Supplementary Information

The present and future of sustainability disclosure in equity investment funds'

pre-contractual documents:

mapping ESG discourse through STM

Content summary

This PDF file includes:
• Supplementary text;
• Supplementary Figures (Figures S1 to S4);
• Supplementary Tables (Tables S1 to S6);

Supplement Section S1 contains additional materials related to the results. It includes: the values of the parameters used for the final STM estimation (Table S1); two summary tables of the top 5 tokens ordered by FREX score (in Table S2) or by probability (in Table S3) for each topic and by Article class; a table containing estimated topic propensity covariate effects (Table S4); four plots (Figure S1) representing the Topic networks of four sustainability-related topics, for Article-8 Vs Article 9 KIIDs.

Robustness checks and description of parameters used in the STM are described in Section S2 of the Supplement. The section includes: the values of the parameters used in experiments for choosing the number of topics K (Table S5); two Figures, one representing the STM's Lower bound and Heldout likelihood (Figure S2) as a function of K, and the other one representing the Semantic Coherence and Exclusivity (Figure S3) as a function of K.

Finally, details about the dataset, methods and code are given in Section S3, which contains:
a detailed description of the data selection and collection process; a set of wordclouds representing the Frequency of geographical funds composition by Article class and Asset manager (Figure S4), the list of stopwords removed from the text during the KIIDs preprocessing step (Table S6); additional details and references related to the methods, in particular with respect to the Structural Topic Model.

S1 Supplementary Results

The results in the paper and in this supplement have been obtained using the parameter values described in Table S1 for estimating the STM. For not listed parameters, *stm* (V.1.3.6) library's default values have been used. The model converged after 71 iterations. The value of the Lower Bound at the final iteration is -4989081

Table S1: Parameter values used in the final STM estimation.

parameter	value
K	35
Topic prevalence	~ ARTICLE + ASSET_MANAGER + TYPE + RISK + GEO_AREA
Topic content	~ ARTICLE
max.em.its	250
em.tol	1e-06
init.type	"Spectral"

Table S2: Topics summary, with top 5 tokens ordered by token FREX score, by topic and by Article.

Topic	Topic Prop.	Article 6 (top 5 tokens by FREX)	Article 8 (top 5 tokens by FREX)	Article 9 (top 5 tokens by FREX)
1	0.3%	vary, fees, time, charged, fees->charged	fees, vary, charged, fees->charged, one	vary, fees, periodically, charged, time
2	4.01%	ishares, half, yearly, plc, bgf	bf1, bgf, arrangements, along, umbrella	bgf, within->bgf, umbrella, half, yearly
3	4.39%	launch, launch->date, class->launch, data->available, insufficient->data	launch, launch->date, insufficient->data, data->available, class->launch	launch, launch->date, annual->past, class->launch, insufficient->performance
4	2.04%	assessed, assets->directly, asset->classes, indirectly, outside	assets->directly, wide->range, safety, wide, monitoring	payments, transfer, delivery, delayed, delayed->delivery
5	0.51%	eur->hgd, hgd, gbs, solactive->gbs, gbs->developed	model, business->model, investment->case, manager->aims, analysing->esg	reference->indicator, control, pollution->control, energy->efficiency, across->ten
6	1.37%	share->class, share, class, one, benchmark:euro	launched, excluded, performance->data, one->year, exit->charges	one->year, blank, intentionally->blank, left, left->intentionally
7	1.24%	allows, assets->representing, proportion, representing, index->constituents	minimum->volatility, msci->world, world, volatility->index, world->minimum	index->constituents, regulation->aims, index->based, allows, assets->representing
8	0.09%	invest, derivatives, involved, vary->periodically, net->dividend	derivatives, invest, involved, time, vary->periodically	invest, derivatives, vary->periodically, one, fund
9	6.05%	ongoing->charge, exact, addendum, last->year, ending->december	ongoing->charge, addendum, financial->year, charges->made, country->specific	exact, exact->charges, charges->made, include->detail, financial->year
10	2.81%	actively->manages, manages, team->actively, substantial->business, risk->exposure	various, reference->benchmark, integrates->sustainability, financial->product, sustainability->factors	footprint, carbon->footprint, asset->weighted, controversial, footprint->intensity
11	2.23%	fundname->investment, funds->s.a, supervised, sicav, home	sicav, home, home->page, fundname->investment, funds->s.a	fundname->investment, funds->s.a, supervised, sicav, home
12	12.53%	adequately, adequately->captured, mean->risk, captured, risk->typically	adequately, adequately->captured, default, materially->relevant, vis	adequately, adequately->captured, mean->risk, captured, factors->include
13	1.23%	managed->based, june, fundname->funds, july, russian	absorb, absorb->fundname, adopted, created, fee->structure	european->players, players, top->european, top, master
14	13.54%	revenue, another->collective, charge->paid, pay->fees, stock->brokers	figure, maximum->figures, might->pay, charges->figure, charges->shown	advisor, actual, actual->entry, financial->advisor, charge->paid
15	2.06%	currency->exchange, get->back, originally, originally->invested, rates	get->back, originally, originally->invested, result, level	rise, fall, difficult, accumulation->eur, shareclass
16	0.47%	target->benchmark, deducted, efficiently, kong->stock, three	morningstar, distributors->premises, internet->website, premises, manager->invests	comparator->benchmark, target->benchmark, efficiently, manager->invests, regions
17	2.82%	results, future->results, investment->returns, reliable->indicator, class->net	applicable->entry, appointed, appointed->representative, events, applicable	useful, useful->indication, reliable->indicator, history, performance->history
18	2.18%	use->financial, ensure, specific->investment, appropriate->given, active->risk	main->business, indirect->exposure, appropriate->given, active->risk, determined->appropriate	sustainable->energy, nutrition, oil, fdis, energy->companies
19	4.33%	associated->risks, failure, material->risks, lowest->risk, investment->risks	associated->risks, investment->risks, shareholder, techniques, risk->potentially	simulated->net, associated->risks, material->risks, failure, lowest->risk

Table S2 *continued*

Topic	Topic Prop.	Article 6 (top 5 tokens by FREX)	Article 8 (top 5 tokens by FREX)	Article 9 (top 5 tokens by FREX)
20	0.16%	epra, ftse->epra, nareit, developed->index, index->ftse	real, real->estate, estate, ftse, estate->investment	ftse, developed->europe, nareit->developed, epra->nareit, nareit
21	2.07%	equity->fund, met, calculated->nav, excluding->initial, enhanced->equity	nordic, calculated->nav, excluding->initial, met, msci->nordic	global->climate, historical->performance, calculated->nav, excluding->initial, nav->excluding
22	4.21%	emea, privacy, privacy->policy, another->share, emea->privacy	privacy, privacy->policy, emea->privacy, another->share, asset->management	emea, privacy, privacy->policy, emea->privacy, another->share
23	2.09%	actively->managing, mainly->invests, accessible, fund->mainly, reducing->risks	actively->managing, finance->disclosure, sustainable->finance, accessible->via, accessible	data->quality, methodologies, growth->prospects, appear, distributing->dividends
24	1.26%	subscribe, lending->transactions, appropriate->considering, assets->eligible, assets->predominantly	sri, issuer->capped, low->carbon, carbon->select, sri->low	etf->plc, irl, subscribe, exchanges, climate->paris
25	1.77%	benchmark->index, float, adjusted, certain->eligible, eligible->third	started, fundname->robur, robur, 1st, 1st->january	benchmark->index, 1st, 1st->january, 31st, 31st->december
26	1.34%	companies->anywhere, built, level->analysis, research->team, stock->level	improving, demonstrate->improving, systematically->reinvested, bank->business, business->days	themebot, change->solutions, address->climate, engagement, develop->solutions
27	3.43%	limited->value, calculated->net, diagram, performances->displayed, april	limited->value, pioneer->u.s, calculated->net, diagram, performances->displayed	limited->value, calculated->net, diagram, performances->displayed, fundname->index
28	5.1%	fundname->luxembourg, periodical, periodical->reports, applying, connection	fundname->luxembourg, periodical->reports, address, periodical, allée	fundname->luxembourg, periodical->reports, periodical, ucits, luxembourg
29	1.03%	grade, mmis, specified, debt, overnight	ties, certain->sectors, exclude->direct, whereby, apply->fundname	market->capitalisation, capitalisation, equity->securities, small->market, transport
30	4.25%	fund->management, additional->information, convert, fundname->fund, online	fil->investment, fil, obtained->free, tax->advisor, full->entry	management->luxembourg, online, obtained->free, easy, fundname->easy
31	3.39%	percentage->change, expressed, deduction, annual->performance, excluded	end, reliable->indicator, chart->shows, percentage, unit	calendar, calendar->year, shows, chart->shows, percentage->change
32	0.32%	order, reduce, portfolio, adviser, please->see	mmis, maximum->fee, conditions->none, running->costs, agencies	subscription, conditions->none, running->costs, units, conversion
33	3.32%	broad->discretion, meaningfully, meaningfully->different, main->part, broad	norms->based, assets->excluding, excluding->cash, sustainable->investments, social->objectives	positions, limited->number, sectors, ideas, usage
34	1.62%	market->liquidity, hours, maintain->market, makers, market->makers	higher->potential, different->share, equity->funds, fund->value, investment->within	paris, sample, companies->included, extra, composed
35	0.42%	international->markets, potentially, risky, fluctuate->considerably, investment->mainly	horizon, investment->horizon, justified, recommended->investment, rewards->potentially	considerably, horizon, investment->horizon, justified, recommended->investment

The FREX metric (Bischof and Airoldi, 2012) aims to identify words that are both frequent in a topic and exclusive to it, striking a balance between these two traits.

Table S3: Topics summary, with top 5 tokens ordered by token probability, by topic and by Article.

Topic	Topic Prop.	Article 6 (top 5 tokens by prob.)	Article 8 (top 5 tokens by prob.)	Article 9 (top 5 tokens by prob.)
1	0.3%	vary, fees, time, fund, charged	fees, vary, charged, one, time	fees, vary, time, charged, fund
2	4.01%	fund, funds, assets, ishares, fundname	fund, funds, fundname, assets, document	fund, bgf, funds, fundname, assets
3	4.39%	performance, launch, launch->date, date, past->performance	performance, launch, launch->date, past, past->performance	performance, launch, launch->date, date, past->performance
4	2.04%	fund, emerging, emerging->markets, index, markets	fund, index, investment, characteristics, invest	fund, markets, emerging, emerging->markets, assets
5	0.51%	europe, msci->europe, cap, mid, mid->cap	esg, fund, investment, manager, score	esg, fund, benchmark, reference->indicator, investment
6	1.37%	share, share->class, class, fund, one	performance, launched, charges, excluded, class	performance, launched, charges, class, share->class
7	1.24%	index, fund, net, prospectus, value	world, msci->world, world->index, hedged, return	index, fund, net, prospectus, value
8	0.09%	invest, derivatives, fund, investment, performance	invest, derivatives, time, fund, involved	invest, derivatives, one, fund, vary->periodically
9	6.05%	charges, charge, taken, year, share	charges, charge, share, share->class, class	charges, charge, taken, share, charges->taken
10	2.81%	fund, benchmark, investment, management, prospectus	fund, benchmark, investment, prospectus, value	fund, investment, benchmark, portfolio, carbon
11	2.23%	fundname, fund, sicav, share, investment	fundname, fund, share, sicav, remuneration	fundname, fund, sicav, share, remuneration
12	12.53%	risk, fund, risks, category, higher	risk, fund, risks, category, higher	risk, fund, risks, category, higher
13	1.23%	funds, fundname->funds, equity, june, fund	fund, investment, funds, fundname->funds, investment->policy	fund, european, fundname->funds, funds, top
14	13.54%	charges, fund, costs, taken, charge	charges, fund, taken, year, charge	charges, fund, costs, taken, charge
15	2.06%	fund, level, risk, investment, currency	fund, level, risk, investment, currency	europe, msci->europe, result, certain, fund
16	0.47%	fund, investment, china, value, including	fund, investment, benchmark, manager, investment->manager	fund, investment, companies, investment->manager, manager
17	2.82%	performance, past, benchmark, fund, past->performance	performance, past, past->performance, fund, benchmark	useful, useful->indication, performance, past->performance, past
18	2.18%	fund, investment, index, shares, fdis	fund, securities, investment, index, shares	fund, investment, esg, shares, securities
19	4.33%	risks, risk, fund, share, share->class	risks, risk, fund, share->class, class	risks, risk, fund, class, share->class
20	0.16%	ftse, epra, ftse->epra, nareit, global	real, estate, real->estate, ftse, global	ftse, nareit, epra, ftse->epra, nareit->developed
21	2.07%	index, performance, fund, net, share->class	index, performance, fund, return, net	index, performance, fund, eur, net
22	4.21%	fundname, policy, fund, management, funds	fundname, policy, fund, funds, management	fundname, policy, fund, management, funds

Table S3 *continues on next page*

Topic	Topic Prop.	Article 6 (top 5 tokens by prob.)	Article 8 (top 5 tokens by prob.)	Article 9 (top 5 tokens by prob.)
23	2.09%	fund, investment, management, portfolio, investments	fund, investment, management, portfolio, characteristics	fund, investment, management, companies, portfolio
24	1.26%	fund, index, fundname, shares, securities	index, fund, securities, fundname, currency	fund, index, etf, securities, shares
25	1.77%	shares, index, fund, market, securities	performance, past, past->performance, fund, calculated	performance, past->performance, past, fund, benchmark->index
26	1.34%	fund, benchmark, investment, portfolio, derivatives	fund, investment, esg, companies, benchmark	investment, companies, fund, solutions, change
27	3.43%	fund, fundname, reference, performance, index	fund, launched, performance, reference, euro	fund, reference, msci, launched, index
28	5.1%	luxembourg, ucits, fund, fundname, prospectus	luxembourg, ucits, fund, fundname, prospectus	luxembourg, ucits, fund, fundname, prospectus
29	1.03%	fund, securities, investment, shares, assets	fund, investment, shares, market, assets	fund, esg, investment, shares, market
30	4.25%	fund, luxembourg, management, fundname, available	fund, luxembourg, prospectus, available, share	fund, luxembourg, remuneration, fundname, prospectus
31	3.39%	performance, fund, chart, year, launched	performance, fund, year, index, charges	performance, fund, launched, year, chart
32	0.32%	portfolio, order, reduce, adviser, specific	taken, investment, fund, determined, based	maximum, taken, conversion, fund, investment
33	3.32%	fund, investment, derivatives, benchmark, companies	investment, fund, derivatives, companies, benchmark	sectors, positions, fund, limited->number, investment
34	1.62%	index, fund, shares, market, performance	fund, index, risk, companies, equity	index, fund, esg, companies, tracking
35	0.42%	investment, markets, potentially, fund, term	investment, markets, potentially, indicator, fund	investment, markets, potentially, indicator, fund

Table S4: Estimated effects of covariates on topic propensities (only covariates that are statistically significant at the 0.1 significance level have been included).

Topic	Covariate (affecting topic propensity)	estimated effect	std.error	statistic	p.value
1	ASSET_MANAGER=SWEDBANK	0,21383	0,02835	7,54221	0,00000
1	ASSET_MANAGER=UBS	0,01002	0,00502	1,99672	0,04591
1	TYPE (SECTION)=PAST_PERF	-0,00362	0,00217	-1,66702	0,09558
1	TYPE (SECTION)=RISK_AND_REW	0,00888	0,00251	3,53928	0,00041
2	ASSET_MANAGER=BLACKROCK	0,19170	0,00815	23,51100	0,00000
2	TYPE (SECTION)=INFO	0,19908	0,00746	26,68106	0,00000
3	ASSET_MANAGER=JP MORGAN	0,19039	0,00800	23,79251	0,00000
3	TYPE (SECTION)=PAST_PERF	0,21993	0,00763	28,83123	0,00000

4	ARTICLE= ART. 8	-0,00867	0,00507	-1,70893	0,08753
4	ASSET_MANAGER=FIDELITY	0,17492	0,01120	15,61162	0,00000
4	ASSET_MANAGER=SCHRODERS	0,05962	0,02369	2,51718	0,01186
4	TYPE (SECTION)=OBJ_AND_INV_POL	0,08960	0,00598	14,99193	0,00000
4	TYPE (SECTION)=RISK_AND_REW	0,00976	0,00542	1,79996	0,07193
5	ASSET_MANAGER=BLACKROCK	-0,03459	0,00396	-8,73235	0,00000
5	ASSET_MANAGER=BNP PARIBAS	-0,02416	0,00627	-3,85030	0,00012
5	ASSET_MANAGER=EURIZON	-0,03037	0,00976	-3,11250	0,00187
5	ASSET_MANAGER=FIDELITY	-0,02888	0,00529	-5,46340	0,00000
5	ASSET_MANAGER=JP MORGAN	-0,03215	0,00396	-8,12294	0,00000
5	ASSET_MANAGER=NORDEA	-0,03330	0,00484	-6,87436	0,00000
5	ASSET_MANAGER=SCHRODERS	-0,03528	0,00902	-3,90875	0,00009
5	ASSET_MANAGER=SWEDBANK	-0,03225	0,01010	-3,19371	0,00141
5	ASSET_MANAGER=UBS	-0,03097	0,00523	-5,92584	0,00000
5	TYPE (SECTION)=OBJ_AND_INV_POL	0,01822	0,00372	4,89427	0,00000
5	TYPE (SECTION)=PAST_PERF	0,01855	0,00356	5,21450	0,00000
6	ARTICLE= ART. 8	0,00903	0,00490	1,84307	0,06538
6	ARTICLE= ART. 9	-0,01129	0,00622	-1,81387	0,06976
6	ASSET_MANAGER=BLACKROCK	0,08710	0,00581	14,97956	0,00000
6	TYPE (SECTION)=PAST_PERF	0,06059	0,00561	10,79982	0,00000
6	GEO_AREA=Continental Europe	-0,11618	0,07050	-1,64806	0,09941
7	ARTICLE= ART. 8	-0,01944	0,00467	-4,16117	0,00003
7	ARTICLE= ART. 9	-0,01649	0,00643	-2,56691	0,01029
7	ASSET_MANAGER=BLACKROCK	-0,04983	0,00593	-8,40773	0,00000
7	ASSET_MANAGER=BNP PARIBAS	-0,02699	0,00904	-2,98722	0,00283
7	ASSET_MANAGER=EURIZON	-0,03497	0,01567	-2,23133	0,02571
7	ASSET_MANAGER=FIDELITY	-0,03757	0,00788	-4,77092	0,00000
7	ASSET_MANAGER=JP MORGAN	-0,03623	0,00546	-6,63259	0,00000
7	ASSET_MANAGER=NORDEA	-0,02262	0,00782	-2,89284	0,00384
7	ASSET_MANAGER=SCHRODERS	-0,03558	0,01349	-2,63692	0,00839

7	ASSET_MANAGER=SWEDBANK	-0,04013	0,01651	-2,43058	0,01511
7	ASSET_MANAGER=UBS	-0,03464	0,00790	-4,38460	0,00001
7	TYPE (SECTION)=OBJ_AND_INV_POL	0,06117	0,00574	10,66202	0,00000
7	GEO_AREA=Pacific region, excluding Japan	0,10398	0,05358	1,94063	0,05236
8	ASSET_MANAGER=FIDELITY	0,01132	0,00350	3,22986	0,00125
8	ASSET_MANAGER=SCHRODERS	0,01119	0,00586	1,90938	0,05627
8	ASSET_MANAGER=SWEDBANK	0,04375	0,02058	2,12584	0,03357
8	TYPE (SECTION)=OBJ_AND_INV_POL	0,00891	0,00173	5,14278	0,00000
8	TYPE (SECTION)=RISK_AND_REW	0,00225	0,00117	1,91508	0,05554
9	(Intercept)	0,25178	0,08186	3,07563	0,00211
9	ASSET_MANAGER=JP MORGAN	0,19148	0,00877	21,84353	0,00000
9	ASSET_MANAGER=NORDEA	0,12364	0,01237	9,99531	0,00000
9	ASSET_MANAGER=SCHRODERS	0,05154	0,02640	1,95217	0,05098
9	ASSET_MANAGER=SWEDBANK	0,11137	0,03262	3,41366	0,00065
9	TYPE (SECTION)=INFO	-0,29987	0,00850	-35,26636	0,00000
9	TYPE (SECTION)=OBJ_AND_INV_POL	-0,30066	0,00841	-35,73160	0,00000
9	TYPE (SECTION)=PAST_PERF	-0,30057	0,00841	-35,73975	0,00000
9	TYPE (SECTION)=RISK_AND_REW	-0,30068	0,00843	-35,66463	0,00000
10	ARTICLE= ART. 8	0,04868	0,00584	8,33191	0,00000
10	ARTICLE= ART. 9	0,03868	0,00770	5,02306	0,00000
10	ASSET_MANAGER=BLACKROCK	-0,08514	0,00723	-11,77100	0,00000
10	ASSET_MANAGER=BNP PARIBAS	-0,11050	0,01135	-9,73978	0,00000
10	ASSET_MANAGER=FIDELITY	-0,10756	0,01045	-10,28972	0,00000
10	ASSET_MANAGER=JP MORGAN	-0,10716	0,00737	-14,53243	0,00000
10	ASSET_MANAGER=NORDEA	-0,09563	0,00917	-10,42844	0,00000
10	ASSET_MANAGER=SCHRODERS	-0,09924	0,01826	-5,43455	0,00000
10	ASSET_MANAGER=SWEDBANK	-0,11547	0,02125	-5,43445	0,00000
10	ASSET_MANAGER=UBS	-0,09422	0,01007	-9,35509	0,00000
10	TYPE (SECTION)=OBJ_AND_INV_POL	0,12677	0,00683	18,55173	0,00000
10	GEO_AREA=Middle East and North Africa (MENA)	0,11901	0,07170	1,65972	0,09704

11	ASSET_MANAGER=NORDEA	0,20130	0,00800	25,15786	0,00000
11	ASSET_MANAGER=SWEDBANK	0,10856	0,02737	3,96563	0,00007
11	TYPE (SECTION)=INFO	0,10871	0,00576	18,86875	0,00000
12	ASSET_MANAGER=BNP PARIBAS	-0,10088	0,01556	-6,48319	0,00000
12	ASSET_MANAGER=EURIZON	-0,06360	0,02911	-2,18459	0,02897
12	ASSET_MANAGER=FIDELITY	-0,18910	0,01364	-13,86539	0,00000
12	ASSET_MANAGER=JP MORGAN	-0,18900	0,00902	-20,94911	0,00000
12	ASSET_MANAGER=SCHRODERS	-0,14722	0,02521	-5,84094	0,00000
12	ASSET_MANAGER=SWEDBANK	-0,21647	0,02791	-7,75550	0,00000
12	ASSET_MANAGER=UBS	-0,04950	0,01377	-3,59408	0,00033
12	TYPE (SECTION)=RISK_AND_REW	0,58449	0,00862	67,81530	0,00000
13	ARTICLE= ART. 8	0,02233	0,00370	6,03420	0,00000
13	ARTICLE= ART. 9	0,02291	0,00499	4,59158	0,00000
13	ASSET_MANAGER=BLACKROCK	-0,05118	0,00463	-11,04275	0,00000
13	ASSET_MANAGER=BNP PARIBAS	-0,05349	0,00737	-7,25952	0,00000
13	ASSET_MANAGER=EURIZON	-0,04969	0,01234	-4,02592	0,00006
13	ASSET_MANAGER=FIDELITY	-0,05681	0,00636	-8,93775	0,00000
13	ASSET_MANAGER=JP MORGAN	-0,05984	0,00465	-12,86404	0,00000
13	ASSET_MANAGER=NORDEA	-0,05681	0,00550	-10,32598	0,00000
13	ASSET_MANAGER=SCHRODERS	-0,05871	0,01127	-5,21151	0,00000
13	ASSET_MANAGER=SWEDBANK	-0,06781	0,01265	-5,35855	0,00000
13	ASSET_MANAGER=UBS	-0,04736	0,00605	-7,83216	0,00000
13	TYPE (SECTION)=PAST_PERF	0,06004	0,00460	13,06343	0,00000
14	(Intercept)	0,71622	0,08417	8,50876	0,00000
14	ASSET_MANAGER=BNP PARIBAS	-0,06101	0,01711	-3,56593	0,00037
14	ASSET_MANAGER=JP MORGAN	-0,18888	0,00892	-21,17106	0,00000
14	ASSET_MANAGER=NORDEA	-0,11928	0,01278	-9,33199	0,00000
14	ASSET_MANAGER=SCHRODERS	-0,04857	0,02783	-1,74524	0,08101
14	ASSET_MANAGER=SWEDBANK	-0,22312	0,02807	-7,94840	0,00000
14	TYPE (SECTION)=INFO	-0,65385	0,00936	-69,83332	0,00000

14	TYPE (SECTION)=OBJ_AND_INV_POL	-0,65125	0,00934	-69,71461	0,00000
14	TYPE (SECTION)=PAST_PERF	-0,65116	0,00916	-71,05092	0,00000
14	TYPE (SECTION)=RISK_AND_REW	-0,65125	0,00941	-69,23759	0,00000
15	ASSET_MANAGER=FIDELITY	0,20007	0,00974	20,54365	0,00000
15	ASSET_MANAGER=SCHRODERS	0,10019	0,02483	4,03585	0,00006
15	TYPE (SECTION)=RISK_AND_REW	0,09854	0,00578	17,04038	0,00000
16	ASSET_MANAGER=EURIZON	0,06370	0,01621	3,93089	0,00009
16	ASSET_MANAGER=SCHRODERS	0,19200	0,02492	7,70544	0,00000
16	ASSET_MANAGER=SWEDBANK	0,03227	0,01567	2,05862	0,03959
16	TYPE (SECTION)=OBJ_AND_INV_POL	0,01568	0,00270	5,80550	0,00000
17	ARTICLE= ART. 8	-0,01475	0,00535	-2,75741	0,00585
17	ASSET_MANAGER=EURIZON	0,11215	0,02406	4,66218	0,00000
17	ASSET_MANAGER=FIDELITY	0,19177	0,01051	18,24757	0,00000
17	ASSET_MANAGER=UBS	0,11502	0,01094	10,51039	0,00000
17	TYPE (SECTION)=PAST_PERF	0,13074	0,00628	20,82085	0,00000
17	GEO_AREA=Canada	0,16629	0,09560	1,73947	0,08202
18	ARTICLE= ART. 9	0,02078	0,00685	3,03176	0,00244
18	ASSET_MANAGER=BLACKROCK	0,09738	0,00670	14,53389	0,00000
18	ASSET_MANAGER=EURIZON	0,04409	0,02031	2,17072	0,03000
18	TYPE (SECTION)=OBJ_AND_INV_POL	0,10120	0,00661	15,32053	0,00000
19	ASSET_MANAGER=JP MORGAN	0,19948	0,00801	24,90726	0,00000
19	ASSET_MANAGER=SWEDBANK	0,06987	0,03275	2,13370	0,03292
19	ASSET_MANAGER=UBS	0,02295	0,01203	1,90804	0,05645
19	TYPE (SECTION)=RISK_AND_REW	0,21680	0,00752	28,81608	0,00000
20	ARTICLE= ART. 8	-0,00686	0,00266	-2,58247	0,00984
20	ARTICLE= ART. 9	-0,01409	0,00344	-4,08910	0,00004
20	ASSET_MANAGER=BLACKROCK	-0,01281	0,00314	-4,07880	0,00005
20	ASSET_MANAGER=FIDELITY	-0,00849	0,00369	-2,29966	0,02151
20	ASSET_MANAGER=JP MORGAN	-0,00690	0,00259	-2,65894	0,00787
20	ASSET_MANAGER=UBS	-0,00833	0,00353	-2,36318	0,01816

20	TYPE (SECTION)=OBJ_AND_INV_POL	0,00634	0,00229	2,77119	0,00561
20	TYPE (SECTION)=PAST_PERF	0,00968	0,00252	3,83639	0,00013
21	ASSET_MANAGER=NORDEA	0,16759	0,00886	18,92026	0,00000
21	ASSET_MANAGER=UBS	0,01671	0,00966	1,73025	0,08365
21	TYPE (SECTION)=PAST_PERF	0,09104	0,00571	15,94752	0,00000
22	ASSET_MANAGER=BNP PARIBAS	0,03872	0,01360	2,84744	0,00443
22	ASSET_MANAGER=JP MORGAN	0,19921	0,00804	24,78338	0,00000
22	ASSET_MANAGER=SWEDBANK	0,04228	0,02548	1,65912	0,09716
22	TYPE (SECTION)=INFO	0,21451	0,00747	28,71745	0,00000
23	ASSET_MANAGER=NORDEA	0,18612	0,00811	22,95197	0,00000
23	TYPE (SECTION)=OBJ_AND_INV_POL	0,09720	0,00557	17,43456	0,00000
24	ASSET_MANAGER=UBS	0,19885	0,00907	21,93470	0,00000
24	TYPE (SECTION)=OBJ_AND_INV_POL	0,04938	0,00416	11,86258	0,00000
24	TYPE (SECTION)=PAST_PERF	0,00725	0,00424	1,70911	0,08750
24	GEO_AREA=Developed and Emerging Markets	-0,20824	0,09600	-2,16925	0,03011
25	ARTICLE= ART. 8	-0,01763	0,00517	-3,40711	0,00066
25	ARTICLE= ART. 9	-0,02243	0,00747	-3,00472	0,00267
25	ASSET_MANAGER=BLACKROCK	0,03666	0,00727	5,03940	0,00000
25	ASSET_MANAGER=BNP PARIBAS	0,15025	0,01261	11,91188	0,00000
25	ASSET_MANAGER=SWEDBANK	0,13555	0,03305	4,10161	0,00004
25	TYPE (SECTION)=OBJ_AND_INV_POL	0,03262	0,00630	5,17918	0,00000
25	TYPE (SECTION)=PAST_PERF	0,04603	0,00524	8,78191	0,00000
25	RISK (CATEGORY)=5	0,02686	0,01567	1,71391	0,08661
26	ARTICLE= ART. 8	-0,01528	0,00422	-3,62510	0,00029
26	ASSET_MANAGER=BLACKROCK	-0,01047	0,00498	-2,10390	0,03544
26	ASSET_MANAGER=BNP PARIBAS	0,11425	0,01114	10,25220	0,00000
26	ASSET_MANAGER=JP MORGAN	0,04218	0,00526	8,02162	0,00000
26	TYPE (SECTION)=OBJ_AND_INV_POL	0,06757	0,00541	12,50070	0,00000
27	(Intercept)	0,09438	0,05608	1,68278	0,09249
27	ARTICLE= ART. 8	-0,00935	0,00550	-1,70146	0,08892

27	ARTICLE= ART. 9	-0,01249	0,00753	-1,65863	0,09726
27	ASSET_MANAGER=BLACKROCK	-0,11433	0,00667	-17,13127	0,00000
27	ASSET_MANAGER=BNP PARIBAS	-0,10946	0,01041	-10,51661	0,00000
27	ASSET_MANAGER=EURIZON	-0,11324	0,01872	-6,05054	0,00000
27	ASSET_MANAGER=FIDELITY	-0,11631	0,00939	-12,38731	0,00000
27	ASSET_MANAGER=JP MORGAN	-0,11023	0,00673	-16,38694	0,00000
27	ASSET_MANAGER=NORDEA	-0,11174	0,00812	-13,75616	0,00000
27	ASSET_MANAGER=SCHRODERS	-0,04776	0,02007	-2,37989	0,01736
27	ASSET_MANAGER=SWEDBANK	-0,11743	0,01934	-6,07202	0,00000
27	ASSET_MANAGER=UBS	-0,11572	0,00925	-12,50454	0,00000
27	TYPE (SECTION)=PAST_PERF	0,14246	0,00639	22,29667	0,00000
28	(Intercept)	0,14897	0,08242	1,80756	0,07074
28	ASSET_MANAGER=BLACKROCK	-0,19942	0,00889	-22,43109	0,00000
28	ASSET_MANAGER=BNP PARIBAS	-0,19794	0,01444	-13,71123	0,00000
28	ASSET_MANAGER=EURIZON	-0,20021	0,02676	-7,48070	0,00000
28	ASSET_MANAGER=FIDELITY	-0,19971	0,01327	-15,05152	0,00000
28	ASSET_MANAGER=JP MORGAN	-0,19960	0,00869	-22,96414	0,00000
28	ASSET_MANAGER=NORDEA	-0,19942	0,01104	-18,05584	0,00000
28	ASSET_MANAGER=SCHRODERS	-0,19914	0,02387	-8,34343	0,00000
28	ASSET_MANAGER=SWEDBANK	-0,14794	0,02752	-5,37523	0,00000
28	ASSET_MANAGER=UBS	-0,19915	0,01278	-15,57689	0,00000
28	TYPE (SECTION)=INFO	0,25779	0,00807	31,94299	0,00000
29	ARTICLE= ART. 8	0,01580	0,00430	3,67582	0,00024
29	ARTICLE= ART. 9	0,01115	0,00560	1,99132	0,04650
29	ASSET_MANAGER=BLACKROCK	0,05056	0,00592	8,54161	0,00000
29	ASSET_MANAGER=JP MORGAN	-0,00828	0,00475	-1,74469	0,08111
29	TYPE (SECTION)=OBJ_AND_INV_POL	0,05638	0,00524	10,75332	0,00000
29	RISK (CATEGORY)=7	-0,05762	0,02286	-2,52079	0,01174
29	GEO_AREA=Asia Pacific region, including Australia an Japan	0,16134	0,05810	2,77721	0,00551

30	ASSET_MANAGER=BNP PARIBAS	0,15659	0,01376	11,37759	0,00000
30	ASSET_MANAGER=EURIZON	0,20448	0,02452	8,33908	0,00000
30	ASSET_MANAGER=FIDELITY	0,20042	0,01220	16,42457	0,00000
30	ASSET_MANAGER=SCHRODERS	0,10609	0,02710	3,91548	0,00009
30	ASSET_MANAGER=SWEDBANK	0,04480	0,02510	1,78432	0,07444
30	ASSET_MANAGER=UBS	0,19658	0,01166	16,85329	0,00000
30	TYPE (SECTION)=INFO	0,20972	0,00749	27,99634	0,00000
31	ASSET_MANAGER=BLACKROCK	0,12338	0,00778	15,86841	0,00000
31	ASSET_MANAGER=EURIZON	0,08419	0,02586	3,25580	0,00114
31	ASSET_MANAGER=SCHRODERS	0,04200	0,02437	1,72348	0,08487
31	ASSET_MANAGER=SWEDBANK	0,07365	0,03270	2,25223	0,02435
31	ASSET_MANAGER=UBS	0,03145	0,01103	2,85182	0,00437
31	TYPE (SECTION)=PAST_PERF	0,14889	0,00721	20,65264	0,00000
31	RISK (CATEGORY)=5	0,03781	0,02277	1,66054	0,09687
32	ASSET_MANAGER=BLACKROCK	0,00965	0,00318	3,03873	0,00239
32	ASSET_MANAGER=BNP PARIBAS	0,07260	0,01030	7,04920	0,00000
32	TYPE (SECTION)=INFO	-0,01563	0,00296	-5,27681	0,00000
32	TYPE (SECTION)=PAST_PERF	-0,01485	0,00309	-4,80887	0,00000
32	TYPE (SECTION)=RISK_AND_REW	-0,01486	0,00291	-5,10463	0,00000
33	ARTICLE= ART. 8	0,01838	0,00605	3,03818	0,00239
33	ASSET_MANAGER=JP MORGAN	0,14577	0,00755	19,30147	0,00000
33	TYPE (SECTION)=OBJ_AND_INV_POL	0,15657	0,00704	22,24744	0,00000
34	ARTICLE= ART. 8	-0,02815	0,00493	-5,70628	0,00000
34	ARTICLE= ART. 9	-0,01658	0,00673	-2,46182	0,01386
34	ASSET_MANAGER=BLACKROCK	-0,03777	0,00605	-6,24253	0,00000
34	ASSET_MANAGER=BNP PARIBAS	0,01864	0,01059	1,76048	0,07839
34	ASSET_MANAGER=EURIZON	-0,03673	0,01612	-2,27871	0,02273
34	ASSET_MANAGER=FIDELITY	-0,02770	0,00838	-3,30796	0,00095
34	ASSET_MANAGER=JP MORGAN	-0,03112	0,00591	-5,26719	0,00000
34	ASSET_MANAGER=NORDEA	-0,03480	0,00717	-4,85291	0,00000

34	ASSET_MANAGER=SCHRODERS	-0,03564	0,01455	-2,44885	0,01437
34	ASSET_MANAGER=SWEDBANK	0,10275	0,03155	3,25669	0,00114
34	ASSET_MANAGER=UBS	-0,02144	0,01027	-2,08772	0,03688
34	TYPE (SECTION)=OBJ_AND_INV_POL	0,07867	0,00605	13,00656	0,00000
35	ASSET_MANAGER=BNP PARIBAS	0,08424	0,00754	11,16875	0,00000
35	TYPE (SECTION)=RISK_AND_REW	0,01896	0,00228	8,32421	0,00000

Figure S1: Topic networks for four sustainability-related topics (1st row: Topic 4; 2nd row: Topic 5; 3rd row: Topic 10; 4th row: Topic 26) for Article-8 Vs Article 9.

Bigram network for Topic: 5
Nodes & edges (N&E) percentile thresholds: 0.95
Reference faction (N&E width scale): ART.8 (0.5) & ART.9 (0.5)
Compared factions (N&E color scale): ART.8 Vs ART.9

Bigram network for Topic: 10
Nodes & edges (N&E) percentile thresholds: 0.95
Reference faction (N&E width scale): ART.8 (0.5) & ART.9 (0.5)
Compared factions (N&E color scale): ART.8 Vs ART.9
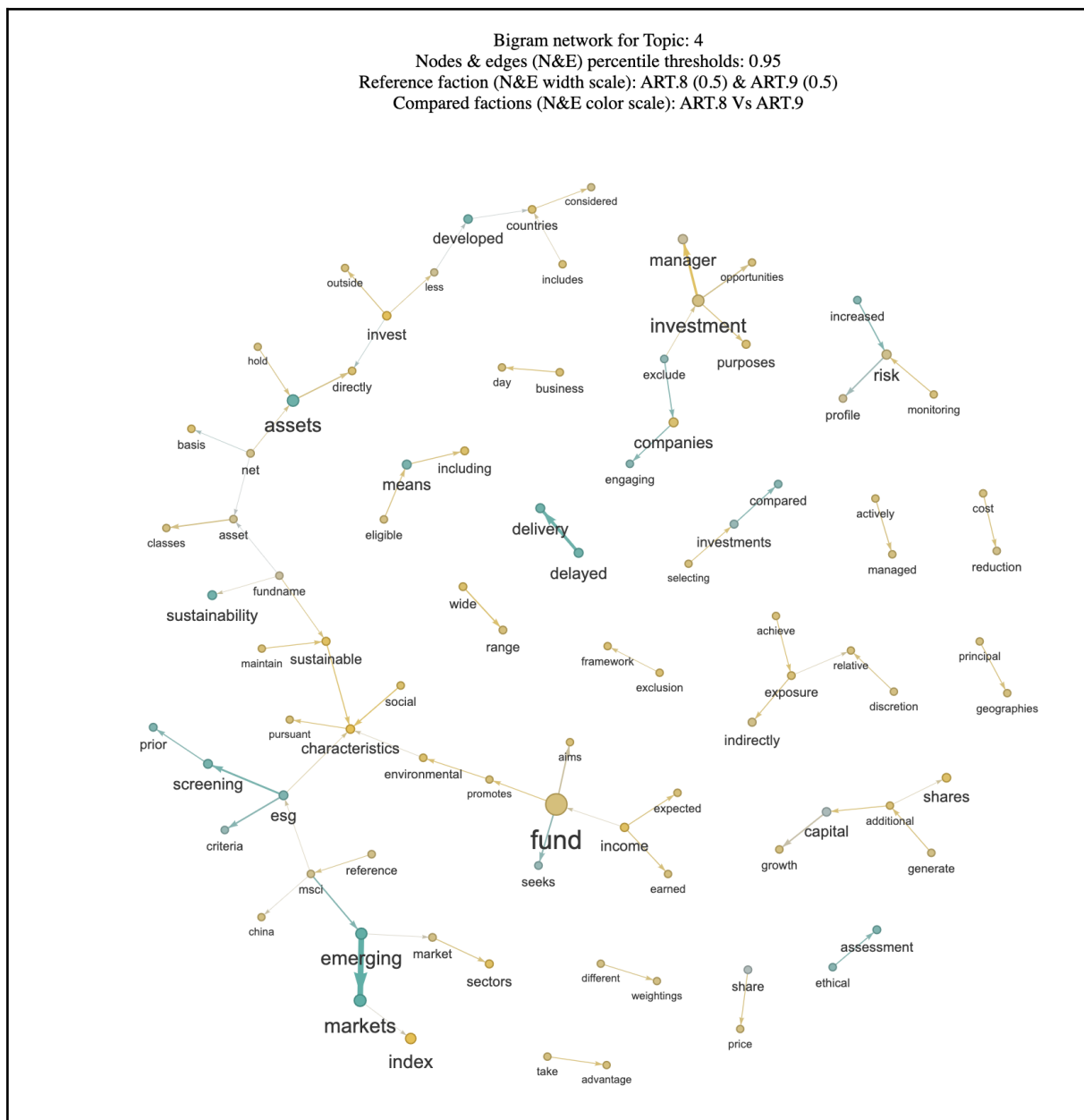
Bigram network for Topic: 26
Nodes & edges (N&E) percentile thresholds: 0.95
Reference faction (N&E width scale): ART.8 (0.5) & ART.9 (0.5)
Compared factions (N&E color scale): ART.8 Vs ART.9

S2 Robustness Checks

Table S5: Parameters used in experiments with different values for K from 5 to 100.

parameter	value(s)
K	5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100
Number of experiment replications for each K	10
Maximum Expectation Maximization Iterations	100
Expectations Maximization Tolerance (em.tol)	1e-05
init.type	"Spectral"
Proportion of docs to be held out	0.25 (25%)
M	10

Figure S2: Mean values of Lower bound and Heldout likelihood (based on 10 replications) of estimated STM for different values of K. Labels represent the number of topics K.
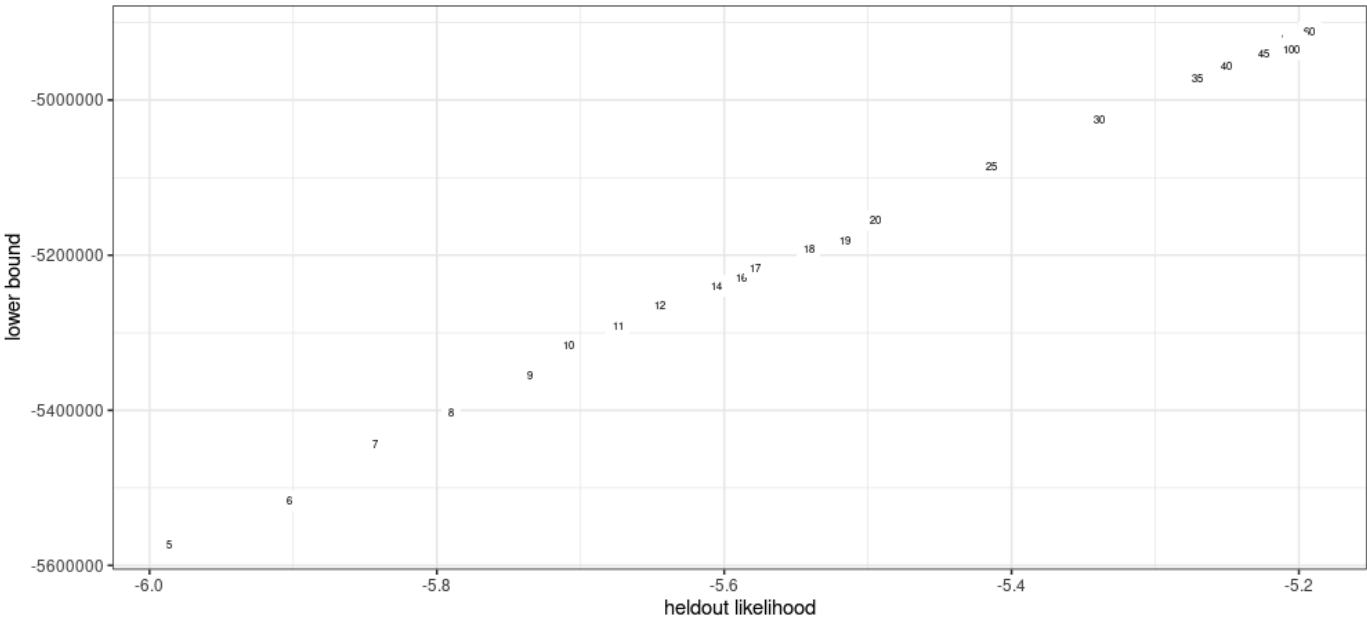
Labels represent the number of topics.
To make them comparable lower bounds values (y axis) have been corrected following the method proposed by Bishop (2006).
For each topic number K in the interval [5, 100] the STM model has been estimated 10 times.
At each estimation a different random sample containing 25% of the observations has been heldout.
The heldout sample has been used to compute the heldout likelihood.

Figure S3: Semantic Coherence and Exclusivity (based on 10 replications) of estimated STM for different values of K. Labels represent the number of topics K.
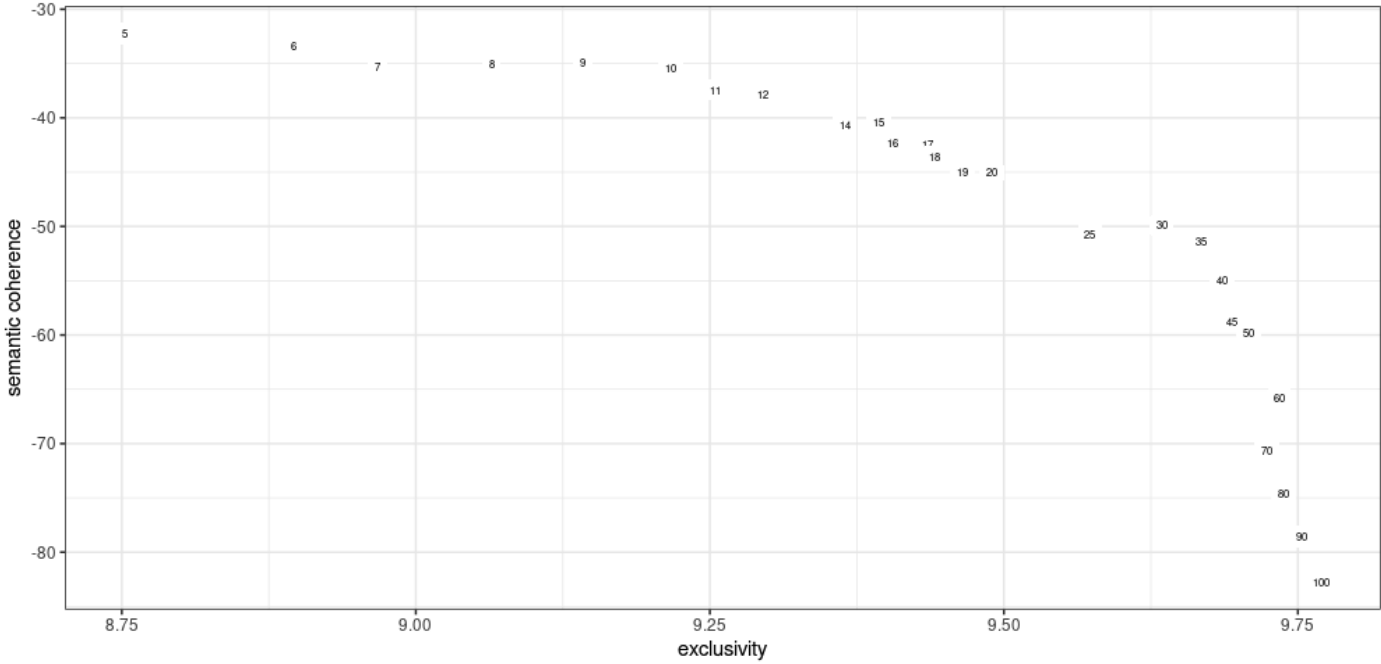
Labels represent the number of topics.
For each topic number K in the interval [5, 100] the STM model has been estimated 10 times.
At each estimation a different random sample containing 25% of the observations has been heldout.

S3 Data and Methods

Regarding the sample selection, to collect our dataset, which is available at the following URL: https://github.com/carlosantagiustina/mapping_ESG_discourse_through_STM, we relied on a report (URL: http://global.morningstar.com/equitydisclosures) by Morningstar published in 2022, mapping several important dimensions of sustainable investment funds (including flows, players and geography).

In Morningstar (2022), top 20 asset managers are ranked in exhibit 19 by article 8 and article 9 funds assets combined, based on SFDR data collected from prospectuses on 91% of funds available for sale in the EU, as of 31 Dec 2021. We select KIIDs from the first ten Asset managers in this list (i.e., JPMorgan, UBS, Amundi, BNP Paribas, Blackrock, Fidelity, Schroders, Swedbank, Nordea, Eurizon) to build our sample. The ten selected asset managers represented together about 28% of the market share for article 8 and 9 funds, as of 31-12-2021.

The reason for choosing funds from top 10 asset managers identified in Morningstar (2022) as the targeted population is twofold:

- From a conceptual point of view, we need to focus on asset managers that are concerned with sustainable issues because we are interested in exploring the relation between sustainability narratives in the funds' KIIDs and their SFDR classification, to see how (and in which terms) different SFDR classes frame their sustainability narratives, in a way that is coherent (or not) with their classification. Therefore, we need asset managers that manage funds from different SFDR classes, that is, we are seeking within group variation. The 10 selected asset managers meet this criterion.

- From a practical/methodological point of view, we select Top 10 asset managers in Morningstar (2022), by Art. 8 and Art. 9 class, because topic modeling methods, like STM, perform better if the corpus of documents (KIIDs) is large. Furthermore, to obtain an human-interpretable model and limit the parameter size of the STM (i.e., each asset manager requires an additional level in the ASSET_MANAGER categorical covariate), we are ideally seeking to have a large sample of Art. 8 and Art. 9 funds with a limited number of asset managers. Also for this reason we select only 10 Asset managers among the Top 20 list in Morningstar (2022).

For the 10 chosen asset managers, we identified all Art.8 and Art.9 funds from www.morningstar.lu and www.morningstar.ie. We then proceeded to each Asset manager's website and downloaded the KIIDs available/downloadable (in a standardized PDF format) through the websites of the asset managers at the time of dataset creation (June-July 2022 for Articles 8 and 9). For Article 6 funds, for each of the 10 selected asset managers, we randomly selected a sample of KIIDs in Nov.-Dec. 2022. Since not all asset managers' websites are structured the same way, and asset managers are diverse in terms of number and types of funds, the KIIDs sample selection approach was customized to obtain a sample that is, as much as possible, representative of the targeted population of asset managers also for non-sustainable funds. That is, Art. 6 equity funds in EUR from the 10 selected asset managers.

Figure S4: Frequency of geographical funds composition, by Article class and Asset manager.
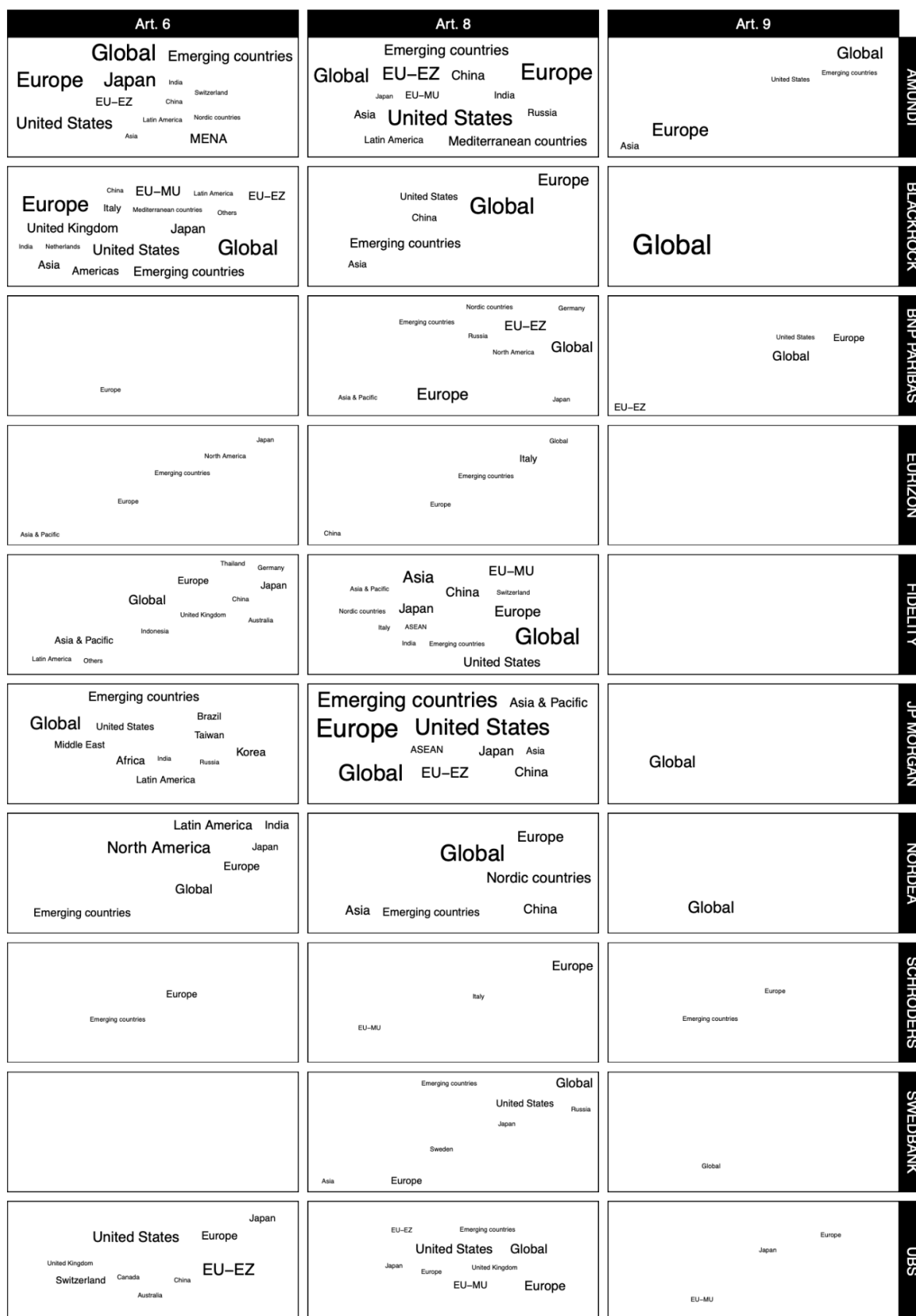
Font size is proportional to the frequency of a specific geographic-area for funds by Article class and Asset manager

Table S6: List of stopwords (based on Quanteda stopword list) used to prune the token-by-document matrix before the STM estimation step.*

"i"	"me"	"my"	"myself"	"we"	"our"	"ours"	"ourselves"	"you"	"your"	"yours"	"yourself"
"yourselves"	"he"	"him"	"his"	"himself"	"she"	"her"	"hers"	"herself"	"it"	"its"	"itself"
"they"	"them"	"their"	"theirs"	"themselves"	"what"	"which"	"who"	"whom"	"this"	"that"	
"these"	"those"	"am"	"is"	"are"	"was"	"were"	"be"	"been"	"being"	"have"	"has"
"had"	"having"	"do"	"does"	"did"	"doing"	"would"	"should"	"could"	"ought"	"i'm"	"you're"
"he's"	"she's"	"it's"	"we're"	"they're"	"i've"	"you've"	"we've"	"they've"	"i'd"	"you'd"	"he'd"
"she'd"	"we'd"	"they'd"	"i'll"	"you'll"	"he'll"	"she'll"	"we'll"	"they'll"	"isn't"	"aren't"	"wasn't"
"weren't"	"hasn't"	"haven't"	"hadn't"	"doesn't"	"don't"	"didn't"	"won't"	"wouldn't"	"shan't"	"shouldn't"	
"can't"	"cannot"	"couldn't"	"mustn't"	"let's"	"that's"	"who's"	"what's"	"here's"	"there's"	"when's"	
"where's"	"why's"	"how's"	"a"	"an"	"the"	"and"	"but"	"if"	"or"	"because"	"as"
"until"	"while"	"of"	"at"	"by"	"for"	"with"	"about"	"against"	"between"	"into"	"through"
"during"	"before"	"after"	"above"	"below"	"to"	"from"	"up"	"down"	"in"	"out"	"on"
"off"	"over"	"under"	"again"	"further"	"then"	"once"	"here"	"there"	"when"	"where"	"why"
"how"	"all"	"any"	"both"	"each"	"few"	"more"	"most"	"other"	"some"	"such"	"no"
"nor"	"not"	"only"	"own"	"same"	"so"	"than"	"too"	"very"	"will"	"sub"	"can"
"also"	"may"	"could"	"therefore"	"hence"	"www"	"will"	"r.l"	"s.à"	"i.e"	"s.à.r.l"	

*In addition to stopwords, by using the following RegEx function (gsub("AMUNDI|BLACK[-]?ROCK|B[.]?N[.]?P[.]?[-]?PARIBAS|EURIZON|FIDELITY|J[.]?P[.]?[-]?MORGAN|[^A-Z](MORGAN)[^A-Z]|NORDEA|SCHRODERS|SWEDBANK|[^A-Z](UBS)[^A-Z]", x=data$TEXT, replacement = " FUNDNAME ",perl = T,ignore.case = T)), the names of Asset Manager have been substituted with the more general token "FUNDNAME". This allows to mitigate the influence of Asset Manager specific labels in the topics of the STM.

Regarding the estimation of the Structural Topic Model, as summarized in the manuscript, the choice of the number of topics parameter (K=35) was based on the four criteria suggested by Roberts et al. (2019), two of them (see Figure S2) are used to evaluate the in-sample and out-of-sample fitness of the model, the remaining two serve to evaluate the average quality of the topics (see Figure S3). In particular, for each value of K described in Table S5, we estimated 10 times the STM, holding out each time a different random sample of the corpus, containing 25% of the total number of documents, and using the complement of the holdout sample for the STM estimation, containing 75% of the total number of documents. Then, we compute for each K the averages of four information criteria. As Figure S2 suggests, the fitness of the STM is rapidly increasing as a function of K for low values of K, ranging from 5 to 35, and then decelerates and stabilizes for K values larger than 45. On the other hand, Figure S3 shows a clear tradeoff between the average semantic coherence and exclusivity of the topics, with K=35 being on the "efficient frontier" and exhibiting a balanced tradeoff between the two criteria. Based on the aforementioned results, we choose K=35 and estimate the STM with the whole dataset using the parameters specified in Table S1. With respect to the default parameters of the *stm* library, we have lowered the Expectations Maximization

Tolerance to 1e-06, to ensure that the model only interrupts iterations when the difference between one iteration and the next is very small. The model is initialized using a deterministic technique called *Spectral initialization*. The model converges after 71 iterations.

For a detailed discussion of Structural Topic Modelling and the role on topic content and topic propensity covariates, please refer to Roberts et al. (2019) and to other Methods Papers made available by the authors at the following link: https://www.structuraltopicmodel.com/

For a detailed discussion on the bigrams and the topic networks (Figure S1) construction process please refer to the Materials and methods section (subsection: *Constructing faction-specific narrative networks*) in Santagiustina et al. (2022), and to the paper's Supplement available at the following link: https://doi.org/10.1371/journal.pone.0270236.s001

To facilitate the exploration of the topics discussed in the KIIDs two additional tables (Table S2 and Table S3) are provided in this Supplement. These two tables show top 5 tokens by topic and by Article class, ordered by FREX score (in Table S2) or by probability (in Table S3).

References

- Bischof, J., & Airoldi, E. M. (2012). Summarizing topical content with word frequency and exclusivity. *Proceedings of the 29th international conference on machine learning* (icml-12) (pp. 201-208).

- Morningstar Manager Research (February, 2022) SFDR Article 8 and Article 9 Funds: 2021 in Review http://global.morningstar.com/equitydisclosures

- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). STM: An R package for structural topic models. *Journal of Statistical Software*, 91, 1-40

- Santagiustina, C. R. M. A., & Warglien, M. (2022). The architecture of partisan debates: The online controversy on the no-deal Brexit. *PLoS one*, 17(6), e0270236 doi.org/10.1371/journal.pone.0270236