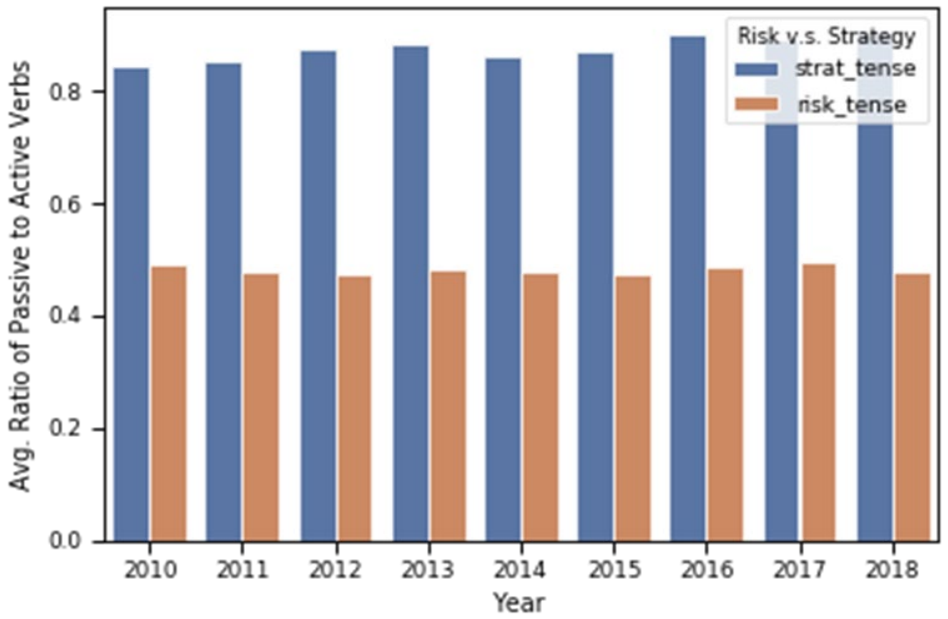
Readability



- IS sections = harder to read than PR sections in 76% of observation by an average 1.6 higher grade level
- PR sections are longer but "easier" to read ~ inherent tension brevity/ clarity in mandatory disclosures?

Tense





- IS uses more passive verbs than PR sections
- Results are time stable

Hypotheses

Semantic complexity obfuscates *(hides!)* **negative information**

Higher readability = lower returns & more risk

Semantic complexity may, however, be necessary because of complex investment strategies

Readability effects are mitigated by high derivative interaction terms.

Risk & strategy disclosures provide different information signals

Both affect returns

Principal risks will have a greater effect on volatility measures

Principal strategy will have a greater effect on diversification measures

Results

- **Semantic complexity obfuscates** *(hides!)* **negative information**
  - True!
    - Risk section readability grade level increase is associated with a decrease in fund return of 2.7% ***
  - Not fully explained by semantic complexity

Table 4: Risk Section Fixed Effects Regression Results.

	Return	Fund Risk		
	Return (mo.)	Std._Dev.	#_Sec.	%_Top_3
Readability(Fog)	-0.027****	-0.0403***	0.0083	-0.0590***
		(0.0131)	(0.0110)	(0.0118)
High_Deriv_Count	0.0675***	-0.1734***	0.0118	-0.1247***
	(0.0153)	(0.0205)	(0.0173)	(0.0185)
Readability×DerivCount	-0.0204**	0.0397***	0.0242**	-0.0496***
	(0.0101)	(0.0134)	(0.0114)	(0.0122)

Results

- **Risk and strategy disclosures provide different information signals**

	Prediction
Returns	Same
Volatility	Different Risk more important
Diversification #	Different Strategy more important
Diversification %	Different Strategy more important

Results

- **Risk and strategy disclosures provide different information signals** ✓ True!
  - Both affect returns
  - Principal risks will have a greater effect on volatility measures
  - Principal strategy will have a greater effect on diversification measures

☒ NOT True!

	Prediction	Results
Returns	Same	**Risk has bigger & negative effect**
Volatility	Different Risk more important	No real difference
Diversification= #	Different IS more important	Mixed results with interaction term
Diversification %	Different IS more important	Mixed results with interaction term

Interpretation & Recommendations

RISK SECTION

- **Finding:**
  - Bloating & obfuscation (e.g., growing length & negative returns)
- **Recommendation:**
  - Increased enforcement against kitchen sink disclosures

STRATEGY SECTON

- **Finding:**
  - Complexity is not from bloating or obfuscation, but may reflect necessary complexity
- **Recommendation:**
  - SEC specific guidance for IS content

Mutual Fund Reg Tech

3 case studies

3. Tone & Sentiment

Reading between the lines to discover sentiment

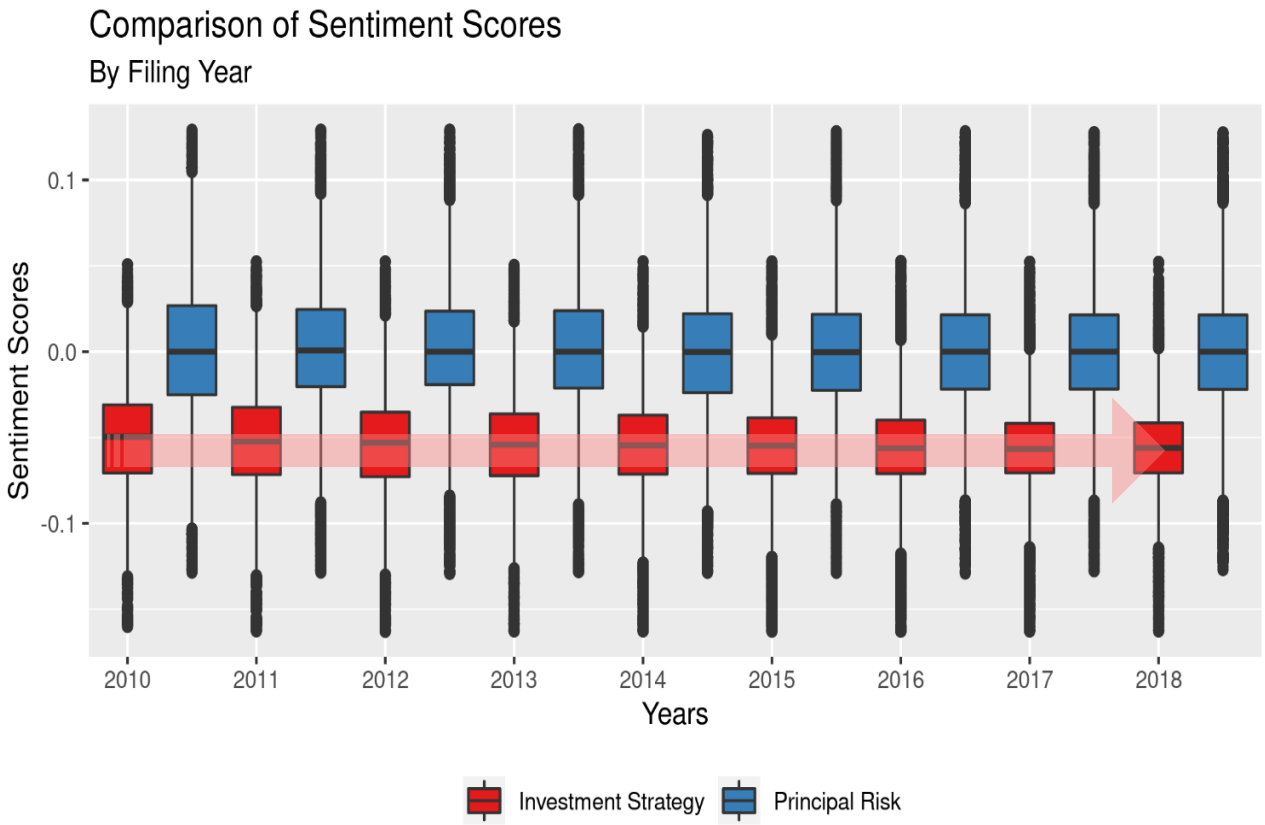
Emerging Risks

Mapping the evolving risk landscape with topic modeling

Plain English

Disclosures have gotten longer and more complex

Sentiment Scores



- Challenging intuition
- Investment strategy language becomes more similar and more negative over time

Georgia State University | COLLEGE OF LAW

Want **more**?

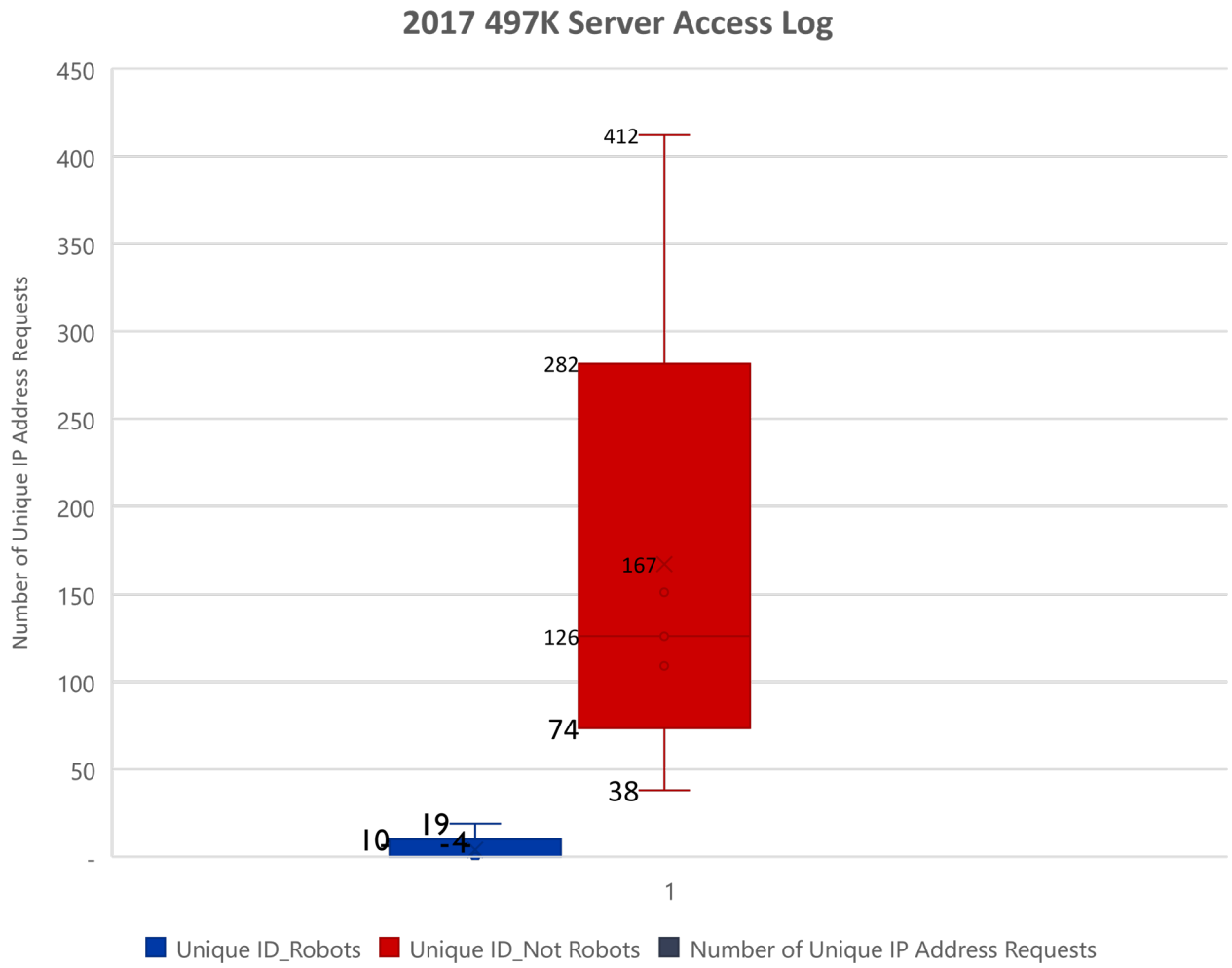
Professor Anne Tucker
GSU College of Law
amtucker@gsu.edu

https://sites.google.com/view/gsu-mutual-fund-research

The Promise & Peril of Plain English Mutual Fund Disclosures (SSRN)

Appendix

Server Access Log



- Some individual (non robot IPs) access 497k filings
- Average = 138 requires per filing, but undercounts real access to information

Back

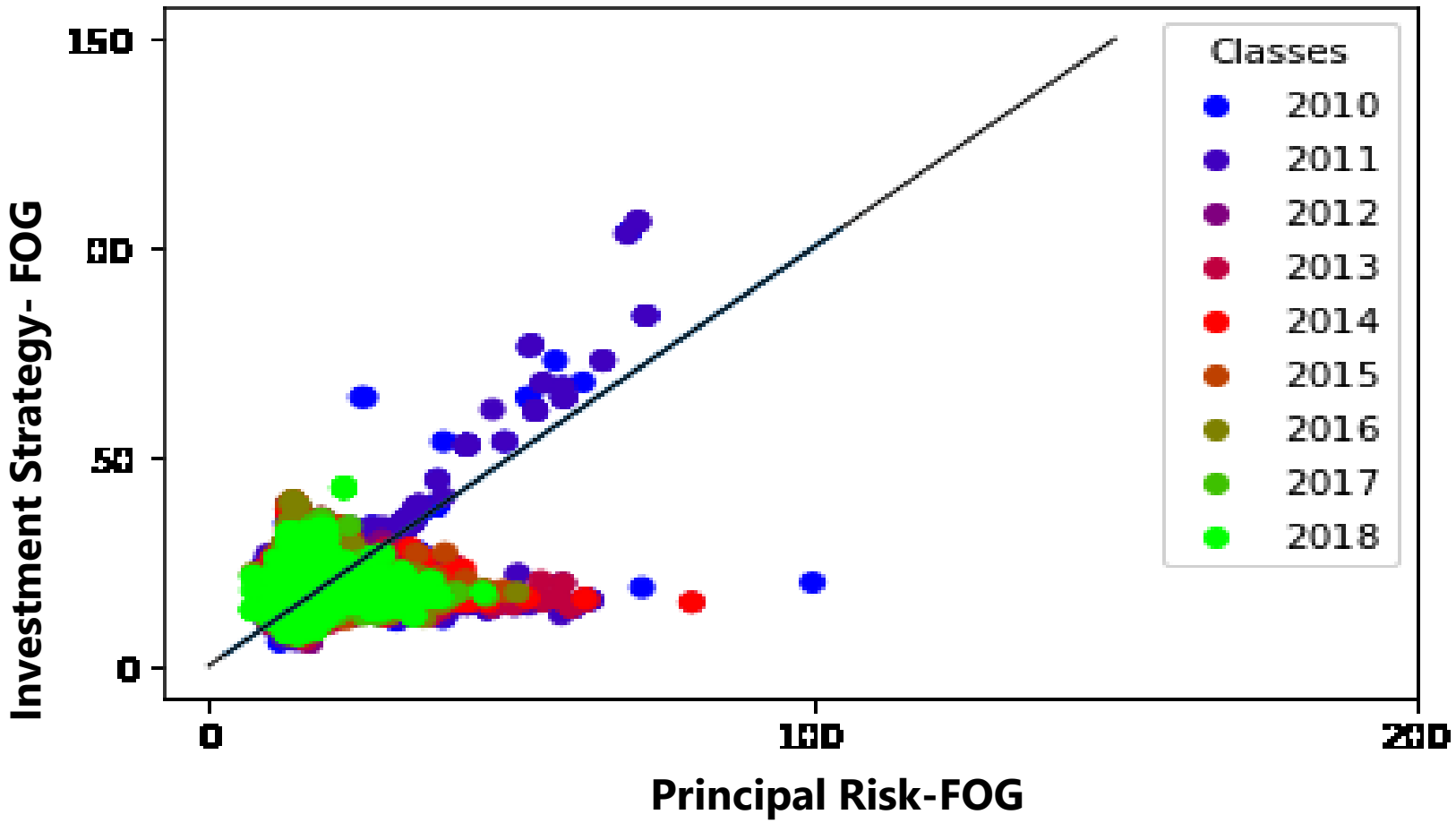
Spirit of regulation: Length

	2010	2011	2012	2013	2014	2015	2016	2017	2018
Corpus									
pages	5.25	5.34	5.52	5.73	5.84	6.22	6.24	6.25	6.77
Invest. Strategy									
pages	0.71	0.80	0.84	0.86	0.88	0.92	0.97	0.96	0.97
% of corpus	14%	15%	15%	15%	15%	15%	16%	15%	14%
Principal Risks									
pages	1.65	1.82	2.03	2.22	2.43	2.77	2.90	3.02	3.26
% of corpus	31%	34%	37%	39%	42%	45%	46%	48%	48%

Back

Intro Reg Tech Mutual Funds Emerging Risks Plain English Sentiment **Appendix**

Readability: ARI IS vs. PR scores

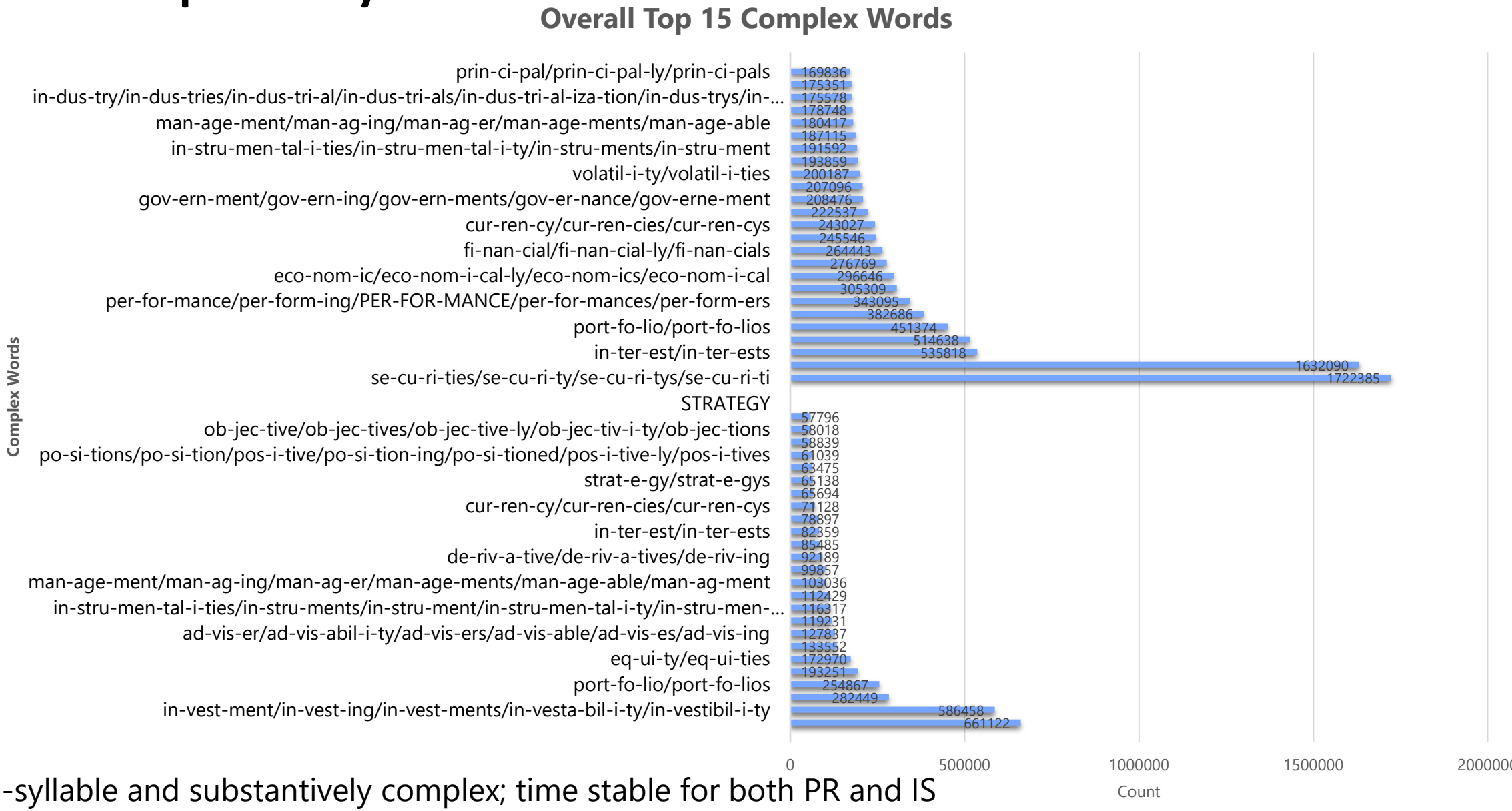


- Investment strategy ARI scores are larger than the Risk ARI scores in 70% of observations.
- IS disclosures are harder to read (average 1.5 higher grade level) than principal risk disclosures for the same fund, a result that is statistically significant at $p < 0.001$.

Back

Intro Reg Tech Mutual Funds Emerging Risks Plain English Sentiment **Appendix**

Complexity



- Multi-syllable and substantively complex; time stable for both PR and IS

Dictionary: Sentiment framework

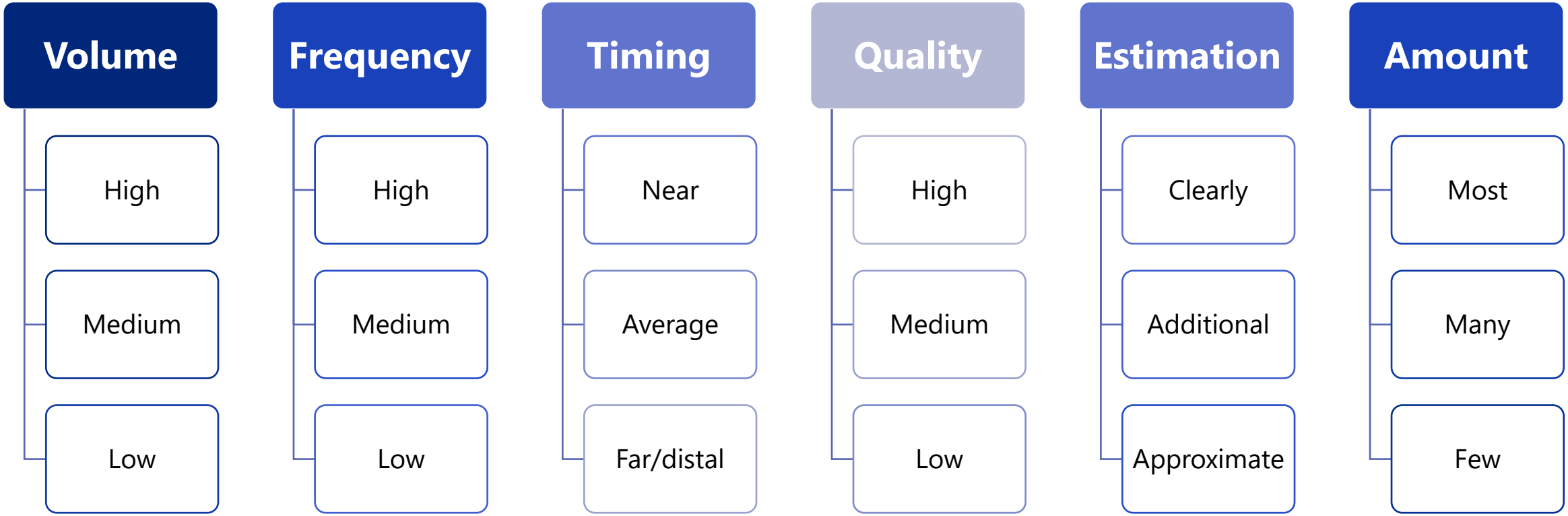
Loughran McDonald Dictionary

Positive | Negative | | Litigious | Modal | Uncertainty | | Strong | Weak

Tucker_Smelcer_Xia Scoring Framework/Dictionary

POLARITY			SUBJECT	VOLUME					
Positive	**Negative**	**Negator**	**Legal**	**Modal** (Scored based on LM)		**Uncertainty** Higher uncertainty = strong; less = weak		**Degree** (high volume/freq./quality = strong; low= weak; middle= moderate)	
+2	**-2**	**Reverse polarity**	**-1**	Strong (i.e., must)	*1.25 in same direction	Strong (i.e., arbitrary)	*.25 in same direction	Strong (i.e., always/best)	*1.25 in same direction
				Moderate (i.e., ought)	*.75 in same direction	Moderate (i.e., assume)	*.75 in same direction	Moderate (i.e., risky/generally)	*.75 in same direction
				Weak (i.e., seem)	*.25 in same direction	Weak (i.e., approx.)	*1.25 in same direction	Weak (i.e., seldom/poor)	*.25 in same direction
				Single word score capped for these three categories to ie., RISK is both uncertainty and degree but can only be scored once here.					
				If negative, add to negative dictionary and have the score reflected. i.e., RISK is negative so given a volume score + polarity score					
* indicates a multiple so you X1.25 or X.75, etc.									

Dictionary: Scoring Framework



Georgia State University | COLLEGE OF LAW

TSX vs. LM Dictionaries

TSX Unique Words (not in LM)	766
LM Unique (not in TSX)	2426
Common Words (TSX/LM)	1443
TSX % Unique (not in LM)	34.68%
LM % Unique (not in TSX)	62.70%
TX Common Words % (with LM)	65.32%
LM Common Words % (with TSX)	37.30%
TSX words frequency in corpus	35,304,414.00
LM words frequency in corpus	18,321,156.00
TSX unique frequency in corpus	18,071,481.00
LM unique frequency in corpus	1,088,223.00

Sentiment Scores: Summary Statistics

Disclosure Text	N	Mean	Std. Dev.	Min	1st Q.	Med.	3rd Q.	Max
Invest. Strategy	76,363	-0.055	0.029	-0.163	-0.071	-0.055	-0.038	0.053
Principal Risk	76,582	0.0003	0.038	-0.129	-0.022	0.000	0.023	0.129

- IS scores are more negative than PR scores (counter intuitive?)
- correlation between the IS and PR scores is positive and statistically significant, but very small ($r = 0.0502$, $p < 0.0001$).