Climate Change News Indices: Are They Reflected in Japanese Stock Prices?

Kazuhiro Hiraki[1] Noriyasu Kaneda[1]
Tomonori Kimata[1] Tomohiro Matsue[2]

[1]BOJ, [2]BOJ (currently Sumitomo Life Insurance)

CEBRA Annual Meeting 2023
July 7, 2023

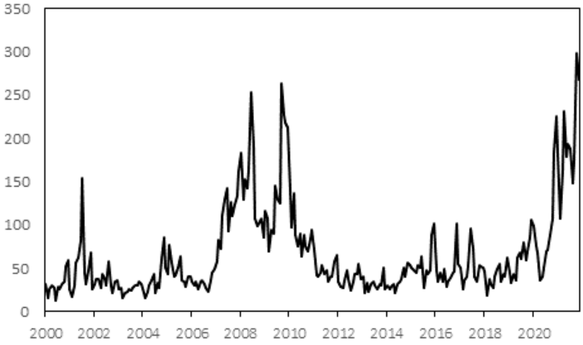


図 1. **Monthly numbers of climate-related articles (Nikkei, morning eds.)**

Elevated attentions in Japanese media (newspapaer)

1. 2008–09: Kyoto protocol, Toyako G7 summit, Copenhagen COP
2. 2020–: Recent movements for zero carbon emission

Research questions

Q: Are attentions/concerns on climate risk reflected in Japanese stock prices?

1. How to quantify and dissect "attentions/concerns on climate risk" ?
   - news articles and natural language processing (NLP) (e.g., Engle et al., 2020)
   - climate change is multi-faceted issue $\Rightarrow$ use topic modeling tech to dissect
     - physical risk (natural disasters)
     - transition risk (env. regulations, energy policy, ESG investment...)
2. How to examine whether climate risk is "reflected in stock prices" ?
   - regard climate indices as (additional) risk factors
   - test whether factor exposure to climate factors earn risk premium (α)

What we do and find

1. Apply NLP tech (LDA) to Nikkei newspaper data and make climate indices

   ⇒ We make 10 climate-related indices, e.g., "env. tech," "green finance," "env. regulation," "natural disasters," "int'l summit"

   - (to the authors knowledge) first study using a Japanese corpus
   - LDA is an unsupervised topic modeling tech that dissect a corpus into prespecified number of topics
   - LDA is very popular in recent economics & finance literature

2. Calculate "climate beta" for each stock and climate index. Then, examine relationship between beta exposure and risk premium

   - results suggest "green finance" and "env. regulation" betas are priced

   ⇒ beta-sorted spread portfolio α's sometimes exceed 0.4% per month!

Selected literature additional literature

- Application of NLP in economics & finance
  - Survey and Seminal papers: Gentzkow et al. (2019, JEL), Loughran and McDonald (2011, JF), Baker et al. (2016, QJE), Engle et al. (2020, RFS)
  - Application of LDA: Bandiera et al. (2017, JPE), Larsen and Thorsrud (2019, JoE), Nimark and Pitschner (2019, JET), Faccini et al. (2021), Dierckx et al. (2021)
  - Textual analysis and climate risk: Engle et al. (2020), Kapfhammer et al. (2020), Faccini et al. (2021), Bessec and Fouquau (2021), Bua et al. (2022), Ardia et al. (2022), Sautner et al. (2021)
- Climate change and equity risk premium
  - Theory: Pedersen et al. (2021), Pastor et al. (2021), Avramov et al. (2021)...
  - Hedging: Krueger et al. (2020), Giglio, Kelly, Stroebel (2021), Stroebel and Wurgler (2021), Engle et al. (2020), Pastor et al. (2021), Alekseev et al. (2022)...

1. Introduction
   - Research questions
   - Literature
2. Making climate news indices
   - Pre-processing textual data
   - Topic modeling (LDA)
   - Climate-related articles in Nikkei newspaper
3. Climate risk and Japanese equity risk premium
   - Climate beta and portfolio analysis
   - Results of asset pricing tests
4. Conclusions

Making climate news incides: Procedures

1. Pre-processing textual data detail
   - cleaning documents (e.g., dropping unnec. symbols)
   - tokenization (Japanese doesn't use space bw. words)
   - remove unnec. words (grammatical words, stop words...)
   - lemmatization (basic form of verbs and adjectives)
   - transform into Bag-of-words (BoW) expression
2. Estimation of topic model (LDA)
   - determine number of topics
   - interpret estimated topics
3. Time-aggregate LDA topic shares to make indices
   - decomposition of daily total number of articles into each LDA topic

LDA (Latent Dirichlet Allocation)

- LDA is a relatively mature unsupervised NLP tech that estimates topics from a corpus (Blei et al., 2003)
  - Number of topics is set manually beforehand
- Expresses the structure of corpus by **topics** and **topic distributions** detail
  - A **topic** is a prob. distr. over list of all unique words appearing in the corpus (frequently appearing words characterize topics)
  - **Topic distr.** differ across articles, showing how each article loads on topics

◎ No training data (i.e., topic-tagged data) required

$\Rightarrow$ Easy to use $\Rightarrow$ main reason for widespread use in economics & finance

× Estimation of latent topics $\Rightarrow$ no a priori economic meaning

- Need human judgment on number of LDA topics & interpretation of topics

Data

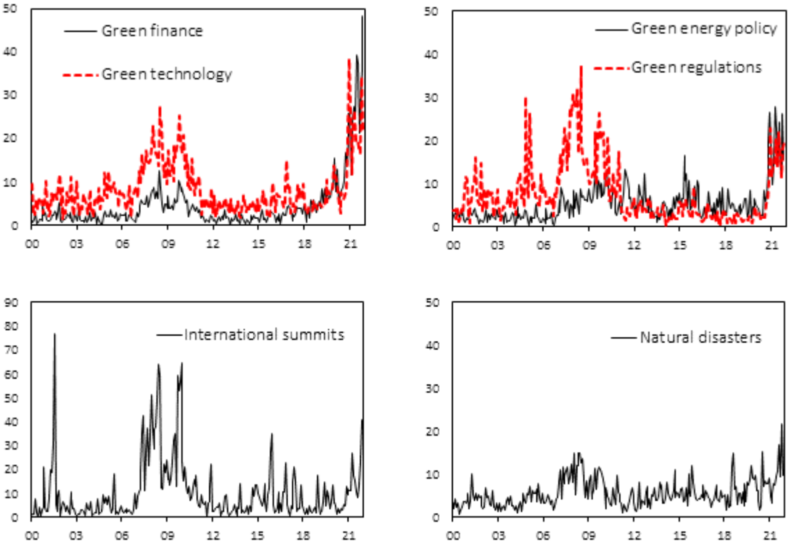
- Nikkei newspaper (morning eds.)
  - Nikkei is the like of WSJ and FT, specializes in news on economy, financial markets and business activities
- Period: January 2000 – November 2021
  - 17,275 articles containing at least one of...
    - ★ Climate change (気候変動)
    - ★ Global warming (温暖化)
    - ★ Greenhouse effect (温室効果)
- 10 LDA topics to be optimal based on the coherence measure and manual inspection
- Stock returns and market cap data from Bloomberg
- Fama-French factors from Ken French's website (converted to yen-denominated returns)

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
	Int'l summit	**security**	**env. regulation**	**env. technology**	**life style**
1	EU	US	emission	car	world
2	conference	China	volume	technology	economy
3	climate	US	reduce	hydorogen	coronavirus
4	target	president	target	use	problem
5	change	administration	carbon	development	society
6	framework	Trump	CO	CO	research
7	Europe	Biden	trading	EV	come
8	prime minister	policy	firm	automobile	climate
9	leaders	economy	achieve	environment	needs
10	international	problem	introduce	battery	change
11	measures	support	fiscal year	fuel	technology
12	agreement	election	measures	volume	environment
13	UN	ideology	industry	sales	international
14	COP	diplomacy	ratio	production	important
15	China	politics	environment	electoricity	think

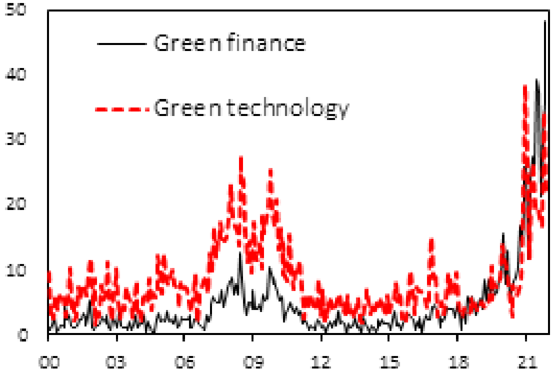
	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
	market	**energy policy**	**tax & soc. sec.**	**nat. disasters**	**green finance**
1	price	power generation	policy	temparature	firm
2	oil	energy	economy	influence	investment
3	production	electoric power	tax	disaster	environment
4	world	renewable	public finance	rise	finance
5	demand	nuclear pwr plnt	measures	meteorology	ESG
6	market	thermal pwr plnt	prime minister	Tokyo	climate
7	dollar	coal	reform	damage	disclosure
8	China	energy	budget	earth	change
9	investment	nuclear power	government	emerge	enterprise
10	energy	solar	system	sea	institution
11	volume	plan	fiscal year	cliamte	information
12	proportion	possible	amendment	observation	management
13	fuel	wind power	administration	projection	risk
14	resource	volume	income	research	fund
15	supply	rollout	support	typhoon	asset mngment

Word clouds in Japanese

Climate news indices: selected 6 topics

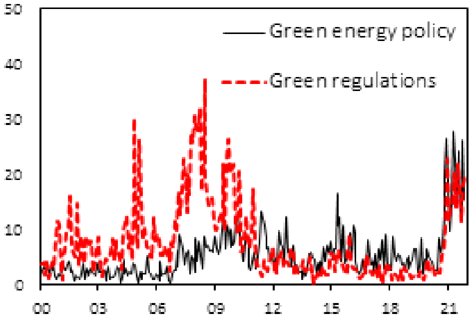


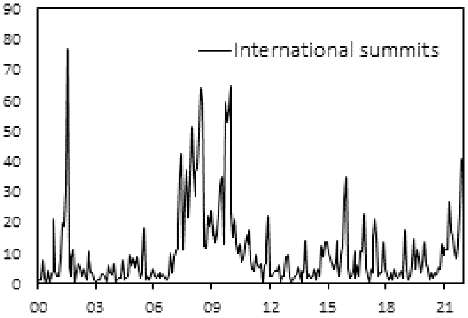
Climate news indices: topics behind 2020– increase



- "Env. tech." ⇒ green business strategies & related capital expenditure
- "green finance" ⇒ ESG investment, disclosure, central banks' policies

Climate news indices: topics behind 2008–09 increase





- “Env. regulation” $\Rightarrow$ emission regulations, emission trading
- “Int’l summit” $\Rightarrow$ COP and G8/G20 summits

Climate news indices: Natural disaster indices

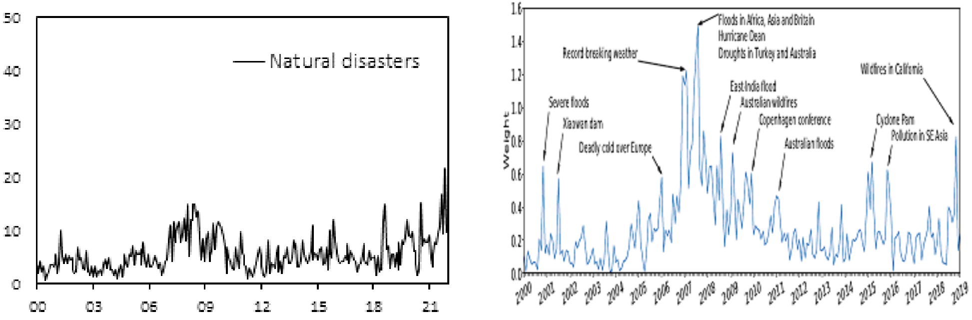


図 2. **left: our JP-based index, right: Faccini et al.'s (2021) EN-based index**

- Natural disaster news (and domestic political news) are local by nature
- Different developments in different countries...

Climate news indices: takeaways

- Huge differences bw. different topics
  - Corroborates multi-faceted nature of climate change
  - Naturally, effects on asset prices will be different by topics
- Different topics drive increases of attention
  - 2008-09: public sector-related (int'l summit, env. regulation)
  - 2020–: private sector-related (env. tech., green finance)
- Different developments compared with English-based indices
  - Local nature of natural disasters (Alekseev et al., 2022 etc.)
  - Curation by media reflecting country-wise political, economic, and cultural background

Climate beta

Climate beta

$$\text{Excess return}_t^i = \alpha_t^i + \beta_{j,t}^{CC,i} CI_{j,t} + \beta^i f_t + \varepsilon_t^i$$

$CI_{j,t}$: j-th climate index, f_t: other risk factors (e.g., FF3)

- Calculate climate beta for each stock-month
  - Sensitivity of stock returns to changes in attention for climate change
  - analogous to "market beta," but to the exposure to climate change risk
  - Beta estimation is based on rolling window estimation
- Is climate beta relevant to stock returns?
- Advantage of analyzing climate beta
  - can be calculated for any stocks (with sufficient time-series data)
  - on the other hand, very narrow coverage of ESG scores etc.

Portfolio analysis

- Portfolio analysis: a standard method to examine relationship bw. individual stock characteristics and cross-section of returns
- Sorting indiv. stocks by char. variable to construct portfolios, and analyze return of them
  1. At the end of each month, sort stocks by climate beta value
  2. Construct N sorted-portfolios (typically 5 or 10)
  3. Construct long-short spread portfolio that longs highest-beta port and shorts lowest-beta port (weighting: value-weight vs equal-weight)
  4. test whether α of the spread port is significant (which factor model?)
  5. If significant, abnormal return difference is driven by beta difference
  6. climate beta exposure is priced in stock returns

Asset pricing test: full sample

表 1. **Alphas of β^{CC}-sorted long-short portfolios**

	VW, TSE1				EW, ALL			
	10		5		10		5	
	FF3	FF4	FF3	FF4	FF3	FF4	FF3	FF4
Panel A: monthly freq. climate beta (60 months)								
Environ. regulation	0.05	-0.05	-0.12	-0.14	0.32***	0.29***	0.22**	0.22***
	(0.23)	(-0.21)	(-0.67)	(-0.77)	(2.91)	(2.78)	(2.47)	(2.63)
Green finance	-0.43**	-0.4**	-0.46***	-0.38**	-0.06	-0.07	-0.05	-0.04
	(-2.22)	(-2.10)	(-2.90)	(-2.43)	(-0.33)	(-0.42)	(-0.36)	(-0.29)
Energy policy	0.05	-0.07	0.13	0.1	0.29	0.22	0.16	0.13
	(0.17)	(-0.20)	(0.59)	(0.43)	(1.51)	(1.31)	(1.21)	(1.08)
Panel B: daily freq. climate beta (12 months)								
Natural disaster	-0.32	-0.25	-0.12	-0.13	-0.16	-0.17	-0.15*	-0.14*
	(-1.47)	(-1.31)	(-0.84)	(-1.01)	(-1.52)	(-1.62)	(-1.83)	(-1.68)

upper: alpha in bps. lower: *t*-stat. sample: Jan. 2000–

Asset pricing test: latter sample (2011–)

表 2. **Alphas of β^{CC}-sorted long-short portfolios**

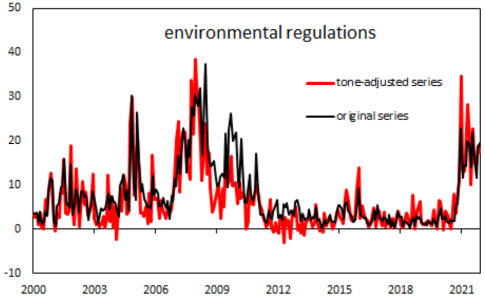
	VW, TSE1				EW, ALL			
	10		5		10		5	
	FF3	FF4	FF3	FF4	FF3	FF4	FF3	FF4
Panel A: monthly freq. climate beta (60 months)								
Environ. regulation	0.11	0.01	-0.1	-0.15	0.43***	0.39***	0.29**	0.3**
	(0.30)	(0.03)	(-0.40)	(-0.60)	(2.75)	(2.64)	(2.32)	(2.53)
Green finance	-0.42	-0.44	-0.46*	-0.35	-0.08	-0.1	-0.02	-0.02
	(-1.26)	(-1.53)	(-1.72)	(-1.43)	(-0.31)	(-0.40)	(-0.10)	(-0.10)
Energy policy	0.01	-0.06	0.25*	0.26**	0.39**	0.35**	0.26**	0.27***
	(0.05)	(-0.19)	(1.82)	(2.19)	(2.19)	(2.06)	(2.34)	(2.66)
Panel B: daily freq. climate beta (12 months)								
Natural disaster	-0.7**	-0.57**	-0.29	-0.3	-0.25**	-0.25**	-0.23**	-0.2**
	(-2.06)	(-2.10)	(-1.44)	(-1.62)	(-2.08)	(-2.17)	(-2.14)	(-2.04)

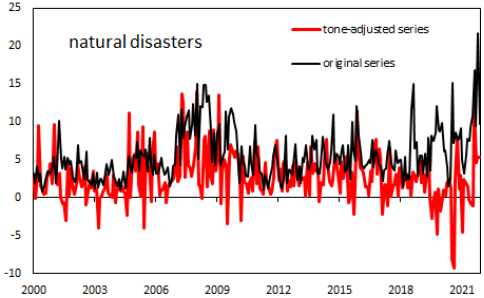
upper: alpha in bps. lower: *t*-stat. sample: Jan. 2011–

Signs of alphas: Interpretation

- How to interpret positive and negative alpha signs?
  - positive α for "env. regulation" and "energy policy" factors
  - negative α for "green finance" and "natural disasters" factors
- Hedge portfolios are expected to earn lower returns (i.e., $\alpha < 0$)
  - risk-averse investors are willing to pay premium for hedge
  - climate concerns $\uparrow \Leftrightarrow$ deterioration of investment env. (Merton's ICAPM)
- By construction, climate index $\uparrow \Rightarrow$ beta-sorted long-short port return $\uparrow$
  - For bad news index (i.e., index $\uparrow$ means bad news) $\Rightarrow \alpha < 0$ (hedge)
  - For good news index (i.e., index $\uparrow$ means good news) $\Rightarrow \alpha > 0$ (risky)
- Let's check the tone of news indices (next slide)

Tone adjustment





- Tone adjustment: Izumi lab's sentiment dictionary
  - sentiment scores are calculated based on relevance to stock return performance (positively toned words are associated with positive stock returns)
- Env. reg.: two series are similar $\Rightarrow$ orig. series is a good news index
- Nat. dis.: two series go opposite $\Rightarrow$ orig. series is a bad news index

Conculsions

- We construct climate change indices by applying NLP tech to Nikkei newspaper data
- Our 10 climate indices corroborate multi-faceted nature of climate change and seem to capture time-varying concerns specific to Japanese audience
- Asset pricing exercises suggest that some climate factors are priced in Japanese stocks in a statistically and economically significant manner
  - More signifiant results in latter period ("natural disaster" & "energy policy")
- Studying determinants of climate betas is an important remaining issue
  - climate betas may evolve as nature of risk and people's perception change
  - More firm-level environmental and ESG data coverage is key for good investment decisions and research

Literature: Role of media, more NLP literature

- Media as information intermediaries
  - delegated information collecting/curating agent: Nimark and Pitchner (2019, JET), Larsen and Thorsrud (2019, JoE), Larsen et al. (2021, JME), ter Ellen et al. (2020, JMCB). Chahrour et al. (2021, AER).
  - updating beliefs by news: McCombs and Shaw (1972), Ball-Rokeanch and DeFleur (1976), Gentzkow and Shapiro (2006, JPE), Chong and Druckman (2007), Sampei and Aoyagi-Usui (2009)
- NLP in econ. & finance, survey: Gentzkow et al. (2019, JEL)
- Early literature: Tetlock (2007, JF), Loughran and McDonald (2011, JF), Baker et al. (2016, QJE), Engle et al. (2020, RFS)
- LDA analysis: Bandiera et al. (2017, JPE), Larsen and Thorsrud (2019, JoE), Nimark and Pitschner (2019, JET), Faccini et al. (2021), Dierckx et al. (2021)

Literature: env./ESG variables and firms'/stock performance

- Environemntal/ESG variables and performance (ROA, Tobin's q, capital cost)
  - cleaner and better: Aggarwal and Dow (2012), Matsumura et al. (2014), Friede et al. (2015), Garvey et al. (2018), Tzouvanas et al. (2020), Nishitani and Kokubu (2012), Fujii et al. (2013), Aruga et al. (2022)
  - dirtier and better: Wang et al. (2014), Busch et al. (2020)
- Environmental/ESG variables and stock returns
  - dirtier and higher: Renneboog et al. (2008), Bolton and Kacperczyk (2021, 2022), Hsu et al. (2021), Ilhan et al. (2021), Cao et al. (2021)
  - cleaner and higher:In et al. (2019), Choi et al. (2020), Cheema-Fox et al. (2021), Yuyama et al. (2019), Goshima and Yagi (2021)
- Measurement errors and bias in environmental and ESG variables
  - Berg et al. (2020), Aswani et al. (2021), Avramov et al. (2021)

back

Appendix: details of textual data pre-processing (1)

Idea of NLP

Convert documents to (huge) vetor/matrix expression and quantitatively analyze

- cleaning documents in Japanese: remove two-byte symbols (e.g., ※■☆) and URLs, convert two-byte alphabets and digits, replace figures to 0, etc.
- tokenization (wakachigaki): Japanese doesn't seprate words by space ⇒ need to separate sentences into words.
  - There are standard R/Python packages for this

example of wakachigaki

「東京は日本の首都です (Tokyo is the capital of Japan)」

⇒ 「東京、は、日本、の、首都、です」

Appendix: details of textual data pre-processing (2)

- Remove uninformative words and lemmatization
  - grammatical words (e.g., article "the") are not informative from topic modeling perspective (we keep nouns, verbs, and adverbs only)
  - Stop words: too frequently appearing words do not help identifying topics (e.g., I, you).
  - Lemmatization: conjugated words to base form

Good night, good night, parting is such sweet sorrow

$\Downarrow$

good, night, good, night, part, sweet, sorrow

Appendix: details of textual data pre-processing (3)

- Transform cleaned word list into numerical expression
- LDA uses Bag-of-Words (BoW) expression
- BoW expression ignores context (word ordering). More recent techs (e.g., transformer) can consider context

Good night, good night, parting is such sweet sorrow

$\Downarrow$

good, night, good night, part, sweet, sorrow

word	good	night	part	sweet	sorrow
ID	1	2	3	4	5
count	2	2	1	1	1

back

LDA: theoretical background

- $v \in V$: word v and list of words V
- $d \in D$: document d and collection of documents (corpus) D
- $p_{d,v}$: probability weight of picking word v in document d
- $x_{d,v}$: counts of the appearance of word v in document d
- Likelihood of given corpus is $\prod_{d \in D} \prod_{v \in V} p_{d,v}^{x_{d,v}}$

LDA models prob. distr. $p_{d,v}$ by "topic" β_k^v and "topic share" θ_d^k

$$p_{d,v} = \theta_d^1 \beta_1^v + \cdots + \theta_d^K \beta_K^v$$

- β_k^v is $V \times 1$ vector (prob. distr. over list of words)
- $(\theta_d^1, \ldots, \theta_d^K)$ is prob. weight over topics

- Estimate topics β_k^v and topic shares θ_d^k to maximize likelihood back

LDA analysis: estimated topics by word clouds ($K = 10$) English

back

Climate beta and environmental variables

表 3. **Correlation between climate betas and CO2 emission**

	int'l summit	security	env. reg.	env. tech	life style	N
FM-type	-0.01	0.01	0.04	-0.02	-0.02	407.2
pooled	-0.01	0.01	0.06	-0.02	-0.01	6923
	market	energy pol.	tax &ss	nat. disaster	green fin.	
FM-type	0.01	-0.01	-0.03	0.01	-0.10	
pooled	0.04	-0.03	-0.02	-0.00	-0.08	

climate betas are those estimated by monthly-freq (60months) and FF3

- No visible linear correations for most indices
- "Green finance" beta and CO2 emissions exhibit (slight) negative correlation
  - Less CO2 emission ⇒ higher green finance climate beta

  ⇒ Greener stocks' prices increase more when "green finance" articles increase
- Caveat: very narrow coverage of CO2 data (cross-section: ca. 400 stocks)