Supplementary material for Keller et al., "News Media Coverage of Climate Change in India 1997-2016: Using Automated Content Analysis to Assess Themes and Topics", *Environmental Communication*, **2019.**

Appendix. Further considerations regarding our method.

Figure S1. Absolute occurrences, theme: "Climate Change Impacts".

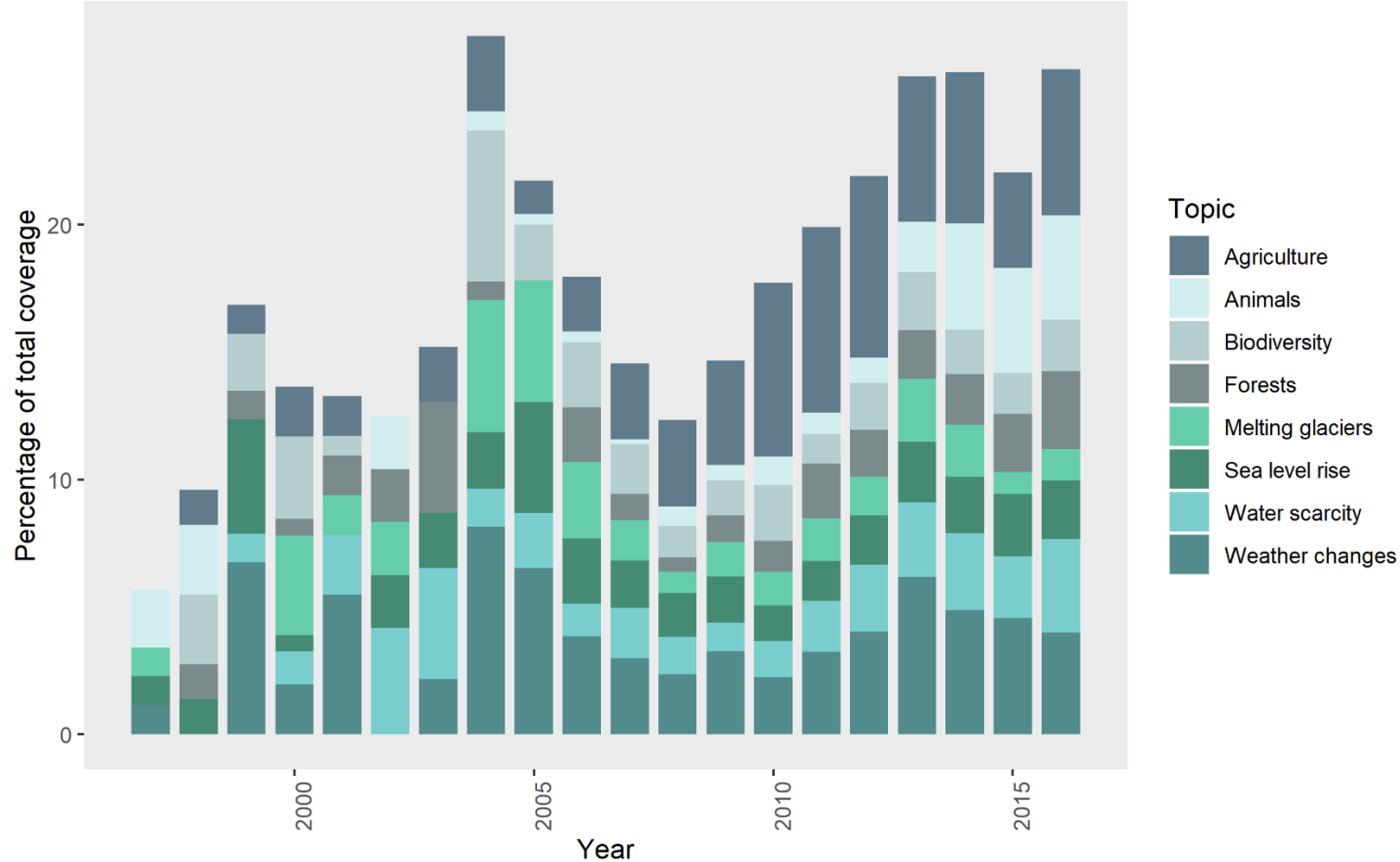**Figure S2.** Percentage of total coverage, theme: "Climate Change Impacts".

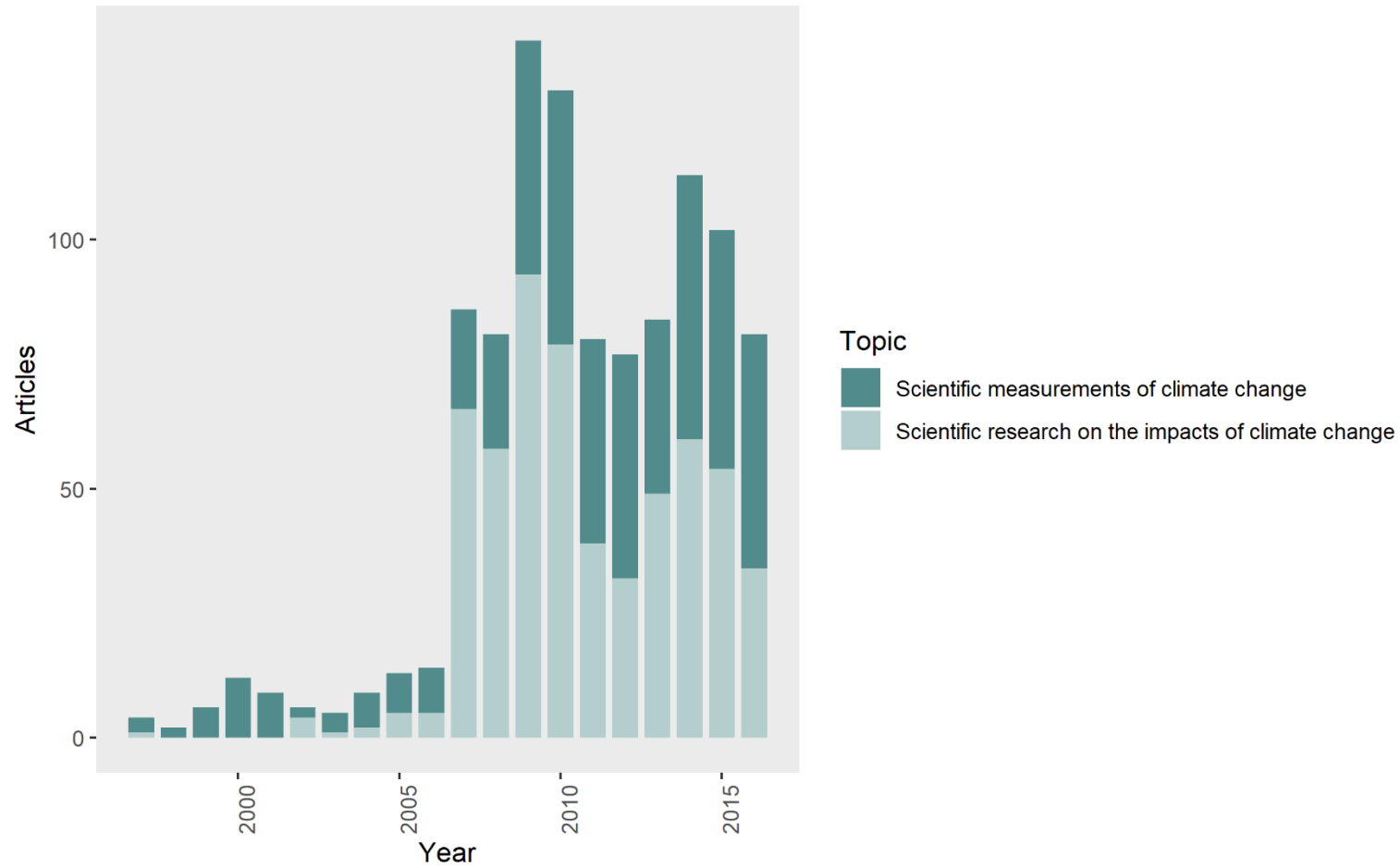**Figure S3.** Absolute occurences, theme: "Climate Science".

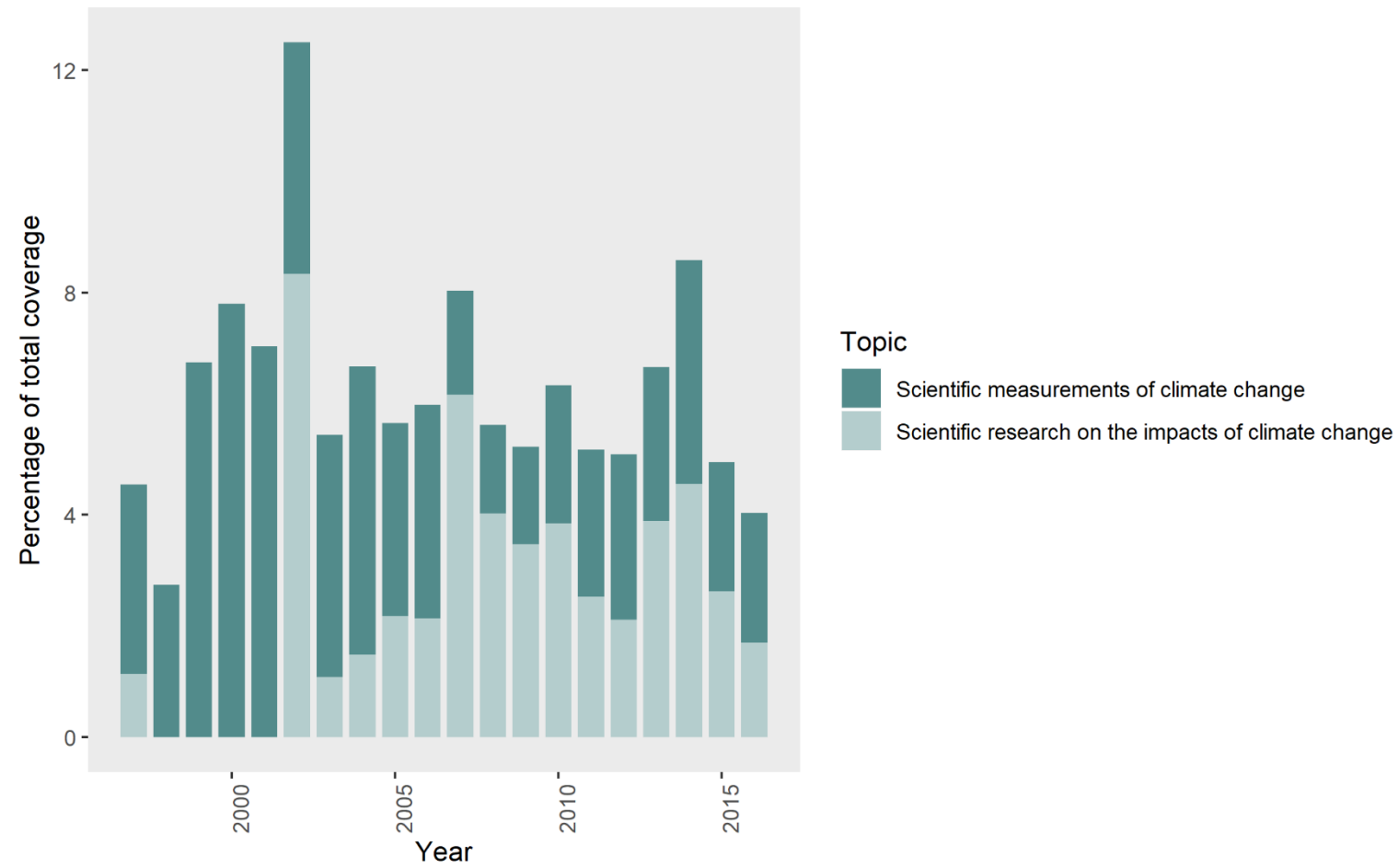**Figure S4.** Percentage of total coverage, theme: "Climate science".

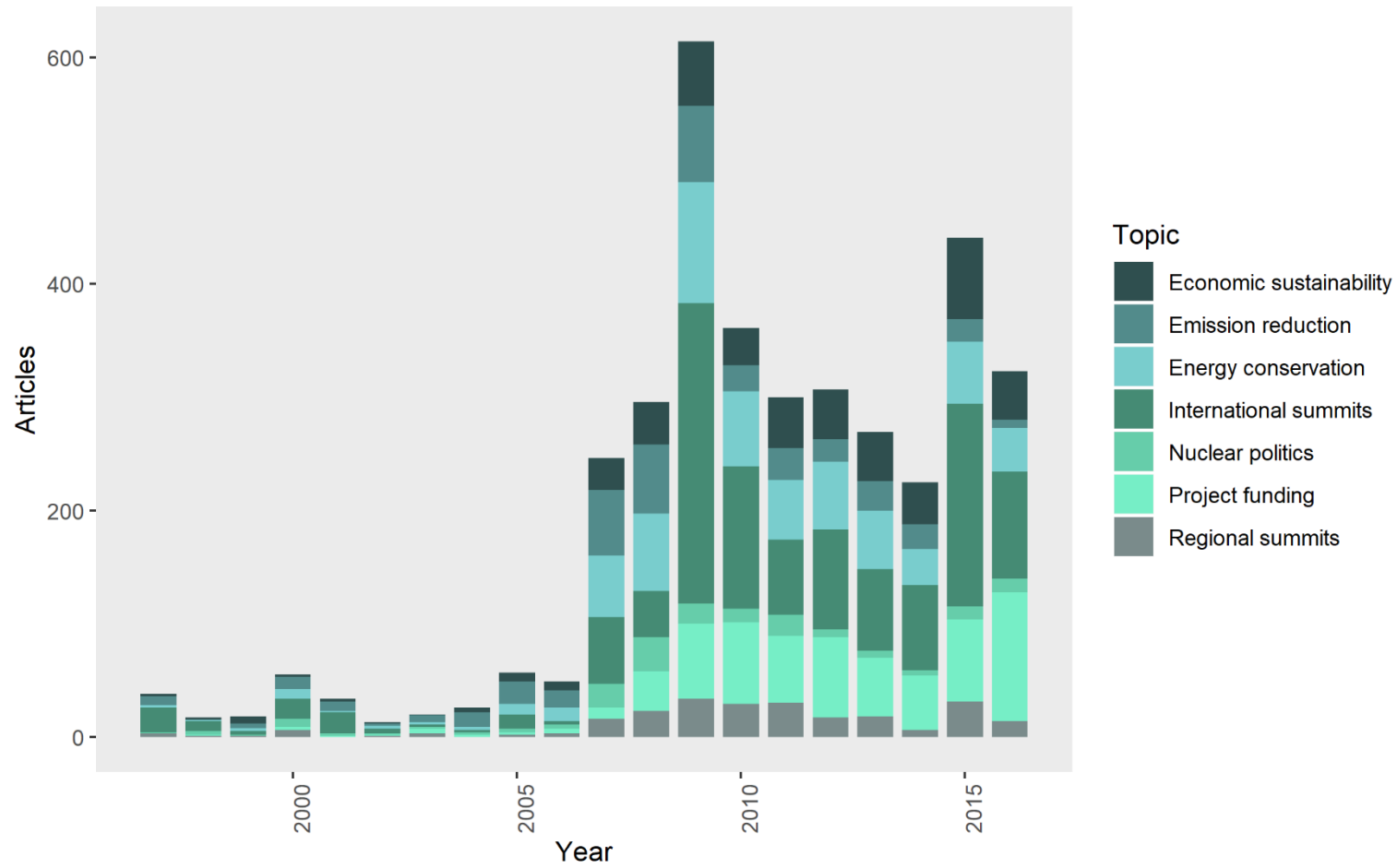**Figure S5.** Absolute occurences, theme: "Climate Politics".

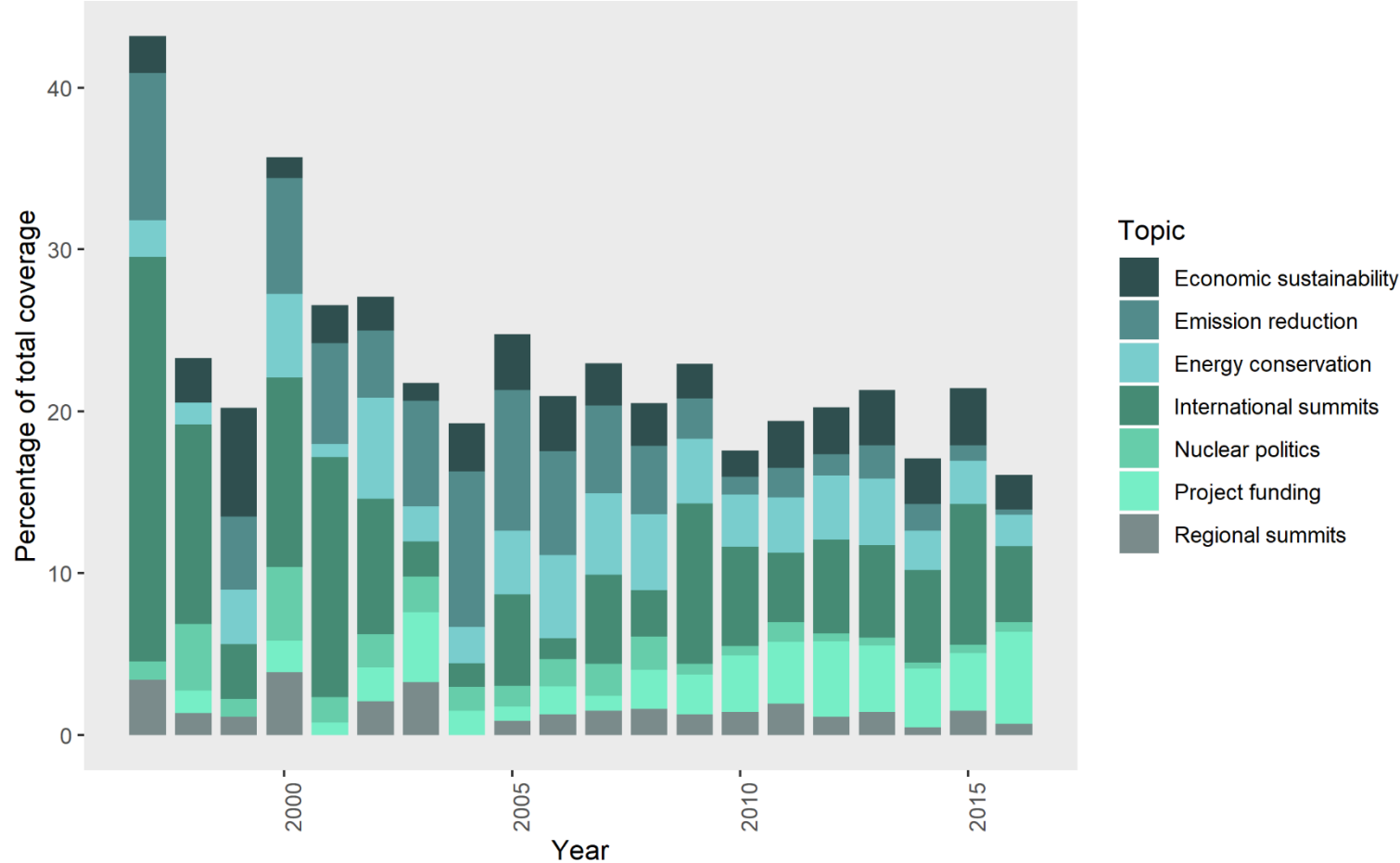**Figure S6.** Percentage of total coverage, theme: "Climate Politics".

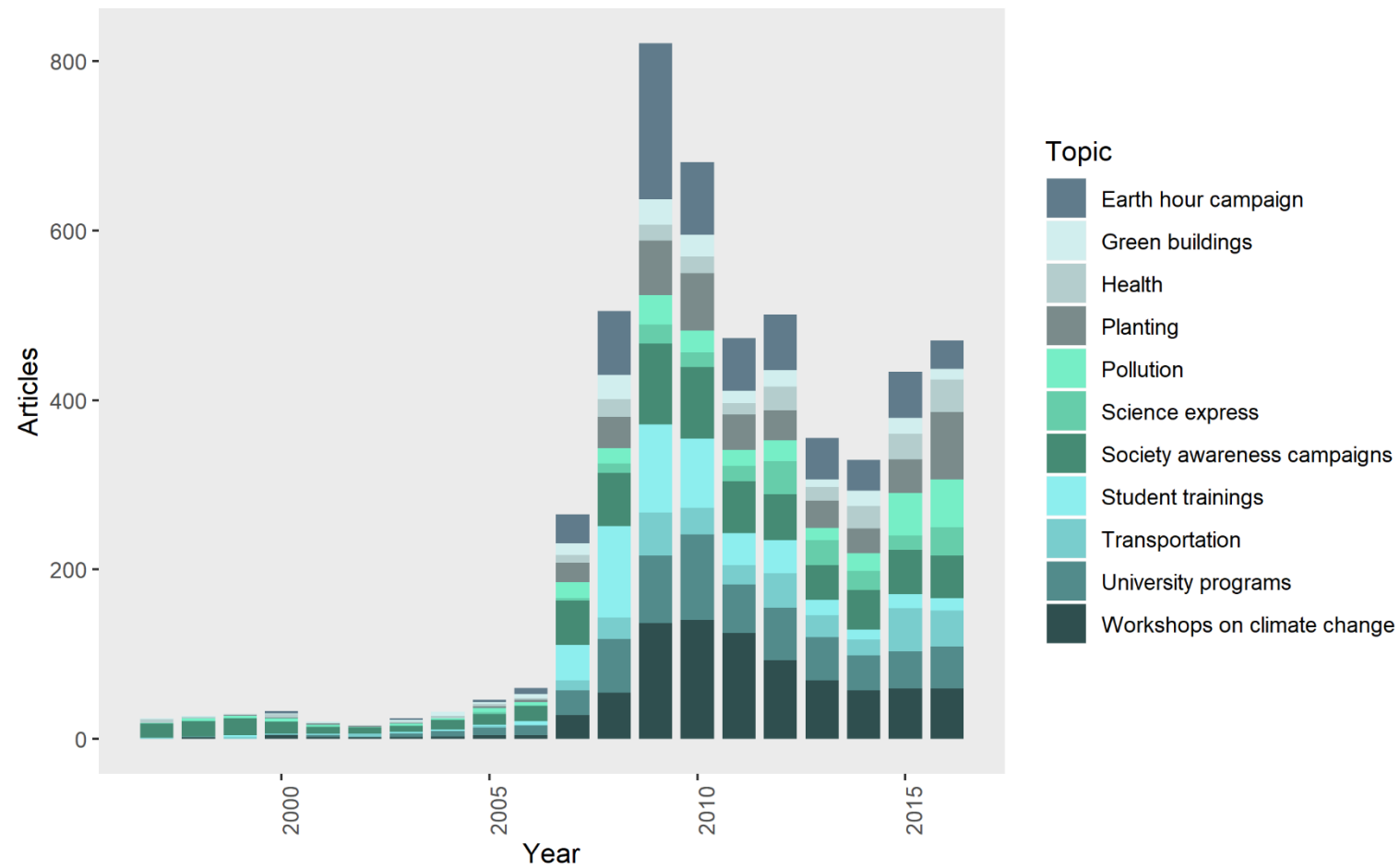**Figure S7.** Absolute occurences, theme: "Climate and Society".

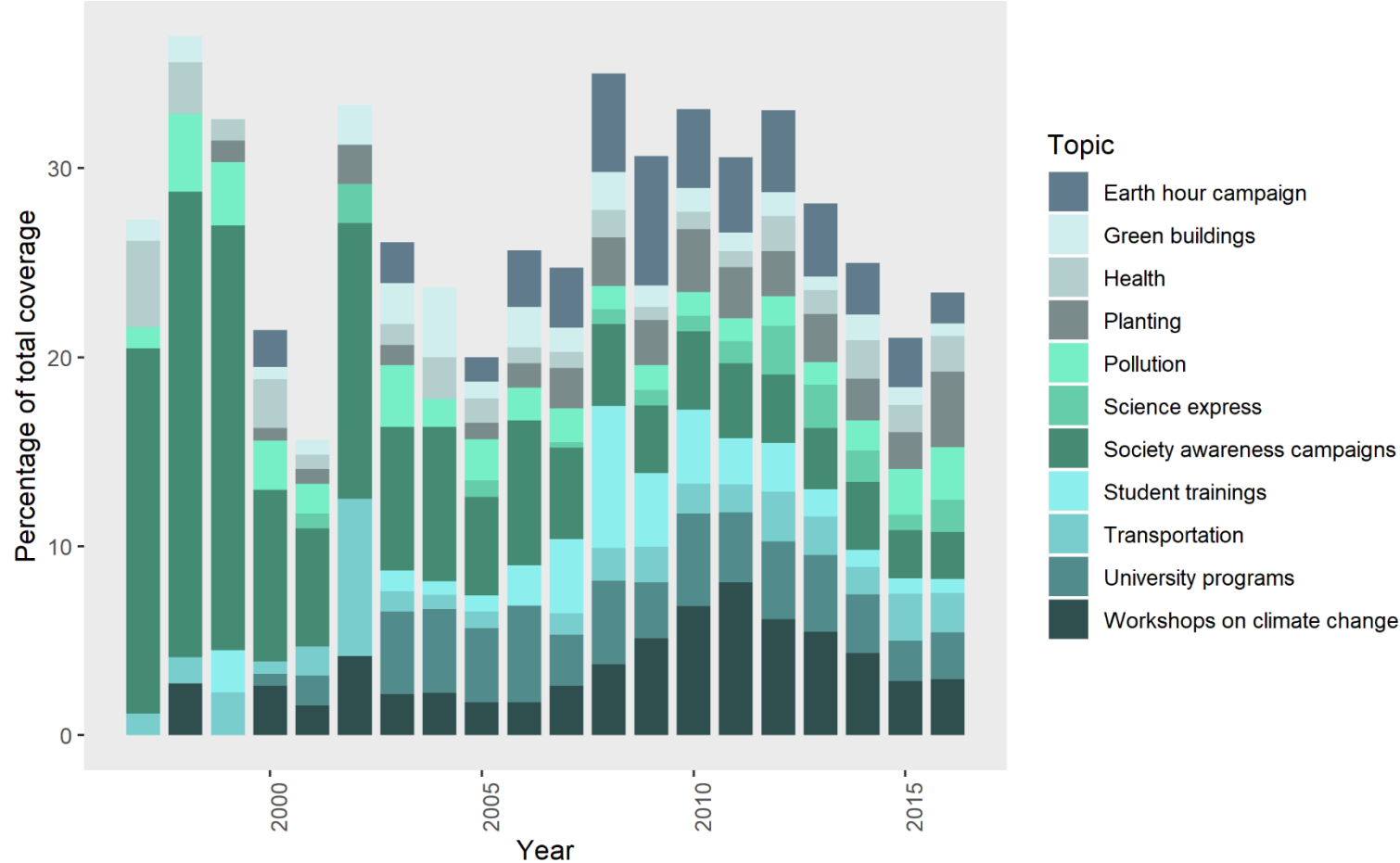**Figure S8.** Percentage of total coverage, theme: "Climate and Society".

Table S1. Robustness and external validity tests.

Label	*The Hindu/ Times of India*	k	seed	Word intrusion	Topic intrusion	Reproducibility and validity	Overall evaluation
Theme: "Climate Change Impacts"							
Agriculture	H, T	100%	100%	75%	100%	Topic appears in both newspapers; stable; high external validation	High
Weather changes	H, T	100%	90%	75%	100%	Topic appears in both newspapers; stable; high external validation	High
Water scarcity	H, T	100%	100%	75%	100%	Topic appears in both newspapers; stable; high external validation	High
Sea level rise	H, T	60%	80%	75%	100%	Topic appears in both newspapers; relatively stable; high external validation	High
Animals	H, T	30%	10%	100%	100%	Topic appears in both newspapers; not stable (merges with *Biodiversity* and *Forests*); high external validation	Medium
Biodiversity	T	0%	0%	75%	100%	Topic appears only in *Times of India*; not stable (merges with *Forests*); high external validation	Low
Forests	H	10%	20%	75%	100%	Topic appears only in *The Hindu*; not stable (merges with *Biodiversity*); high external validation	Medium
Melting glaciers	H, T	40%	60%	100%	100%	Topic appears in both newspapers; relatively stable (often turns from melting glaciers to loss of tourism); high external validation	High
Theme: "Climate Science"							
Scientific research on the impacts of climate change	H, T	90%	90%	100%	100%	Topic appears in both newspapers; stable; high external validation	High
Scientific measurements of climate change	H, T	40%	20%	75%	100%	Topic appears in both newspapers; not stable (often turns from measurements to specifications of data systems in orbit); high external validation	Medium

Theme: "Climate Politics"							
International summits	H, T	100%	100%	75%	100%	Topic appears in both newspapers; stable; high external validation	High
Energy conservation	H, T	100%	90%	100%	100%	Topic appears in both newspapers; stable; high external validation	High
Project funding	H, T	100%	90%	25%	100%	Topic appears in both newspapers; stable; high external validation for documents, low for words as topic was hard to recognize based on top 5 words for external reviewers	High
Economic sustainability	T	10%	10%	75%	50%	Topic appears only in *Times of India*; not stable; medium external validation due to broad coverage of different issues	Low
Emission reduction	T	100%	100%	75%	100%	Topic appears only in *Times of India*; stable; high external validation	High
Regional summits	H	10%	30%	100%	100%	Topic appears only in *The Hindu*; not stable (often merges with political background topics); high external validation	Medium
Nuclear politics	H	40%	40%	0%	100%	Topic appears only in *The Hindu*; relatively stable (often merges with background topic *American politics*); high external validation for documents, low for words topics was hard to recognize based on top 5 words for external reviewers	Medium
Theme: "Climate Change and Society"							
Workshops on climate change	H, T	30%	50%	25%	100%	Topic appears in both newspapers; not stable (often splits up into two more specific event related topics); high external validation for topics, less for words as top terms do not explain the topic easily	Medium
Society awareness campaigns	H	10%	20%	100%	75%	Topic appears only in *The Hindu*; not stable (seems to be noisy as it contains many different awareness campaigns); high external validation	Low
Earth hour campaign	H, T	90%	100%	25%	100%	Topic appears in both newspapers; stable; high external validation for documents, low for words as reviewers did not know the Indian campaign	High
University programs	H, T	100%	100%	100%	100%	Topic appears in both newspapers; stable; high external validation	High

Label	H/T					Description	
Student trainings	H	70%	80%	50%	25%	Topic appears in both newspapers; stable; high external validation for documents, low for words as top five terms are specific for Indian coverage and were not recognized by reviewer	Medium
Planting	H, T	90%	100%	50%	100%	Topic appears in both newspapers; stable; high external validation	High
Transportation	H, T	50%	30%	100%	100%	Topic appears in both newspapers; not stable (merges with "*Pollution*"); high external validation	Medium
Pollution	T	50%	60%	75%	100%	Topic appears only in *Times of India*; relatively stable (with more topics stronger focus on waste disposal and air pollution); high external validation	Medium
Health	H, T	100%	100%	100%	100%	Topic appears in both newspapers; stable; high external validation	High
Science express	H, T	50%	70%	75%	100%	Topic appears in both newspapers; relatively stable; high external validation	High
Green buildings	H, T	20%	20%	100%	100%	Topic appears in both newspapers; not stable (often turns to stronger focus on constructing sustainable cities); high external validation	Medium

Column "Label: The first column contains the name of the topic under discussion.

Column "The Hindu/Times of India": The second column illustrates whether the topic appears when running the topic model separately for *The Hindu* and the *Times of India*. Here, we can see that the newspapers mostly overlap when it comes to the variety of topics, but that some topics only appear in either one of the newspapers. Specifically, there seem to be differences related to "Climate Politics", a them that is discussed differently by the newspapers.

Column "k": The third column relates to running the model with ten different k, more and less topics than our final solution (k between 35 and 45). Based on the results, we check which topics merge or differentiate further when running the model with a lower or higher number of k. To analyze which topics seem to be the most robust or the furthest from other topics in the 38-k-solution, we counted in how many runs of the LDA from $k = 35$ to $k = 45$ the topic was reproduced similarly to our final solution, meaning that at least 7 out of the 10 words with the highest probability stayed within the top-10-words of the reproduced solution. It should be considered, however, that this is a very conservative approach to check for reproducibility. If topics show low percentage values (meaning that, for example, the topic *Biodiversity* could not be reproduced with a different k-solution) this does not infer that the topic only emerges in our final model. In the example of *Biodiversity*, the topic might simply merge with other topics, e.g. become a bit broader and also include the protection of *Forests* or *Animals* leading to a less exact conformity on top terms of this topic. Also, some topics will not be robust with a lower number of k as topics have to merge when running the model with less topics than the proposed solution of $k = 38$. For the future, we would advise to take on more advanced procedures by comparing the results of LDA based on similarity measures of the word-topic-assignment vectors of different models (e.g. Niekler & Jähnichen, 2012).

Column "seed": The fourth column relates to running the model with 10 different random seeds. Percentage agreement was then calculated as described before.

Columns "word intrusion", "topic intrusion": Word and topic intrusion relates to external validation of each topic by human coders, in percentage of agreement. Within the word intrusion test, reviewers were presented with a topic consisting of 6 terms: Five of these were terms with the highest probability for this topic. The additional term, in contrast, did not fit: While the sixth term was chosen to be within the top 10 terms of any of the other topics, it did not score among the top 500 words for the topic to be validated. The external reviewers then had to identify the term not suitable to describe this topic. For the topic intrusion approach, two external reviewers were presented with a document and three topics. While the document had very high probability for this topic, it had close to zero probability for the other two which were randomly chosen from all other topics. Reviewers then had to identify the topic best suitable to describe the specific article.

Column "reproducibility and validity", "overall evaluation": Both columns are evaluations of the previous measures in the table. They are meant to give the reader an overview as to how reproducible and valid the different topics are.

Table S2. Background topics.

k	Label	Occurrence (% of articles)	Top-5 Words
Background topics			
3	*Student competitions*	771 (4.23%)	student, school, children, competit, educ
10	*International Indo-China affairs*	667 (3.66%)	india, china, indian, countri, issu
14	*City news/travel*	635 (3.48%)	say, year, time, peopl, work
8	*Mining policy*	628 (3.44%)	ministri, state, govern, environ, court
23	*Cultural events*	520 (2.85%)	film, festiv, art, global, music
22	*American politics*	366 (2.01%)	polit, american, presid, obama, state
28	*Indian politics*	321 (1.76%)	minist, govern, state, prime, singh
12	*Economic growth*	309 (1.69%)	cent, year, india, countri, price
9	*Awards/honors*	273 (1.5%)	women, award, right, villag, gandhi
6	*Public events*	236 (1.29%)	colleg, sri, nagar, road, templ

Explanations on background topics

In addition to the 28 climate-relevant topics that were sorted into four overarching themes, the topic modeling generated ten additional topics from the article corpus that deal with climate change only in passing. We have categorized these topics as *Background topics* similar to previous studies (Maier et al., 2018). Other than climate-relevant topics, they provide contexts in which climate change related key words occur. *Student competitions* is the most frequent topic in this category. Articles related to this topic cover different types of school competitions, such as badminton, yoga, quiz competitions, or various entertainment programs. Climate change plays a subordinate role here. These articles are part of our sample as the events are for example held on World Anti-Global Warming Day, or students present working models on topics such as satellites, traffic regulations, climate change and its causes. Within the topic *International Indo-China affairs*, long-running boundary negotiations between India and China and fight against terrorism in Pakistan are addressed. In the light of these negotiations, the increasing convergence of both countries' positions on multilateral issues such as climate change or cooperation for a groundbreaking climate protection agreement in Paris are discussed in passing. Under the topic *City news/Travel*, various anecdotal stories from people's everyday life or short stories for children are summarized. In these stories, climate change is often mentioned only in sub-phrases, as a topic that concerns the society. Articles on *Mining policy* deal with mining in India. Since the Ministry of environment, forest and climate change has the task of deciding on the renewal of mining leases, these articles were part of our sample. The topic *Cultural events* includes various social events or festivals at which artists or authors address climate change, however among numerous other topics. Furthermore, celebrities such as Sienna Miller, as part of an organization committed to prevent global warming (the "Global Cool"-project), are reported on in the context of a new documentary on this topic. *American politics* articles deal with US foreign and domestic policy, including disputes between Republicans and Democrats on climate change and immigration. Similarly, the topic of *Indian politics* discusses various domestic policy issues in India, which sometimes have only marginal links to climate change. Articles on *Economic growth* deal with India's economic situation and central influencing factors, such as the oil prices, which are rising due to the monsoons caused by climate change. Articles on public memorial celebrations, spiritual and cultural events or the honoring of social actors who work to combat social evils, caste restrictions and the creation of a classless society are can be summarized as *Awards/honors*. Articles on the topic of *Public events* contain numerous local events that take up literature seminars, meditations or psychological counselling for victims of violence and, in passing, mention climate-change-related events.