

1 Text-Mining the Signals of Climate Change Doubt

2 Constantine Boussalis^a, Travis G. Coan^{b,*}

3 ^a*Department of Political Science, Trinity College Dublin, 3 College Green, Dublin 2, Ireland*

4 ^b*Department of Politics and Exeter Q-Step Centre, University of Exeter, Amory Building,*
5 *Rennes Drive, Exeter, Devon EX4 4RJ, United Kingdom*

6 Abstract

Climate scientists overwhelmingly agree that the Earth is getting warmer and that the rise in average global temperature is predominantly due to human activity. Yet a significant proportion of the American public, as well as a considerable number of legislators in the U.S. Congress, continue to reject the “consensus view.” While the source of the disagreement is varied, one prominent explanation centres on the activities of a coordinated and well-funded countermovement of climate sceptics. This study contributes to the literature on organized climate scepticism by providing the first systematic update of conservative think tank counter-claims in nearly 15 years. Specifically, we 1) compile the largest corpus of climate sceptic claims-making activity to date, collecting over 16,000 documents from 19 organizations over the period 1998 to 2013; 2) introduce a methodology to measure key themes in the corpus which scales to the substantial increase in content generated by conservative think tanks (CTTs) over the past decade; and 3) leverage this new methodology to shed light on the relative prevalence of science- and policy-related discussion among CTTs. We find little support for the claim that “the era of science denial is over”—instead, discussion of climate science has generally increased over the sample period.

7 *Keywords:* climate change, scepticism, text classification, latent Dirichlet
8 allocation

9 1. Introduction

10 Climate scientists overwhelmingly agree that the Earth is getting warmer
11 and that the rise in average global temperature is predominantly due to human
12 activity (IPCC 2014, National Research Council 2010, Oreskes 2004, Doran and
13 Zimmerman 2009, Anderegg et al. 2010, Cook et al. 2013). Yet a sizeable seg-
14 ment of the American public rejects this “consensus view” (Weber and Stern
15 2011) and U.S. climate policy remains in a state of limbo. As of early 2015,
16 one-third of the American public believes that climate change is *not* primarily
17 caused by human activity and only one in ten understands that more than 90% of
18 climate scientists agree on the existence and nature of observed global warming
19 (Leiserowitz et al. 2015). What explains this divergence in views among climate

* Corresponding author: T.Coan@exeter.ac.uk. Replication data and code are provided at <https://www.elsevier.com/locate/elsevier>.
Preprint submitted to Elsevier. January 18, 2016

20 scientists and the American public? What factors promote inaction on compre-
21 hensive climate mitigation policy? These questions have garnered considerable
22 attention in disciplines across the social and behavioural sciences.

23 One prominent explanation investigates the influence of a “well-funded and
24 relatively coordinated ‘denial machine’” on shaping the public’s understanding
25 of climate science (Begley et al. 2007). While a diverse set of actors promote cli-
26 mate scepticism, conservative think tanks (CTTs) play a central role, providing
27 key counter-claims to challenge climate science and obstructing climate policy
28 (McCright and Dunlap 2000). CTTs provide a multitude of services to the cause
29 of climate change scepticism: providing material support and lending credibility
30 to contrarian scientists, sponsoring pseudo-scientific climate change conferences,
31 directly communicating contrarian viewpoints to politicians, and, more gener-
32 ally, disseminating sceptic viewpoints through a range of media to the wider
33 public (Dunlap and McCright 2011). A number of studies also suggest that
34 these organizations are central in obstructing national climate policy (Lahsen
35 2008, Oreskes and Conway 2010) and international climate change mitigation
36 agreements (McCright and Dunlap 2003). The prominence of CTTs in the con-
37 trarian counter-movement has prompted calls for an expansion and improvement
38 of data collection efforts on a range of climate movement and counter-movement
39 activities (Brulle et al. 2012).

40 Despite an active interest in CTTs, few studies have systematically analysed
41 the nature and prevalence of contrarian counter-claims. Aaron McCright and
42 Riley Dunlap’s influential study offers a notable exception, providing a compre-
43 hensive survey of CTT counter-claims from 14 major conservative think tanks
44 over the period 1990-1997. Yet, to our knowledge, there have been no systematic
45 updates to this study over the past 15 years and thus little is known about how
46 contrarian claims have evolved over the last decade. We seek to fill this gap
47 in the literature by 1) compiling the largest corpus of climate sceptic claims-
48 making activity to date, collecting over 16,000 documents from 19 organizations
49 over the period 1998 to 2013; 2) introducing a methodology to measure key
50 themes in the corpus which scales to the exponential increase in content gener-
51 ated by conservative think tanks (CTTs) over the past decade; and 3) leveraging
52 this new methodology to examine the dynamics of policy- and science-related
53 claims over a 16 year period. We argue that understanding CTT counter-claims
54 is of both theoretical and practical significance, as an acceptance of the anthro-
55 pogenic causes of climate change is arguably a necessary condition for progress
56 on reaching a climate agreement and may portend a window for policy action.

57 **2. Understanding contrarian counter-claims**

58 A number of scholars argue that the entrenchment of climate change scep-
59 ticism in American society is not an “accident.” Rather, the dismal state of
60 public understanding of AGW in the United States is largely the result of an
61 orchestrated attack on climate science and individual climate scientists by a

62 constellation of interests that are determined to obstruct policies aimed at miti-
63 gating global warming (Pooley 2010, Oreskes and Conway 2010, Washington and
64 Cook 2011, Mann 2013). For over twenty years, the American public has been
65 subject to waves of information produced by a “well-coordinated, well-funded
66 campaign by contrarian scientists, free-market think tanks and industry” which
67 has “created a paralyzing fog of doubt around climate change” (Begley et al.
68 2007). Employing tactics (and even participants) from similar disinformation
69 campaigns, such as those against the regulation of tobacco and ozone-harming
70 chlorofluorocarbons (CFCs), the counter-movement aims to block climate policy
71 by “manufacturing doubt” about the credibility of individual scientists, misrep-
72 resenting peer-reviewed scientific findings, and exaggerating scientific uncertain-
73 ties (Union of Concerned Scientists 2007, Oreskes and Conway 2010, Greenpeace
74 2010, Dunlap and McCright 2011).

75 While there are a number key actors in what Begley et al. (2007) refer to
76 as the “denial machine” (see Dunlap and McCright 2011 for an overview), the
77 “engine” of information centres on a number of influential CTTs. CTTs seek
78 to manufacture uncertainty in two important ways. First, sceptics have im-
79 plemented a campaign to re-frame the issue of climate change, shifting the
80 story away from consensus and the urgent need for action toward one of “non-
81 problematicity” (Freudenburg 2000, McCright and Dunlap 2003). Communica-
82 tions research repeatedly emphasizes the sensitivity of public perceptions to how
83 an issue is *framed* within the wider information space (Lakoff 2014, Scheufele
84 and Tewksbury 2007). And given the inherent complexity of climate change,
85 “interpretive storylines” surrounding the issue are ripe for manipulation by par-
86 ties on either side of the debate (Nisbet 2009). Second, relying on their image
87 as the “alternative academia” or “counter-intellegentsia,” CTTs play a lead role
88 in constructing viewpoints to challenge orthodox views on climate science and
89 policy (Beder 2001, Austin 2002, Jacques et al. 2008, Dunlap and Jacques 2013).
90 CTT-affiliated contrarian scientists and commentators have generated and dis-
91 seminated numerous counter-claims against climate science and policy action
92 through various forms of media, including books, op-eds, newsletters, policy
93 studies, speeches and press releases (McCright and Dunlap 2000, Jacques et al.
94 2008, Dunlap and Jacques 2013).

95 Studies interested in measuring the prevalence of contrarian claims focus al-
96 most exclusively on the *level* of contrarian information present in media coverage
97 of global warming. These studies have yielded important insights into the preva-
98 lence of skepticism within newspapers (e.g., Boykoff and Boykoff 2004, Painter
99 and Ashe 2012, Schmidt et al. 2013), opinion pieces in print media (Hoffman
100 2011, Elsasser and Dunlap 2013, Young 2013), television (Boykoff 2008, Hart
101 2008, Feldman et al. 2012), and “new media” (O'Neill and Boykoff 2011, Hol-
102 liman 2011, Knight and Greenberg 2011, Sharman 2014, Elgesem et al. 2015).
103 However, few studies systematically analyse the *content* of contrarian claims
104 and even fewer focus specifically on CTTs. To date, McCright and Dunlap
105 (2000) offers the most comprehensive survey of CTT counter-claims on climate
106 change. The authors content analyse a sample of 224 documents related to

107 global warming from 14 major conservative think tanks over the period 1990-
108 1997, with the vast majority of this literature being produced during 1996 and
109 1997. Overall, the analysis suggests that climate scepticism during this period
110 centred on three major counter-claims: 1) the evidentiary basis of global warm-
111 ing is weak or wrong, 2) global warming would be beneficial if it was to occur,
112 and 3) global warming policies would do more harm than good (see [McCright
113 and Dunlap 2000](#) pg. 510, Table 3). For the 1990-1997 period, the study finds
114 that 71% of the documents contained criticisms of the scientific evidence for
115 global warming (Counter-claim 1), only 13.4% discussed the benefits of global
116 warming (Counter-claim 2), and 62.1% provided a discussion on the downsides
117 of climate policy action (Counter-claim 3).

118 McCright and Dunlap’s study provides a unique look at sceptical counter-
119 claims in the mid-to-late 1990s, yet much less is known about how these claims
120 have evolved. Several studies provide a more recent look at the key features of
121 the contrarian discourse more generally. [Elsasser and Dunlap \(2013\)](#) employed
122 John Cook’s list of sceptical arguments (www.skepticalscience.com) to classify
123 203 op-eds over the period 2007-2010. The authors find that personal attacks
124 on Al Gore and scepticism of the IPCC were common throughout the corpus,
125 while “it’s not happening” arguments dominated the discussion, showing up in
126 almost two thirds of the articles. [Sharman \(2014\)](#) examines the climate skeptic
127 blogosphere from March to April of 2012, classifying 171 blog posts as either
128 science- or policy-oriented. The author finds that blogs which are “central” in the
129 blogosphere network tended to focus on discussions of science, while peripheral
130 blogs tended to emphasise policy. Lastly, and more in line with the current
131 study, in a content analysis of documents from the Heartland Institute over the
132 period September-December 2013 ($n = 102$), [Cann \(2015\)](#) finds a considerable
133 drop in discussions of policy when compared to the findings of [McCright and
134 Dunlap \(2000\)](#). As the author acknowledges, however, it is difficult to determine
135 whether this indicates a general move away from policy-oriented claims or is
136 simply a sampling issue associated with focusing on a single organisation for a
137 two month period. More generally, this limitation applies equally to the analysis
138 of op-eds and blogs as well: the existing evidence provides segmented glimpses of
139 the evolution of contrarian claims over the past decade and a half. The remainder
140 of this study seeks to overcome this limitation by providing a comprehensive look
141 at CTT claim-making activity.

142 **3. Measuring contrarian claims**

143 *3.1. The corpus*

144 To systematically gauge claims-making activity, we retrieved information re-
145 lated to climate change from the websites of 19 well-known North American
146 conservative think tanks and organizations (see online appendix for details).
147 Our choice of organizations, to a large extent, mirrors that of [McCright and
148 Dunlap \(2000\)](#) and the most heavily funded organizations which are identified
149 in [Brulle \(2014\)](#). For each organization, we visited all pages including the terms

150 “climate change” or “global warming” and extracted relevant text and key meta
151 data. There were also instances where pages included links to documents in PDF
152 format, which were typically relatively long policy reports. These PDFs were
153 automatically retrieved, passed through optical character recognition (OCR)
154 software to extract the text, and appended to the list of text retrieved from the
155 HTML code. Audiovisual materials were a minority of the overall set of retrieved
156 pages and were excluded in the current analysis. This process produced more
157 than 16,000 documents over the period from 1998 to 2013.

158 Table 1 provides an overview of the organizations included in the sample. The
159 first two columns display the total number of words and documents published
160 online by each organization over the period of study. To provide a general sense of
161 the types of output, the next five columns provide a tabulation of the documents
162 by type, following the classification scheme used in (McCright and Dunlap 2000,
163 p. 508). Relying heavily on meta-data provided within the URL or the document
164 itself, we categorize the documents by five general types: (A) op-eds, articles and
165 blogs, (B) policy/science reports and analyses, (C) speech/interview transcripts,
166 (D) press releases/open letters, and (E) scientific reviews. More information on
167 the document type coding procedure is available in the online appendix.

168 The table provides a number of insights into the claims-making behaviour
169 of the most important CTTs. First, these organisations have increased their
170 production and dissemination of literature exponentially, from roughly 203 docu-
171 ments over the period 1990-1997 (McCright and Dunlap 2000) to 16,028 docu-
172 ments for the years 1998-2013. Second, the distribution of the document classi-
173 fications suggests that the communication strategy of these organizations varies.
174 Several organisations focus on producing shorter, op-ed style documents (e.g,
175 NCPA), while others focus on producing lengthier policy or science-related re-
176 ports (e.g, George C. Marshall Institute). Third, as expected based on past
177 research, the Heartland Institute is a central actor among CTTs, producing or
178 disseminating a significant portion of the documents in the corpus and focusing
179 on a mix of short articles and longer policy reports. We take a closer look at
180 the claims-making trends of Heartland in Section 6.

181 3.2. *Methods: probabilistic topic modelling*

182 The time and effort associated with reading over 16,000 documents renders
183 traditional content analytic approaches inadequate and/or infeasible and thus
184 the next step is to find a suitable computational model to help make sense of
185 the data. We approach this step using an *unsupervised* approach, exploring
186 the presence of meaningful clusters of terms that appear across documents in
187 the collected corpus. While there is no shortage of clustering algorithms in the
188 literature (Grimmer and King 2011), we utilize the latent Dirichlet allocation
189 (LDA) model originally proposed in Blei et al. (2003). LDA provides a statistical
190 framework for understanding the latent topics or themes running through a
191 corpus by explicitly modelling the random process responsible for producing
192 a document. The LDA model assumes that each document is made up of a
193 mixture of topics, as well as a mixture of words associated with each topic. For

Organization Name	Total Words (thous.)	Total Docs.	Document Type				
			A	B	C	D	E
American Enterprise Institute (AEI)	1,872.53	745	596	61	48	15	25
Cato Institute	772.68	768	712	41	8	6	1
Center for the Study of Carbon Dioxide and Global Change (CO2Science)	2,387.27	4,592	713	0	0	1	3,878
Competitive Enterprise Institute (CEI)	1,743.02	1,461	941	55	0	465	0
Committee for a Constructive Tomorrow (CFACT)	738.52	894	882	12	0	0	0
Citizens for a Sound Economy (CSE)	88.2	111	105	6	0	0	0
Fraser Institute	78.39	81	62	19	0	0	0
Foundation for Research on Economics and the Environment (Free-Eco)	76.64	105	105	0	0	0	0
Heartland Institute	9,900.54	2,930	1,383	1,537	10	0	0
Heritage Foundation	1,825.78	1,652	1,198	431	23	0	0
Hoover Institution	51.06	37	3	32	2	0	0
Hudson Institute	124.61	83	81	2	0	0	0
Manhattan Institute	315.59	199	183	13	3	0	0
George C. Marshall Institute	209.75	101	69	21	11	0	0
National Center for Policy Analysis (NCPA)	469.78	451	376	75	0	0	0
National Center for Public Policy Research (NCPPI)	393.54	639	378	90	0	171	0
Pacific Research Institute	384.68	435	402	7	0	26	0
Reason Foundation	397.12	192	179	13	0	0	0
Science and Public Policy Institute (SPPI)	3,064.88	552	0	552	0	0	0
Total	24,894.58	16,028	8,368	2,967	105	684	3,904

Table 1: *Climate sceptic organizations.* The table displays the total count of words (thousands), the number, and type of documents from 19 well-known conservative think-tanks over the period January 1998 – August 2013. Documents have been classified as follows: (A) op-eds, articles and blogs; (B) policy/science reports and analyses; (C) speech/interview transcripts; (D) press releases/open letters; (E) scientific reviews.

194 instance, the document you are reading at this moment includes a mixture of
195 themes such as “climate scepticism” and “text analysis,” and these themes tend
196 to use different language—the topic “climate scepticism” is likely associated with
197 the word “denial,” whereas the topic “text analysis” is associated with the word
198 “random.” Moreover, this process is probabilistic in the sense that we could have
199 used the term “stochastic” instead of “random” in the previous sentence.

200 This basic generative story provides the basis for a simple hierarchical Bayesian
201 model based on the following assumptions: 1) each word in a text is exchange-
202 able, each text in a corpus is a combination of a specific number of topics (T_k),
203 and each specific topic is represented as a distribution of words (w) over a fixed
204 vocabulary (Blei et al. 2003, Griffiths and Steyvers 2004). The generative struc-
205 ture that produces each document in a corpus is represented as random mixtures
206 of latent topics and their associated distributions of words. Specifically, the LDA
207 assumes that documents are generated from the following probabilistic process:

- 208 1. Each of the k topics are drawn from a topic distribution by

209

$$\theta \sim \text{Dirichlet}(\alpha)$$

210

2. The term distribution β for each topic is represented by

211

$$\beta \sim \text{Dirichlet}(\eta)$$

212

3. For each of the N words w_n :

213

Randomly sample a topic $z_n \sim \text{Multinomial}(\theta)$.

214

Choose a word w_n from $p(w_n|z_n, \beta)$.

215

216 Although this model provides an overly simplified representation of the true
217 data generating process for text, it has been shown to be effective in applied
218 situations and employed in a diverse range of fields, from population biology to
information retrieval (see [Blei 2012](#) for an overview).

219

3.2.1. How many topics?

220

221 LDA requires one to specify the number of topics *a priori*. This presents
222 an obvious challenge when studying contrarian counter-claims, as past research
223 suggest anywhere from 9 claims ([McCright and Dunlap 2000](#)) to 176 “debunked
224 climate myths” (www.skepticalscience.com). While a range of methods have
225 been introduced in the literature to estimate the “natural” number of topics
226 (see [Wallach et al. 2009b](#) for an overview), there remains considerable debate on
227 the utility of data-driven approaches for generating interpretable topics ([Chang
et al. 2009](#)). Moreover, when applying probabilistic topic models to understand
228 social phenomena, the “natural” number of topics is conditional on the particular
229 research question of interest. If answering your question requires a high degree
230 of detail, then using a larger number of topics is advisable; otherwise, little
231 substantively meaningful information is lost by assuming a smaller number of
232 topics ([Quinn et al. 2010](#), [Roberts et al. 2014](#)).

233

234 With little theoretical guidance on the appropriate number of topics, we
235 employ a balanced approach between data-driven methods and a qualitative
236 assessment of the interpretability of the latent space. First, we rely on the topic
237 selection criteria proposed in [Arun et al. \(2010\)](#), which has proven an effective
238 heuristic for determining a reasonable topic number in both real and synthetic
239 datasets (see the online appendix for technical details). Using the Arun et al.
240 procedure as a starting point, we then systematically adjusted the assumed topic
241 number (k) around the “optimal” data-driven result and manually assessed the
242 quality of the topic solutions. While the details of this analysis are available in
243 the online appendix, we find that $k = 53$ offers a suitable balance between having
244 a manageable number of topics, enough detail to assess core substantive themes
245 in climate contrarianism, displaying a reasonable level of “fit” using data-driven
methods, and demonstrating stability across a range of solutions.

246

4. Results

247

4.1. Model estimation and topic interpretation

248

249 We estimate the model using the sparse Gibbs sampler described in [Yao
et al. \(2009\)](#) and the hyperparameter optimization routine utilized in [Wallach](#)

250 et al. (2009a). Consistent with the findings in Wallach et al. (2009a), we found
251 that optimizing α , while fixing β , provided the easiest results to interpret and
252 thus employ this specification. Moreover, given that mixture models such as the
253 LDA are known to produce multimodal likelihood surfaces, we used a number
254 of different random starting values. We found a good deal of stability in the
255 estimated topic distributions across runs, improving our confidence that the
256 model converged on a global optimum.

257 After removing 6 “junk” topics (AlSumait et al. 2009),¹ our final list in-
258 cludes 47 substantively meaningful topics representing a range of issues related
259 to global warming. Table 2 provides a complete list of the estimated topics of
260 the sceptical discourse. To ease interpretation, we produce a descriptive label for
261 each topic by reading the 10 most probable documents and noting the key theme
262 consistent within each sub-sample. The descriptive labels not only provide use-
263 ful information to facilitate topic interpretation, but also offer a first look at one
264 aspect *semantic validity*: the extent to which each topic is coherent in terms of
265 its meaning (Quinn et al. 2010). We also include a set of keywords for each topic
266 based on the word’s “frequency-exclusivity” (FREX), as described in Roberts
267 et al. (2014). FREX offers a balance between the probability (or “frequency”) of
268 a word being associated with a particular topic and the extent to which a word
269 is unique to a topic (i.e., “exclusivity”).

270 Looking at the full list of topics shown in Table 2, the results demonstrate a
271 good level of face validity and are generally consistent with the themes discussed
272 in McCright and Dunlap (2000). These topics touch on a wide range of themes
273 such as scientific integrity and uncertainty, climate change impacts, energy, en-
274 vironmental policy, society, as well as domestic and international politics. And,
275 as expected, the corpus is rife with claims surrounding the uncertainty of cli-
276 mate scientific studies. The notion that human activity, specifically the emission
277 of greenhouse gases into the atmosphere, is leading to a rise in global tempera-
278 tures (topic 1) has been characterized as suffering from a “real-world disconnect”
279 (Heartland Institute, Nov. 11, 2011) and any discussion to the contrary amounts
280 to “alarmism” (Heartland Institute, May 17, 2013). Further, the general agree-
281 ment of scientists on this relationship is repeatedly refuted within the corpus
282 (topic 4) as there is “no consensus on climate change” (NCPR, March 22, 2004).
283 Appeals to long-term natural cycles in temperature (topic 5), as purportedly
284 demonstrated by the Roman and Medieval Warm Periods, are common support
285 for arguments against anthropogenic global warming. This topic is of particular
286 interest as it was not detected in McCright and Dunlap (2000) and has become
287 a common claim among climate sceptics. Studies that support anthropogenic
288 global warming are also deemed to be “fabricated” and have led to a “childish
289 panic.” Typical examples of these arguments include:

¹AlSumait et al. (2009) note that not all topics in an estimated topic model are of equal importance and it is not uncommon to have a set of “junk” topics that pick up common co-occurrences of words with little or no substantive meaning.

Id	S/P	Topic Name	Id	S/P	Topic Name
42	S	Acidification calcif reef bleach coral phytoplankton	20	P	Corporations & env. borelli sharehold greenpeac donor philanthropi
16	S	Alarmism gore morano romm inconveni depot	43	P	Disaster costs insur pension mortgag florida premium
11	S	Climate models simul gcm model cmip coupl	25	P	Economic impact of climate policy baselin discount sector eia mit
1	S	Climate sensitivity to CO2 warm degre cool dioxid warmer	29	P	Emissions reduction carbon scheme credit trade dioxid
46	S	Endangered species butterfli stirl extinct bear polar	10	P	Environmentalism lomborg holdren ehrlich evangel simon
34	S	Forest impacts npp ndvi shrub peatland finzi	38	P	EPA caa epa endanger naaq anpr
19	S	Human health ddt precautionari malaria diseas cancer	2	P	Fossil fuel production shale barrel oil drill pipelin
27	S	IPCC integrity chapter ipcc tsd wg summari	15	P	Govt. agencies fy sec gao omb provis
5	S	Long-term climate trends holocen millenni quaternari mediev palaeo	9	P	Govt. intervention approach intervent principl geoengin outcom
26	S	Monckton monckton graph ppmv brenchley humankind	24	P	Green jobs job stimulu taxpay subsidi green
4	S	No scientific consensus consensu denier oresk agw scientif	44	P	Int'l climate agreements kyoto protocol treati ratifi ratif
30	S	Plant impacts seedl leaf mycorrhiz cultivar elev	17	P	Int'l relations militari nato missil afghanistan iran
45	S	Pollution mercuri ozon toxic asthma particul	31	P	Int'l trade & develop india china chines wto asia
14	S	Scientific misconduct cru mcintyr mann hockey email	39	P	Law court judici lawsuit constitut suprem
3	S	Sea level rise antarct greenland glacier melt antarctica	23	P	Nuclear power hydrogen reactor nuclear technolog cell
12	S	Solar forcing & cloud models cosmic cloud radiat ray aerosol	6	P	Public opinion gallup abc pew cnn cb
40	S	State climate reports viru cessat Nile wigley inch	36	P	Public transportation rail ridership travel passeng vmt
28	S	Storms cyclon storm hurrican tc frequenc	8	P	Renewable energy rp turbin renew wind megawatt
13	S	Temperature station data station giss ushcn fig thermomet	22	P	Reuse & recycle bag mtbe bulb cfl reus
18	P	Agri. Industry corn ethanol biofuel farmer sugar	41	P	State climate policy ghg jersey greenhous wefa rggi
47	P	Auto. fuel standards cafe nhtsa mpg vehicl car	32	P	Tax & spend tax dividend incom fiscal medicaid
35	P	Cap & trade markey waxman lieberman warner cap	21	P	Urban econ. california ab metropolitan schwarzenegg californian
37	P	Climate adaptation goklani adapt stern mitig resili	7	P	US politics republican sen mccain democrat vote
33	P	Conservation timber eagl fisheri perc graze			

Table 2: *A full list of the estimated topics.* The table provides each topic’s unique ID, descriptive label (in bold), and top 5 stemmed keywords based on the FREX score (Roberts et al. 2014). Further, based on the findings from the topic similarity analysis in Section 5.1, we code whether each topic is related to climate science (S) or climate politics & policy (P).

290 Global temperatures have been flat for approximately 15 years now, even though
291 atmospheric carbon dioxide levels rose more than 40 ppm (or more than 10
292 percent) during that time. Rather than being a harbinger of doom and gloom,
293 the approaching 400 ppm carbon dioxide threshold presents still more evidence
294 that humans are not creating a global warming crisis (Heartland Institute, May
295 17, 2013).

296 The existence of the [Medieval Warm Period] had been recognized in the sci-
297 entific literature for decades. But now it was a major embarrassment to those
298 maintaining that the 20th century warming was truly anomalous. It had to be
299 “gotten rid of” (NCPA, Dec. 6, 2006).

300 Many documents also suggest alternate climate forcing inputs such as the sun
301 or cosmic rays (topic 12) as more plausible explanatory factors for climate fluc-
302 tuations than greenhouse gas emissions. The validity and reliability of empirical
303 data used in climate change studies (topic 13) to demonstrate global warming
304 impacts are cast into doubt. Further, the underlying assumptions of climate
305 change models (topic 11) that are referenced in the IPCC assessments are of
306 “dubious merit” (Fraser, July 7, 2004).

307 The results of the LDA model also demonstrate the breadth of topics dis-
308 cussed in documents referencing climate change with important issue linkages
309 across both the domestic and international political economy. Much critical
310 discussion surrounds international mitigation policies (topic 44) as threats to
311 national sovereignty and expected detrimental impacts to the economy (topic
312 25). Renewable energy technologies such as solar and wind (topic 8) as well
313 as biofuels (topic 18) are almost always presented as inadequate solutions on
314 their own. Fossil fuel production (topic 2), on the other hand, is discussed in
315 positive terms, typically in relation to energy independence and technological
316 innovation. For instance, an expansion of oil drilling into the Arctic National
317 Wildlife Refuge (ANWAR) has been framed as an “important part of a pro-
318 consumer energy policy” that will make energy “plentiful and affordable” (CEI,
319 March 14, 2005). The harmful impacts of regulation in the energy sector, such
320 as GHG emissions reductions (topic 29), automobile fuel standards (topic 47)
321 and cap-and-trade policy (topic 35), are also discussed negatively. For instance:

322 Whether the American economy is booming or heading off a fiscal cliff, the right
323 time for a carbon tax is never (Heritage Foundation, January 8, 2013).

324 [A] carbon tax would raise family energy prices by more than \$500 per year, jack
325 up gasoline prices 50 cents per gallon, reduce family income by nearly \$2,000,
326 and cost 1 million jobs by 2016 alone. Since developing nations like China and
327 India will continue increasing their CO2 no matter what the U.S. does, a carbon
328 tax is a bad solution to a still-unproven problem (CFACT, February 15, 2013).

329 Overall, the Lieberman-Warner bill promises substantial hardship for the econ-
330 omy overall, for jobs, and for energy costs. Given current economic concerns and
331 energy prices, this is the last thing the American people need. At the same time,
332 the environmental benefits would likely be small to nonexistent. The Lieberman-
333 Warner bill fails any reasonable cost-benefit test (Heritage Foundation, May 30,
334 2008).

335 Further, the integrity of climate scientists is also frequently questioned, es-
336 pecially in relation to the peer-review process of the IPCC (topic 27) and other
337 perceived violations of scientific integrity (topic 14) such as the so-called “cli-
338 mategate” email controversy of late 2009 which supposedly has dealt a “death
339 blow” to the global warming “fraud” (Heartland Institute, Nov. 21, 2009). Nu-
340 merous documents take aim at the credibility of climate scientists; the following
341 excerpt serving as a typical example.

342 The purloined letters show a climate-science community in full tribal mode, con-
343 spiring to suppress contrary findings in the peer-reviewed literature; excluding
344 contrary peer-reviewed publications from IPCC reports; concealing the shoddy
345 nature of climate data; colluding to hide data and destroy correspondence; and
346 using mathematical tricks to produce ever more alarming-looking charts (Amer-
347 ican Enterprise Institute, Nov. 25, 2009).

348 These conspiracy-based themes are related to a broader trend within the corpus
349 of equating scientific findings on climate change with “alarmism” (topic 16),
350 where individual scientists and activists are presented as fomenting a state of
351 panic based on inconclusive or even fabricated evidence. Al Gore, for example,
352 has been accused of using “distorted evidence” to further a “scare-them-green
353 agenda” (CEI, March 16, 2007). More generally, “global warming alarmists”,
354 such as climate scientist Michael Mann, are accused of being in the business
355 of “spreading myths and misinformation to further their agenda” (Heartland
356 Institute, June 29, 2012). For example:

357 Mann’s claims that human’s [sic] have caused tremendous warming over the last
358 100 years and that the 1990s were the warmest decade are untenable [...] Looking
359 at the data, the global warming scare appears to be merely ‘Mann made’ junk
360 science (NCPA, July 12, 2004).

361 5. Assessing model quality: reliability and validity

362 It is crucial when coding themes to establish sufficient levels of reliability and
363 validity. Traditionally, difficulties associated with determining reliability have
364 plagued content analytic studies, as a single coder’s judgements may be highly
365 subjective. While subsequent studies have shown that relying on multiple coders
366 and establishing sufficient inter-coder reliability may yield consistent measure-
367 ment in repeated trials, few content analytic studies in the literature on climate
368 scepticism report any reliability estimates. This is understandable given that
369 reproducing measures based on traditional methods is a costly endeavour. On
370 the other hand, this is one area where automated approaches excel—improved
371 reliability is often considered a key benefit of employing a computer-assisted
372 approach (Laver and Garry 2000, Laver et al. 2003). Once the text is collected
373 and the model is programmed, the measuring procedure should yield *exactly* the
374 same results in repeated trials.

375 Although the benefits of employing automated methods for reliability are
376 clear, the same cannot be said for validity and thus the onus is on the researcher

377 to establish the soundness of their results when using computer-assisted ap-
378 proaches. [Grimmer and Stewart \(2013\)](#), in a review of the text analysis litera-
379 ture in political science, argue emphatically for the need to “[v]alidate, validate,
380 validate,” stating “that what should be avoided, then, is the blind use of any
381 method without a validation step” (pg. 5). This section devotes considerable
382 attention to this “validation step,” using multiple methods to examine diverse
383 conceptions of validity. Specifically, we 1) provide further evidence of the *se-*
384 *mantic* validity of our findings, 2) assess *predictive* validity via external events,
385 and 3) examine *concurrent* validity by comparing the model output to a human
386 gold standard.

387 5.1. *Semantic validity and topic similarity*

388 While the descriptive labels described in Section 4.1 offer initial support for
389 semantic validity, an additional means of examining this criterion assesses the
390 extent to which topics relate to one another in substantively meaningful ways
391 ([Quinn et al. 2010](#)). Note that a “topic” in the LDA model is represented by
392 a probability distribution—i.e., the distribution of words given the topic—and
393 thus the notion of “topic similarity” centres on the distance between two proba-
394 bility distributions. While there are a number of metrics available for examining
395 the distance between probability distributions, a common approach is to rely on
396 the well-known Kullback-Leibler (KL) divergence or the related Jensen-Shannon
397 divergence (JSD). We examine similarity (or dissimilarity) using the square root
398 of JSD (sometimes referred to as Jensen-Shannon “distance”), which rescales
399 the JSD into a proper metric ([Endres and Schindelin 2003](#), [Osterreicher and
400 Vajda 2003](#)). Intuitively, when two topic distributions are more similar, they
401 will share a smaller JS distance and vice versa. Figure 1 presents this infor-
402 mation graphically by mapping the pairwise distances onto a two dimensional
403 space using classic multi-dimensional scaling ([Gower 1966](#)). Topics that address
404 similar themes—and thus rely on similar words with high probability—should
405 be relatively close to one another in Figure 1, while dissimilar themes should be
406 further way.

407 The results of this analysis are striking. First, we observe a set of meaningful
408 clusters, with topics related to politics, policy and regulation, energy, climate
409 science, and scientific integrity located in distinct areas of the figure. Moreover,
410 when looking *within* the principal areas, the topics also cluster as expected. For
411 instance, considering the “Policy & Regulation” theme, topics associated with
412 government regulation (15 and 38) inhabit the lower portion of the cluster which
413 is closer to the “Domestic & Int’l Politics” cluster, while the upper area deals
414 with themes more associated with government planning (22, 32, and 33). It is
415 not a surprise that *Tax & Spend* (32), for example, is closer to the “Energy”
416 cluster, as most discussions related to energy policy involve burdensome taxes on
417 fossil fuel consumption. Second, the distance between the four main issue areas
418 fits with intuition. As expected, “Energy”, “Policy & Regulation” and “Do-
419 mestic & Int’l Politics” are quite far away from the “Science” cluster. Perhaps
420 most interesting, however, are the findings associated with scientific integrity.

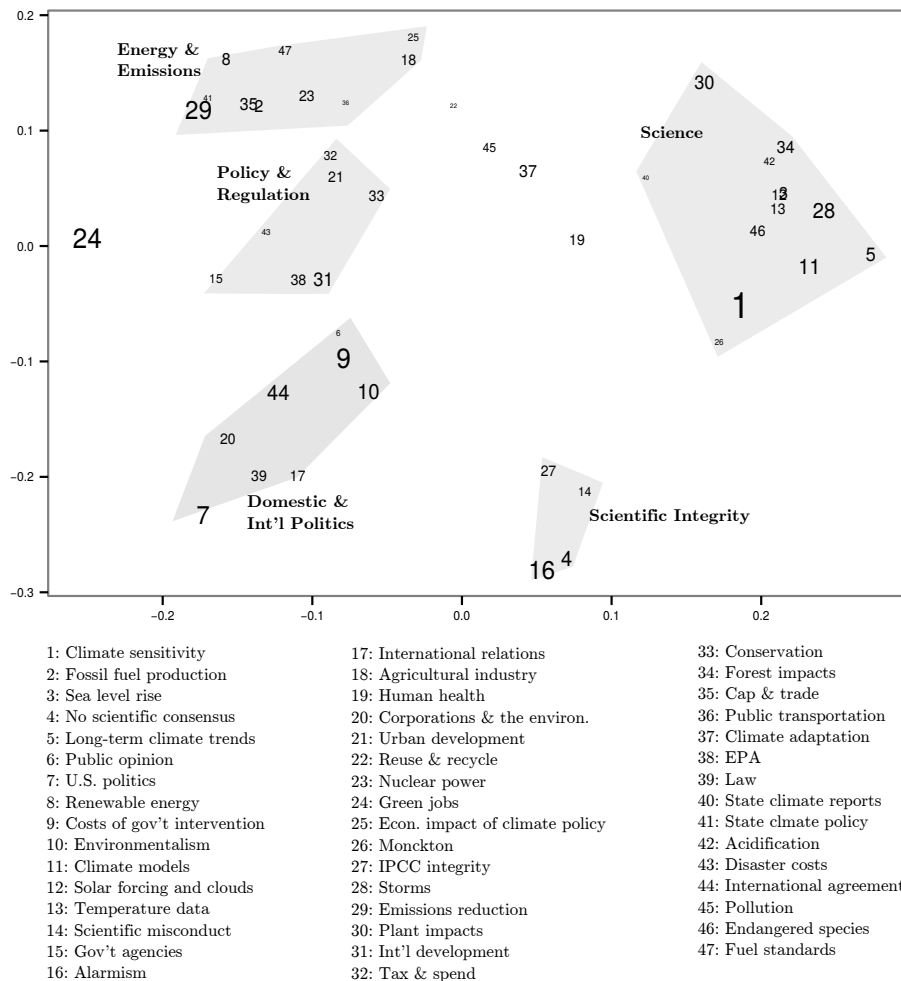


Figure 1: *Topic similarities.* The figure presents Jensen-Shannon distances projected onto a 2D space via multi-dimensional scaling. The size of plotted label corresponds to the number of times the topic was sampled in the corpus and thus gives a rough indication of topic importance. Topics using similar words will be closer together in the figure and vice versa. To ease visualization, we plot the convex hull for each cluster in grey.

421 Not only do topics dealing with scientific misconduct—both regarding scientists
 422 themselves, the scientific consensus on AGW, and the IPCC in general—form
 423 their own distinct cluster, the language used seems to have more in common with
 424 politics than science; that is, scientists are presumed to wield “junk science” to

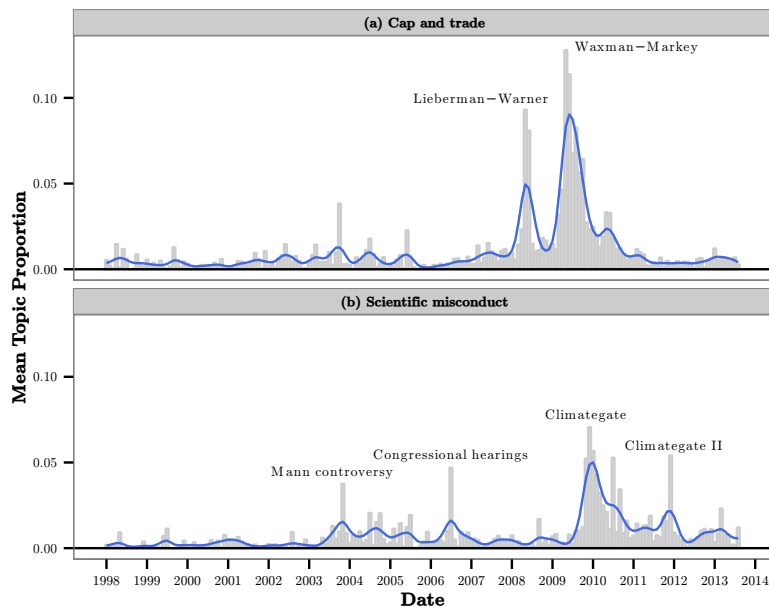


Figure 2: *Predictive validity based on external events.* The graphs illustrate the average monthly topic proportions of four topics over the period January 1998 – August 2013. A local polynomial trend line is included to assist interpretation.

425 achieve political aims. Lastly, a number of topics are at the crossroads between
 426 important issue areas. For example, *Climate adaptation* (37) is located at the
 427 nexus between science and policy, which is not surprising given that adapta-
 428 tion focuses on using climate science to understand the adverse impact of global
 429 warming and implementing policies to prevent or mitigate potential damage.
 430 What is surprising is that a simple model based on word co-occurrences is able
 431 to detect this nuance. Taken together, we find that the 47 topics cluster onto a
 432 smaller set of theoretically meaningful and valid higher-order themes.

433 5.2. Predictive validity and topic dynamics

434 To further assess the quality of our classifications, this section examines
 435 the *predictive validity* of the estimated model—i.e., the extent to which our
 436 topics are predicted by external events (Quinn et al. 2010). However, prior to
 437 examining the relationship between key contrarian claims and external events,
 438 it is necessary to decide on a suitable measure of topic prevalence over time. We
 439 turn to this challenge in the next section.

440 5.2.1. *Measuring topic prevalence over time*

441 There is little agreement in the literature regarding the “best” way to com-
442 bine underlying topic probabilities to produce aggregate level measures and, as
443 with issues of measurement more generally, the appropriateness of an item is
444 often contingent on the research question under consideration. While assumed
445 measures may vary in a number of different ways, the key question for under-
446 standing contrarian claims over time is whether one captures *absolute* or *relative*
447 topic prevalence. An absolute measure allows the “information pie” to grow over
448 time, while its relative counterpart holds the pie constant, instead focusing on
449 the competition among counter-claims within a specified time frame. We rely on
450 two measures—one absolute and the other relative—to formulate the descriptive
451 analysis below. The first (absolute) measure simply sums the topic proportions
452 for a particular topic in a given period of time (e.g. the proportions for the
453 “Alarmism” topic during December 2008), while the second (relative) focuses
454 on the mean topic proportion within a specified time frame. One implicit as-
455 sumption is that each measure gives equal weight to the topic proportions across
456 documents and thus ignores document length. Given the extremely skewed dis-
457 tribution of word lengths in our corpus, however, the proposed measures offer
458 a more stable estimate of topic prevalence and avoid the equally problematic
459 assumption that document importance scales linearly with word length. More-
460 over, estimates using a suitable nonlinear transformation of the word counts
461 (e.g., taking the log) offer virtually identical results in both cases and thus our
462 measurement choice appears robust.

463 5.2.2. *Assessing predictive validity via external events*

464 Figure 2 provides the mean topic proportion for two topics, *Cap & trade* (35)
465 and *Scientific misconduct* (14), for each month over the period from January
466 1998 to August 2013. First, turning to cap-and-trade (see the top panel of
467 Figure 2(a)) two months—May 2008 and August 2009—clearly stand out. The
468 first large peak coincides with the Senate vote on the Lieberman-Warner bill
469 (America’s Climate Security Act of 2007). Significant opposition to the bill found
470 within the corpus largely argues that the legislation would do massive damage
471 to the national economy while offering modest to no environmental benefits.
472 The second significant spike occurs in August 2009, just after House approval of
473 the Waxman-Markey bill (American Clean Energy and Security Act of 2009).
474 Similar types of arguments that were used against the Lieberman-Warner bill
475 also surfaced during the Waxman-Markey period. Following the defeat of the
476 Waxman-Markey bill, we see a sharp decline in discussions surrounding emissions
477 reduction legislation. However, a resurgence of the topic occurs in 2013, with
478 much attention being placed on the dangers of a carbon tax for the economy.

479 Figure 2(b) displays the share of words dealing with a scientific misconduct
480 theme. A sustained period of interest seems to cover the 2003-2005 period,
481 with the release of papers from climate sceptics such as Stephen McIntyre, Ross
482 McKittrick, and Hans von Storch, which criticize Michael Mann’s methodology.
483 The next substantial increase in the topic proportion is observed in July 2006,

484 when Congressional hearings were held on the validity of Mann and colleagues’
485 findings. However, a real break in the series occurs in November-December
486 2009. This is expected since this period coincides with the time when emails of
487 researchers from the Climatic Research Unit (CRU) at the University of East
488 Anglia were hacked, uploaded to the Internet, and subsequently scrutinized by
489 climate sceptics. Following this flurry of attention to scientific integrity during
490 late 2009 and early 2010, a downward trend then follows with significant peaks
491 occurring in July 2010 when the Independent Climate Change Email Review
492 was released and December 2011 which was just after a second round of CRU
493 emails were uploaded to the Internet; an incident named “climategate II” by
494 climate sceptics.

495 Overall, the evidence in Figure 2 suggests that the data produced by the
496 model vary in predictable ways based on closely related external events and, as
497 such, exhibit adequate levels of predictive validity. Moreover, in the interest of
498 space, we limited our discussion to two key topics in the area of climate policy
499 and science. However, many other topics—such as extreme weather, interna-
500 tional negotiations, and energy policy—display similar patterns of predictive
501 validity.

502 5.3. *Assessing concurrent validity via a human “gold standard”*

503 As a last look at validity, we compare the model’s classifications to those of
504 two human coders using a random sample of 300 manually annotated documents.
505 After ensuring a suitable level of inter-coder reliability (Krippendorff’s $\alpha = 0.74$),
506 the coders classified the primary topic or theme of each article using either the
507 47 categories provided in Table 2 or “other” if none of the model-based topics
508 suitably captured the main theme.² Based on these data, the micro-averaged
509 precision and recall for classifying the primary topic are 0.64 and 0.65, respec-
510 tively. These figures are encouraging, as coding a document into 47 categories
511 is a difficult classification task and the model performs considerably better than
512 rolling a 47 sided die or simply choosing the modal value. More importantly
513 for the analysis below, aggregating the topics to produce more general themes
514 or classes greatly improves each measure of performance. When aggregating all
515 the way up to the science label used in Section 6, the precision and recall are
516 0.94 and 0.96, respectively; for the policy label, the precision and recall are 0.94
517 and 0.92, respectively..

518 It is also important to note that assessing a topic model using only the
519 primary topic offers a conservative estimate of performance. Several distinct
520 themes often contribute to a document’s composition and deciding which is

²The coders consisted of one author and a research assistant. In the pilot phase, to get a general sense of the coding task, each coder carried out an initial coding of 10 randomly selected documents, which was followed by an in-depth discussion of coding choices. Following this initial round, the coders went on to code an additional 30 documents and the discussion was repeated. Finally, the coders went through a random sample of 50 documents—this is the sample used to calculate inter-coder reliability.

521 “primary” is often quite difficult for both human and machine. Indeed, allowing
522 documents to be composed of multiple topics—an appropriate assumption for
523 the vast majority of texts in our corpus—is one of the major advantages of using
524 the LDA. Notably, the proportion of documents correctly classified jumps to
525 0.78 if one considers the first two most probable topics based on the model.

526 6. Policy versus science: Is the era of science denial over?

527 In 2013, the World Wildlife Fund-UK’s chief advisor on climate change, Leo
528 Hickman, stated in no uncertain terms that “[t]he real world is leaving behind
529 those who flatly reject the science underpinning the notion that anthropogenic
530 greenhouse gas emissions are warming the planet,” arguing that climate science
531 sceptics are being replaced by “climate policy sceptics.” More recently, in July
532 2015, Elliott Negin from the Union of Concerned Scientists pointed to a more
533 modest retreat: “[deniers] now concede that climate change is real, but reject the
534 scientific consensus that human activity—mainly burning fossil fuels—is driving
535 it.” These arguments are not new. Speculation regarding the decline of scientific
536 scepticism is seen as early as 2002, just two years after McCright and Dunlap’s
537 seminal study. In a leaked memo to the Republican party, conservative strategist
538 Frank Luntz suggests:

539 *The scientific debate remains open. Voters believe that there is no*
540 *consensus about global warming within the scientific community.*
541 *Should the public come to believe that the scientific issues are settled,*
542 *their views about global warming will change accordingly. Therefore,*
543 *you need to continue to make the lack of scientific certainty a pri-*
544 *mary issue in the debate, and defer to scientists and other experts*
545 *in the field [...] The scientific debate is closing [against us] but not*
546 *yet closed. There is still a window of opportunity to challenge the*
547 *science.*³

548 If indeed the window of opportunity for scientific scepticism has closed, this
549 would be a welcome development for proponents of climate action. After all, a
550 general acceptance of anthropogenic global warming is a necessary condition for a
551 comprehensive agreement on climate change mitigation and there is considerable
552 evidence to suggest that acknowledging the scientific consensus on AGW predicts
553 support for climate policy (Ding et al. 2011, McCright et al. 2013, van der Linden
554 et al. 2015). However, based on existing evidence in the literature, it is difficult
555 (if not impossible) to discern whether the era of climate science denial is truly
556 over or if the organised denial of “junk” science remains alive and well.

557 To examine this question, we present evidence on the evolution of the CTT
558 science- and policy-related discourse since the late 1990s. Figure 3(a) presents

³Italics are in original. The full text of the environmental policy section of the Luntz memo can be accessed at https://www.motherjones.com/files/LuntzResearch_environment.pdf.

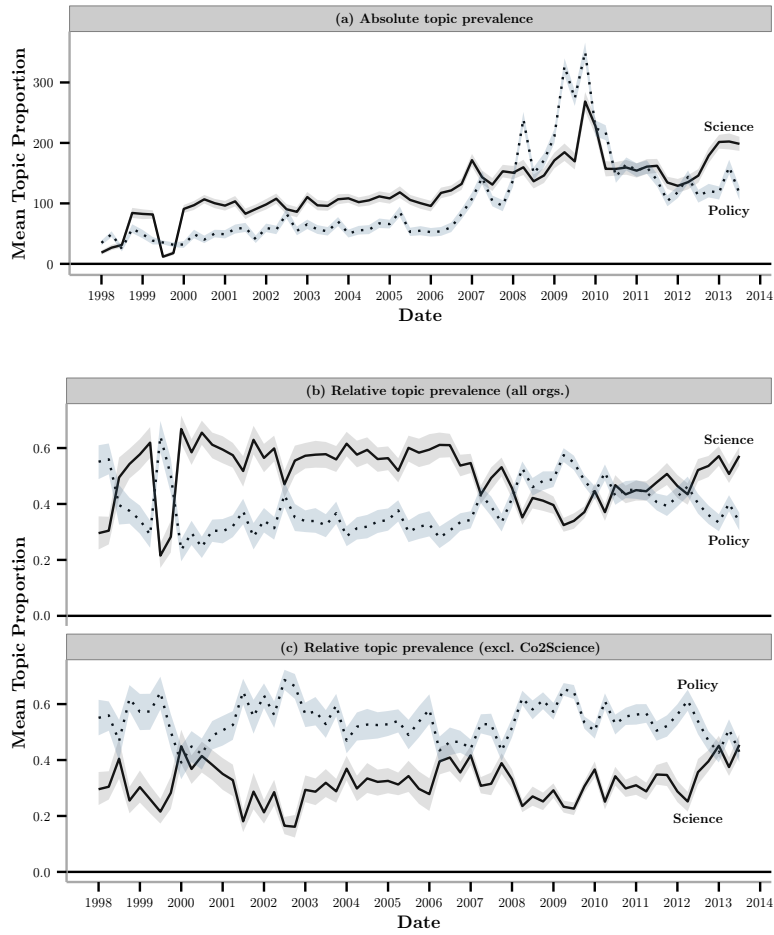


Figure 3: *The evolution of political and science-related discourse.* Panel (a) displays the summed quarterly topic probability of “science” (solid) and “politics & policy” (dotted) related themes for all CTTs in the sample over the period January 1998 – August 2013. These categories are aggregations of the topics based on the codings displayed in Table 2. The bottom panel shows the average quarterly topic probabilities—a relative measure—for the same categories; (b) uses all available data, while (c) excludes Co2Science. The areas around each series represent the bootstrapped 95% confidence interval.

559 the sum of the topic proportions for “science” and “politics & policy” related
560 topics for each quarter over the Q1/1998–Q3/2013 period (absolute measure),
561 while Figures 3(b) and (c) provide mean topic probabilities (relative measure).
562 Each time series also includes an estimate of uncertainty, as measured by a
563 bootstrapped 95% confidence interval.⁴ These categories are aggregations of
564 topics following the codings presented in Table 2. Several aspects of Figure 3 are
565 noteworthy. First, in absolute terms, the intensity of discussion—regardless of
566 whether the focus is on “science” or “politics & policy”—has grown considerably
567 since McCright and Dunlap (2000). Consistent with broader trends in media
568 coverage of climate change, (e.g. Schmidt et al. 2013), the discussion increases
569 until around the time of the Copenhagen conference and the so-called climategate
570 scandal (late 2009–early 2010), and then declines thereafter. Moreover, these
571 data suggest that science-related discussions have been dominant since 2012.
572 We thus find little evidence for the “end of science denial” and yet a rise in
573 “policy sceptics” remains consistent with the data.

574 Second, as demonstrated in Figure 3(b), recent years are marked by a di-
575 vergence between the science and policy series: the relative emphasis on science
576 seems to be gaining in the post-“climategate” era. Nevertheless, this result is
577 largely driven by the influence of one prolific science-oriented CTT, Co2Science,
578 which produces a steady stream of scientific review articles (see Table 1). When
579 excluding this organization, as shown in Figure 3(c), we see that policy-related
580 discussion is frequent, there has been convergence between the frequency of
581 policy and science discussion at key periods, and that aggregate discussions of
582 science appear to be on the rise after 2012.

583 However, aggregating across diverse science and political themes, as shown
584 in Figure 3, masks important heterogeneity in sceptical discourse. Some or-
585 ganizations focus almost entirely on producing science-oriented content (e.g.,
586 Co2Science), others are dedicated to addressing issues surrounding climate pol-
587 icy (e.g., the Heritage Foundation), and still others focus on a range of both
588 science and policy related topics. In the later category, the Heartland Insti-
589 tute stands out as an important counter-movement organisation worthy of a
590 closer look. As proudly trumpeted on its website, Heartland has been described
591 by mainstream news sources as “the world’s most prominent think tank pro-
592 moting scepticism about man-made climate change” (The Economist) and “the
593 primary American organization pushing climate change scepticism” (The New
594 York Times). These “accolades” are not by chance. Judging from our data (see
595 Table 1), it is clear that Heartland has been a front-runner in CTT literature
596 production and has been a leader in public outreach. Indeed, Heartland has been
597 recognized by scholars as a significant contrarian actor and has been prominently
598 studied in past literature on organised climate scepticism (McCright and Dunlap

⁴Note that to remain as consistent as possible with the assumed data generating process, we conducted the bootstrap at the *document* level for each time period of interest in the sample. Specifically, for a given quarter, we sample (with replacement) from the available documents and calculate topic prevalence, repeating this process for 1,000 replicates for each series.

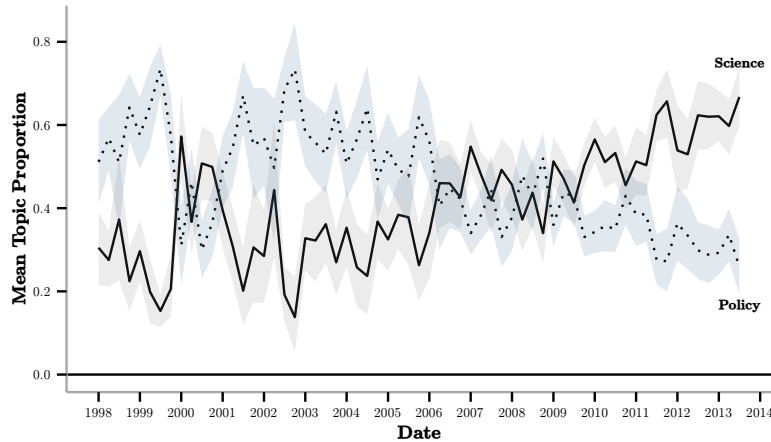


Figure 4: *The Heartland Institute’s political and science-related discourse.* Displays average quarterly topic probabilities for science- and policy-related themes in documents disseminated by Heartland over the period January 1998–August 2013.

599 [2003, Cann 2015](#)).

600 How then, does its discourse on “science” and “politics & policy” related
 601 themes compare to the general trend illustrated in Figure 3? We narrow our
 602 focus on Heartland in Figure 4, which shows how beginning in 2002, we can
 603 observe a steady rise in an emphasis on topics related to science, as well as an
 604 attendant decline in policy-oriented themes. Interestingly, Heartland’s shift to-
 605 wards science-related themes preceded “climategate” by more than 7 years and
 606 actually dovetails with Luntz’s famous “Straight Talk” memo. It is therefore not
 607 a surprise that for a decade it has organized the annual International Conference
 608 on Climate Change (also known as Denial-a-Palooza) which serves as a forum
 609 for climate science deniers,⁵ or that it made headlines in 2012 after launching a
 610 controversial ad campaign which equated climate scientists with Ted Kaczynski
 611 (the Unabomber). The consistent trade-off of attention from policy to science
 612 since 2002 suggests that Heartland has invested heavily in attempting to re-open
 613 the “window of science scepticism.”

614 Another potential source of heterogeneity relates to our categorizations of
 615 science and policy related discussions. It is clear that some topics labelled as
 616 “policy” are only tangentially related to “climate” policy and that there are im-
 617 portant differences between climate science and scientific integrity. We therefore
 618 examine three themes which are directly related to climate science and policy:
 619 “Science,” “Scientific Integrity,” and “Energy and Emissions Policy.” Figure 5

⁵<http://www.desmogblog.com/directory/vocabulary/2782>

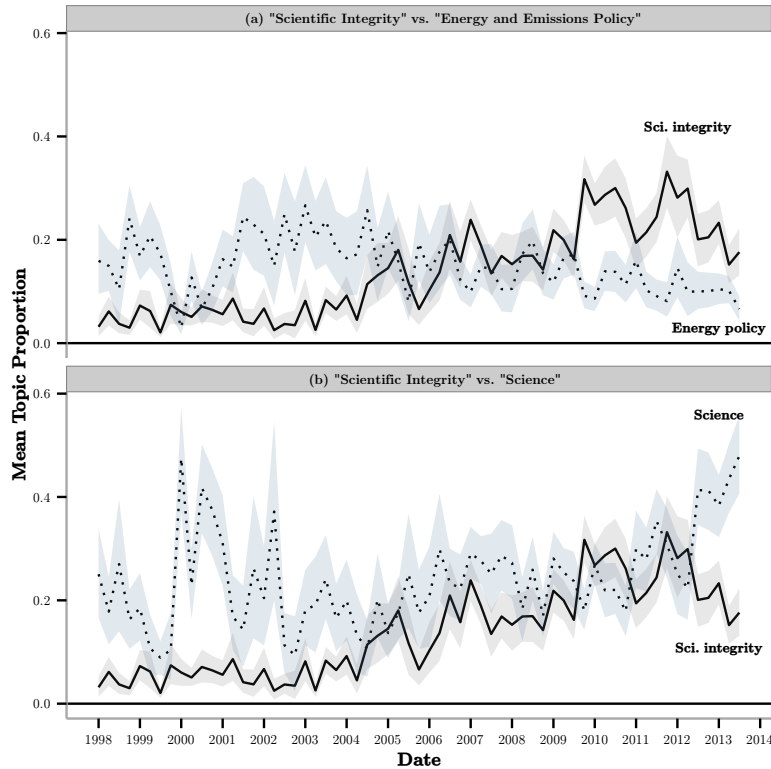


Figure 5: *Climate-specific related themes over time.* The figures show the average quarterly topic proportions of three topic clusters, which are directly related to climate science and policy, as classified in Section 5.1: “Science,” “Scientific Integrity,” and “Energy and Emissions Policy.” Note that Co2Science has been excluded from this analysis. The series covers the period Q1/1998–Q3/2013.

620 provides the results of this comparison. Several features of this figure are notable.
 621 First, considering the “Scientific Integrity” series, there has been an appreciable
 622 rise in the prevalence of integrity-related topics starting in 2004 and peaking in
 623 2011. Second, talk of scientific integrity began to overtake that of energy policy
 624 during 2006 and 2007—which corresponds to a period dominated by *An Inconve-*
 625 *nient Truth* and Al Gore’s acceptance of the Nobel Peace Prize—and proceeded
 626 to become relatively more prevalent in the post-“climategate” era (Figure 5 (a)).
 627 Lastly, while the discussion of climate “Science” was more frequent relative to
 628 “Scientific Integrity” from 1998 to roughly 2004, the two series become inter-
 629 twined for much of the sample period. This suggests that CTTs were just as
 630 likely to question the integrity of individual scientists and scientific bodies than
 631 to discuss alternative scientific viewpoints; though, there has been a percepti-

632 ble break since 2012, with discussions of “Science” once again dominating the
633 conversation.

634 7. Conclusion

635 Despite urgent calls to action among climate scientists, the U.S. government
636 continues to avoid comprehensive climate policy action and the American public
637 remains misinformed on key aspects of the debate. A growing literature draws at-
638 tention to the influence of a well-organized and well-funded movement of climate
639 sceptics. This study provided the first systematic update of the claims making
640 activity of conservative think tanks—a critical piece of the climate counter-
641 movement—since the influential work of [McCright and Dunlap \(2000\)](#). Our key
642 findings include:

- 643 1. The overall level of CTT claims-making has grown rapidly over the past
644 decade and a half, reaching a peak during late 2009–early 2010;
- 645 2. The 19 CTTs studied address a wide range of topics in their written com-
646 munication since [McCright and Dunlap \(2000\)](#), which cluster into distinct
647 themes associated with politics, policy, science, and scientific integrity;
- 648 3. Topics questioning the integrity of individual scientists and scientific bodies
649 appear closer (semantically) to politics than science, suggesting that claims
650 often considered the hallmark of scientific scepticism are rooted in politics;
- 651 4. The era of climate science denial is not over. While the aggregate re-
652 sults demonstrate that both policy and science discussions remain stable
653 throughout the period of study (Figure 3), a detailed analysis of a criti-
654 cal CTT (Figure 4) and a focus on climate change-specific themes (Figure
655 5) reveal the increased importance of both science and scientific integrity
656 discussions over the sample period.
- 657 5. CTTs tend to react to the external environment—i.e., they *counter* claims—
658 and thus studies focusing on narrow intervals of time (or a single organi-
659 sation) are likely sensitive to these contextual factors.

660 It is important to note, however, that the current study has a number of lim-
661 itations. First, we are necessarily restricted to the documents that are publicly
662 available online. It should be noted, however, that these organisations have an
663 incentive to distribute what they produce, which could support validity, but this
664 tendency may be weaker for documents produced further back in time. Second,
665 we do not transcribe video and audio data, which may be included in future
666 work. Third, and more importantly, we do not perform any sentiment analysis
667 on the corpus. For instance, if a document focuses on the Medieval Warm Pe-
668 riod (topic 37), we are assuming that its argument is that natural forces have
669 a stronger climate impact than human activity. Based on our reading of the
670 corpus, as well as our theoretical priors, this is a plausible assumption. Despite
671 these limitations, in providing this corpus to the community, we hope to offer a
672 platform for future work on the claims-making activity of CTTs.

673 **References**

- 674 AlSumait, L., Barbara, D., Gentle, J., Domeniconi, C., 2009. Topic significance
675 ranking of lda generative models. In: ECML.
- 676 Anderegg, W., Prall, J., Harold, J., Schneider, S., 2010. Expert credibility in
677 climate change. Proceedings of the National Academy of Sciences of the United
678 States of America 107, 12107–9.
- 679 Arun, R., Suresh, V., Madhavan, C. V., Murthy, M. N., 2010. On finding the
680 natural number of topics with latent dirichlet allocation: Some observations.
681 In: Advances in Knowledge Discovery and Data Mining. Springer, pp. 391–
682 402.
- 683 Austin, A., 2002. Advancing accumulation and managing its discontents: The us
684 antienvironmental countermovement. Sociological Spectrum 22 (1), 71–105.
- 685 Beder, S., 2001. Research note-neoliberal think tanks and free market environ-
686 mentalism. Environmental Politics.
- 687 Begley, S., Conant, E., Stein, S., Clift, E., Philips, M., 2007. The truth about
688 denial. Newsweek 150 (7), 20–7.
- 689 Blei, D. M., 2012. Probabilistic topic models. Communications of the ACM 55 (4),
690 77–84.
- 691 Blei, D. M., Ng, A. Y., Jordan, M. I., 2003. Latent dirichlet allocation. Journal
692 of Machine Learning Research 3, 993–1022.
- 693 Boykoff, M., 2008. Lost in translation? unites states television news coverage of
694 anthropogenic climate change, 1995-2004. Climatic Change 86, 1–11.
- 695 Boykoff, M., Boykoff, J., 2004. Balance as bias: global warming and the us
696 prestige press. Global Environmental Change 14, 125–136.
- 697 Brulle, R. J., 2014. Institutionalizing delay: foundation funding and the creation
698 of u.s. climate change counter-movement organizations. Climatic Change 122,
699 681–694.
- 700 Brulle, R. J., Carmichael, J., Jenkins, J. C., 2012. Shifting public opinion on
701 climate change: an empirical assessment of factors influencing concern over
702 climate change in the us, 2002–2010. Climatic change 114 (2), 169–188.
- 703 Cann, H. W., 2015. Climate change, still challenged: Conservative think tanks
704 and skeptic frames.
- 705 Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., Blei, D. M., 2009. Reading
706 tea leaves: How humans interpret topic models. In: NIPS.

- 707 Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R.,
708 Way, R., Jacobs, P., Skuce, A., 2013. Quantifying the consensus on anthro-
709 pogenic global warming in the scientific literature. *Environmental Research*
710 *Letters* 8, 024024.
- 711 Ding, D., Maibach, E. W., Zhao, X., Roser-Renouf, C., Leiserowitz, A., 2011.
712 Support for climate policy and societal action are linked to perceptions about
713 scientific agreement. *Nature Climate Change* 1 (9), 462–466.
- 714 Doran, P., Zimmerman, M., 2009. Examining the scientific consensus on climate
715 change. *Eos, Transactions American Geophysical Union* 90 (3), 22–23.
- 716 Dunlap, R. E., Jacques, P. J., 2013. Climate change denial books and conser-
717 vative think tanks: exploring the connection. *American Behavioral Scientist*,
718 0002764213477096.
- 719 Dunlap, R. E., McCright, A. M., 2011. Organized climate change denial. *The*
720 *Oxford handbook of climate change and society*, 144–160.
- 721 Elgesem, D., Steskal, L., Diakopoulos, N., 2015. Structure and content of the dis-
722 course on climate change in the blogosphere: The big picture. *Environmental*
723 *Communication* 9 (2), 169–188.
- 724 Elsasser, S. W., Dunlap, R. E., 2013. Leading voices in the denier choir: Con-
725 servativecolumnists dismissalof global warming and denigration of climate sci-
726 ence. *American Beh* 57 (6), 754–776.
- 727 Endres, D. M., Schindelin, J. E., 2003. A new metric for probability distributions.
728 *IEEE Transactions on Information Theory* 49 (7), 1858–1860.
- 729 Feldman, L., Maibach, E., Roser-Renouf, C., Leiserowitz, A., 2012. Climate on
730 cable: The nature and impact of global warming coverage on fox news, cnn,
731 and msnbc. *The International Journal of Press/Politics* 17 (1), 3–31.
- 732 Freudenburg, W. R., 2000. Social constructions and social constrictions: Toward
733 analyzing the social construction of the naturalizedas well as the natural..
734 *Environment and global modernity*, 103–119.
- 735 Gower, J. C., 1966. Some distance properties of latent root and vector methods
736 used in multivariate analysis. *Biometrika* 53, 325328.
- 737 Greenpeace, 2010. *Dealing in Doubt: The Climate Denial Industry and Climate*
738 *Science: A Brief History of Attacks on Climate Science, Climate Scientists*
739 *and the IPCC*. Amsterdam: Greenpeace International.
- 740 Griffiths, T. L., Steyvers, M., 2004. Finding scientific topics. *Proceedings of the*
741 *National Academy of Sciences of the United States of America* 101, 5228–5235.

- 742 Grimmer, J., King, G., 2011. General purpose computer-assisted clustering and
743 conceptualization. *Proceedings of the National Academy of Science Inagural*
744 *Articles*, 1–8.
- 745 Grimmer, J., Stewart, B. M., 2013. Text as data: The promise and pitfalls
746 of automatic content analysis methods for political texts. *Political Analysis*
747 *21 (3)*, 267–297.
- 748 Hart, P., 2008. Market influences on climate change frames in cnn and fox
749 news channel broadcasts. In: *International Communication Association An-*
750 *ual Meeting*.
- 751 Hoffman, A. J., 2011. Talking past each other? cultural framing of skeptical and
752 convinced logics in the climate change debate. *Organization & Environment*
753 *24 (1)*, 3–33.
- 754 Holliman, R., 2011. Advocacy in the tail: Exploring the implications of climate-
755 gatefor science journalism and public debate in the digital age. *Journalism*
756 *12 (7)*, 832–846.
- 757 IPCC, 2014. *Climate Change 2014: Synthesis Report. Contribution of Working*
758 *Groups I, II and III to the Fifth Assessment Report of the Intergovernmental*
759 *Panel on Climate Change*. IPCC, Geneva, Switzerland.
- 760 Jacques, P., Dunlap, R., Freeman, M., 2008. The organisation of denial: Con-
761 servative think tanks and environmental scepticism. *Environmental Politics*
762 *17 (3)*, 349–385.
- 763 Knight, G., Greenberg, J., 2011. Talk of the enemy: adversarial framing and
764 climate change discourse. *Social Movement Studies* *10 (4)*, 323–340.
- 765 Lahsen, M., 2008. Experiences of modernity in the greenhouse: A cultural anal-
766 ysis of a physicist trio supporting the backlash against global warming. *Global*
767 *environmental change* *18 (1)*, 204–219.
- 768 Lakoff, G., 2014. *The All New Don't Think of an Elephant!: Know Your Values*
769 *and Frame the Debate*. Chelsea Green Publishing.
- 770 Laver, M., Benoit, K., Garry, J., 2003. Extracting policy positions from political
771 text using words as data. *American Political Science Review* *97 (3)*, 311–331.
- 772 Laver, M., Garry, J., 2000. Estimating policy positions from political texts.
773 *American Journal of Political Science* *44 (3)*, 619–634.
- 774 Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinburg, G., Rosenthal, S.,
775 2015. *Climate change in the american mind: March 2015*. Yale University and
776 George Mason University. New Haven, CT: Yale Project on Climate Change
777 Communication.

- 778 Mann, M. E., 2013. The hockey stick and the climate wars: Dispatches from the
779 front lines. Columbia University Press.
- 780 McCright, A., Dunlap, R., 2000. Challenging global warming as a social problem:
781 An analysis of the conservative movement's counter-claims. *Social Problems*
782 47 (4), 499–522.
- 783 McCright, A., Dunlap, R., 2003. Defeating kyoto: The conservative movement's
784 impact on u.s. climate change policy. *Social Problems* 50 (3), 348–373.
- 785 McCright, A. M., Dunlap, R. E., Xiao, C., 2013. Perceived scientific agreement
786 and support for government action on climate change in the usa. *Climatic*
787 *Change* 119 (2), 511–518.
- 788 National Research Council, 2010. Advancing the Science of Climate Change.
789 National Academies Press.
- 790 Nisbet, M. C., 2009. Communicating climate change: Why frames matter for
791 public engagement. *Environment: Science and Policy for Sustainable Devel-*
792 *opment* 51 (2), 12–23.
- 793 Oreskes, N., 2004. The scientific consensus on climate change. *Science* 306, 1686.
- 794 Oreskes, N., Conway, E., 2010. *Merchants of Doubt; How a Handful of Scientists*
795 *Obscured the Truth on Issues from Tobacco Smoke to Global Warming.* New
796 York: Bloomsbury.
- 797 Osterreicher, F., Vajda, I., 2003. A new class of metric divergences on probability
798 spaces and its applicability in statistics. *Annals of the Institute of Statistical*
799 *Mathematics* 55 (3), 639–653.
- 800 O'Neill, S., Boykoff, M., 2011. The role of new media in engaging the public with
801 climate change. *Engaging the public with climate change: behaviour change*
802 *and communication*, 233–251.
- 803 Painter, J., Ashe, T., 2012. Cross-national comparison of the presence of cli-
804 mate skepticism in the print media in six countries, 2007-10. *Environmental*
805 *Research Letters* 7, 1–8.
- 806 Pooley, E., 2010. *The climate war: true believers, power brokers, and the fight*
807 *to save the earth.* Hachette UK.
- 808 Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., Radev, D. R.,
809 2010. How to analyze political attention with minimal assumptions and costs.
810 *American Journal of Political Science* 54 (1).
- 811 Roberts, M. E., Stewart, BrandStewart, B. D., Lucas, C., Leder-Luis, J., Gadari-
812 an, S. K., Albertson, B., Rand, D. G., 2014. Structural topic models for
813 open-ended survey responses. *American Journal of Political Science* 58 (4),
814 10641082.

- 815 Scheufele, D. A., Tewksbury, D., 2007. Framing, agenda setting, and priming:
816 The evolution of three media effects models. *Journal of communication* 57 (1),
817 9–20.
- 818 Schmidt, A., Ivanova, A., Schäfer, M. S., 2013. Media attention for climate
819 change around the world: A comparative analysis of newspaper coverage in
820 27 countries. *Global Environmental Change* 23 (5), 1233–1248.
- 821 Sharman, A., 2014. Mapping the climate sceptical blogosphere. *Global Environ-*
822 *mental Change* 26, 159–170.
- 823 Union of Concerned Scientists, 2007. *Smoke, Mirrors & Hot Air: How Exxon-*
824 *Mobil Uses Big Tobacco’s Tactics to Manufacture Uncertainty on Climate*
825 *Science*. Union of Concerned Scientists.
- 826 van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D., Maibach, E. W., 2015.
827 The scientific consensus on climate change as a gateway belief: Experimental
828 evidence. *PLoS one* 10 (2), e0118489.
- 829 Wallach, H. M., Mimno, D., McCallum, A., 2009a. Rethinking lda: Why priors
830 matter. *NIPS*.
- 831 Wallach, H. M., Murray, I., Salakhutdinov, R., Mimno, D., 2009b. Evaluation
832 methods for topic models. In: *ICML*.
- 833 Washington, H., Cook, J., 2011. *Climate Change Denial: Heads in the Sand*.
834 Earthscan, 711 Third Avenue, New York, NY, 10017.
- 835 Weber, E. U., Stern, P. C., 2011. Public understanding of climate change in the
836 united states. *American Psychologist* 66 (4), 315.
- 837 Yao, L., Mimno, D., McCallum, A., 2009. Streaming inference for latent dirichlet
838 allocation. *kdd*, 2009. *KDD*.
- 839 Young, N., 2013. Working the fringes: the role of letters to the editor in advanc-
840 ing non-standard media narratives about climate change. *Public Understand-*
841 *ing of Science* 22 (4), 443–459.