Text-Mining the Signals of Climate Change Doubt

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.

Keywords: climate change, scepticism, text classification, latent Dirichlet allocation

1. Introduction

Climate scientists overwhelmingly agree that the Earth is getting warmer and that the rise in average global temperature is predominantly due to human activity (IPCC 2014, National Research Council 2010, Oreskes 2004, Doran and Zimmerman 2009, Anderegg et al. 2010, Cook et al. 2013). Yet a sizeable segment of the American public rejects this "consensus view" (Weber and Stern 2011) and U.S. climate policy remains in a state of limbo. As of early 2015, one-third of the American public believes that climate change is *not* primarily caused by human activity and only one in ten understands that more than 90% of climate scientists agree on the existence and nature of observed global warming (Leiserowitz et al. 2015). What explains this divergence in views among climate scientists and the American public? What factors promote inaction on comprehensive climate mitigation policy? These questions have garnered considerable attention in disciplines across the social and behavioural sciences.

One prominent explanation investigates the influence of a "well-funded and relatively coordinated 'denial machine'" on shaping the public's understanding of climate science (Begley et al. 2007). While a diverse set of actors promote climate scepticism, conservative think tanks (CTTs) play a central role, providing key counter-claims to challenge climate science and obstructing climate policy (McCright and Dunlap 2000). CTTs provide a multitude of services to the cause of climate change scepticism: providing material support and lending credibility to contrarian scientists, sponsoring pseudo-scientific climate change conferences, directly communicating contrarian viewpoints to politicians, and, more generally, disseminating sceptic viewpoints through a range of media to the wider public (Dunlap and McCright 2011). A number of studies also suggest that these organizations are central in obstructing national climate policy (Lahsen 2008, Oreskes and Conway 2010) and international climate change mitigation agreements (McCright and Dunlap 2003). The prominence of CTTs in the contrarian counter-movement has prompted calls for an expansion and improvement of data collection efforts on a range of climate movement and counter-movement activities (Brulle et al. 2012).

Despite an active interest in CTTs, few studies have systematically analysed the nature and prevalence of contrarian counter-claims. Aaron McCright and Riley Dunlap's influential study offers a notable exception, providing a comprehensive survey of CTT counter-claims from 14 major conservative think tanks over the period 1990-1997. Yet, to our knowledge, there have been no systematic updates to this study over the past 15 years and thus little is known about how contrarian claims have evolved over the last decade. We seek to fill this gap in the literature by 1) compiling 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) introducing a methodology to measure key themes in the corpus which scales to the exponential increase in content generated by conservative think tanks (CTTs) over the past decade; and 3) leveraging this new methodology to examine the dynamics of policy- and science-related claims over a 16 year period. We argue that understanding CTT counter-claims is of both theoretical and practical significance, as an acceptance of the anthropogenic causes of climate change is arguably a necessary condition for progress on reaching a climate agreement and may portend a window for policy action.

2. Understanding contrarian counter-claims

A number of scholars argue that the entrenchment of climate change scepticism in American society is not an "accident." Rather, the dismal state of public understanding of AGW in the United States is largely the result of an orchestrated attack on climate science and individual climate scientists by a constellation of interests that are determined to obstruct policies aimed at mitigating global warming (Pooley 2010, Oreskes and Conway 2010, Washington and Cook 2011, Mann 2013). For over twenty years, the American public has been subject to waves of information produced by a "well-coordinated, well-funded campaign by contrarian scientists, free-market think tanks and industry" which has "created a paralyzing fog of doubt around climate change" (Begley et al. 2007). Employing tactics (and even participants) from similar disinformation campaigns, such as those against the regulation of tobacco and ozone-harming chlorofluorocarbons (CFCs), the counter-movement aims to block climate policy by "manufacturing doubt" about the credibility of individual scientists, misrepresenting peer-reviewed scientific findings, and exaggerating scientific uncertainties (Union of Concerned Scientists 2007, Oreskes and Conway 2010, Greenpeace 2010, Dunlap and McCright 2011).

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

Studies interested in measuring the prevalence of contrarian claims focus almost exclusively on the *level* of contrarian information present in media coverage of global warming. These studies have yielded important insights into the prevalence of skepticism within newspapers (e.g., Boykoff and Boykoff 2004, Painter and Ashe 2012, Schmidt et al. 2013), opinion pieces in print media (Hoffman 2011, Elsasser and Dunlap 2013, Young 2013), television (Boykoff 2008, Hart 2008, Feldman et al. 2012), and "new media" (ONeill and Boykoff 2011, Holliman 2011, Knight and Greenberg 2011, Sharman 2014, Elgesem et al. 2015). However, few studies systematically analyse the *content* of contrarian claims and even fewer focus specifically on CTTs. To date, McCright and Dunlap (2000) offers the most comprehensive survey of CTT counter-claims on climate change. The authors content analyse a sample of 224 documents related to global warming from 14 major conservative think tanks over the period 1990-1997, with the vast majority of this literature being produced during 1996 and 1997. Overall, the analysis suggests that climate scepticism during this period centred on three major counter-claims: 1) the evidentiary basis of global warming is weak or wrong, 2) global warming would be beneficial if it was to occur, and 3) global warming policies would do more harm than good (see McCright and Dunlap 2000 pg. 510, Table 3). For the 1990-1997 period, the study finds that 71% of the documents contained criticisms of the scientific evidence for global warming (Counter-claim 1), only 13.4% discussed the benefits of global warming (Counter-claim 2), and 62.1% provided a discussion on the downsides of climate policy action (Counter-claim 3).

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

3. Measuring contrarian claims

3.1. The corpus

To systematically gauge claims-making activity, we retrieved information related to climate change from the websites of 19 well-known North American conservative think tanks and organizations (see online appendix for details). Our choice of organizations, to a large extent, mirrors that of McCright and Dunlap (2000) and the most heavily funded organizations which are identified in Brulle (2014). For each organization, we visited all pages including the terms "climate change" or "global warming" and extracted relevant text and key meta data. There were also instances where pages included links to documents in PDF format, which were typically relatively long policy reports. These PDFs were automatically retrieved, passed through optical character recognition (OCR)

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 (NCPPR)	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.

software to extract the text, and appended to the list of text retrieved from the HTML code. Audiovisual materials were a minority of the overall set of retrieved pages and were excluded in the current analysis. This process produced more than 16,000 documents over the period from 1998 to 2013.

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

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

3.2. Methods: probabilistic topic modelling

The time and effort associated with reading over 16,000 documents renders traditional content analytic approaches inadequate and/or infeasible and thus the next step is to find a suitable computational model to help make sense of the data. We approach this step using an *unsupervised* approach, exploring the presence of meaningful clusters of terms that appear across documents in the collected corpus. While there is no shortage of clustering algorithms in the literature (Grimmer and King 2011), we utilize the latent Dirichlet allocation (LDA) model originally proposed in Blei et al. (2003). LDA provides a statistical framework for understanding the latent topics or themes running through a corpus by explicitly modelling the random process responsible for producing a document. The LDA model assumes that each document is made up of a mixture of topics, as well as a mixture of words associated with each topic. For instance, the document you are reading at this moment includes a mixture of themes such as "climate scepticism" and "text analysis," and these themes tend to use different language—the topic "climate scepticism" is likely associated with the word "denial," whereas the topic "text analysis" is associated with the word "random." Moreover, this process is probabilistic in the sense that we could have used the term "stochastic" instead of "random" in the previous sentence.

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

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

$$\theta \sim Dirichlet(\alpha)$$

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

$$\beta \sim Dirichlet(\eta)$$

3. For each of the N words w_n:
   Randomly sample a topic $z_n \sim Multinomial(\theta)$.
   Choose a word w_n from $p(w_n|z_n, \beta)$.

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

3.2.1. How many topics?

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

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

4. Results

4.1. Model estimation and topic interpretation

We estimate the model using the sparse Gibbs sampler described in Yao et al. (2009) and the hyperparameter optimization routine utilized in Wallach et al. (2009a). Consistent with the findings in Wallach et al. (2009a), we found that optimizing α, while fixing β, provided the easiest results to interpret and thus employ this specification. Moreover, given that mixture models such as the LDA are known to produce multimodal likelihood surfaces, we used a number of different random starting values. We found a good deal of stability in the estimated topic distributions across runs, improving our confidence that the model converged on a global optimum.

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

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

[1] AlSumait 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).

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

> The existence of the [Medieval Warm Period] had been recognized in the scientific literature for decades. But now it was a major embarrassment to those maintaining that the 20th century warming was truly anomalous. It had to be "gotten rid of" (NCPA, Dec. 6, 2006).

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

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

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

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

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

Further, the integrity of climate scientists is also frequently questioned, especially in relation to the peer-review process of the IPCC (topic 27) and other perceived violations of scientific integrity (topic 14) such as the so-called "climategate" email controversy of late 2009 which supposedly has dealt a "death blow" to the global warming "fraud" (Heartland Institute, Nov. 21, 2009). Numerous documents take aim at the credibility of climate scientists; the following excerpt serving as a typical example.

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

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

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

5. Assessing model quality: reliability and validity

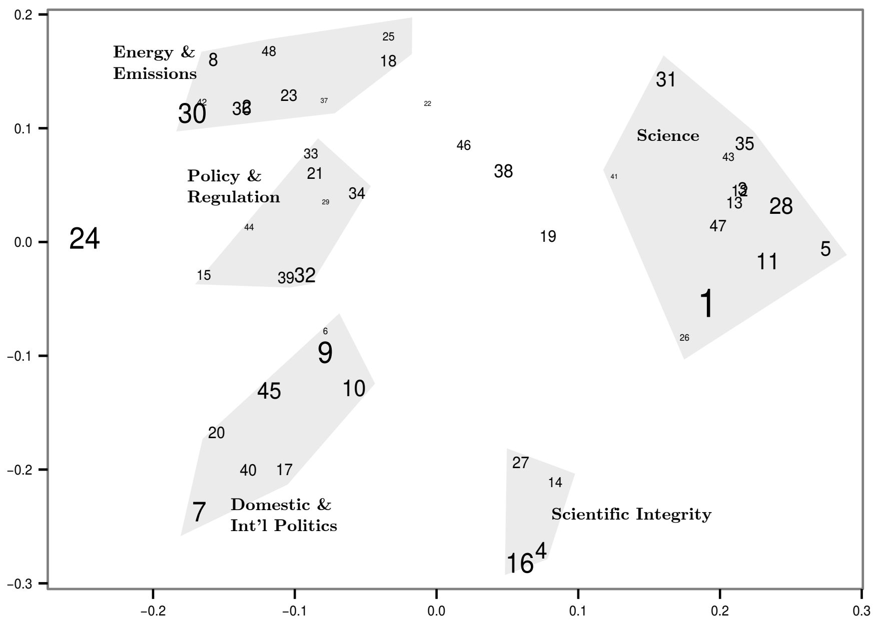
It is crucial when coding themes to establish sufficient levels of reliability and validity. Traditionally, difficulties associated with determining reliability have plagued content analytic studies, as a single coder's judgements may be highly subjective. While subsequent studies have shown that relying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

Although the benefits of employing automated methods for reliability are clear, the same cannot be said for validity and thus the onus is on the researcher to establish the soundness of their results when using computer-assisted approaches. Grimmer and Stewart (2013), in a review of the text analysis literature in political science, argue emphatically for the need to "[v]alidate, validate, validate," stating "that what should be avoided, then, is the blind use of any method without a validation step" (pg. 5). This section devotes considerable attention to this "validation step," using multiple methods to examine diverse conceptions of validity. Specifically, we 1) provide further evidence of the *semantic* validity of our findings, 2) assess *predictive* validity via external events, and 3) examine *concurrent* validity by comparing the model output to a human gold standard.

5.1. Semantic validity and topic similarity

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

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



1: Climate sensitivity
2: Fossil fuel production
3: Sea level rise
4: No scientific consensus
5: Long-term climate trends
6: Public opinion
7: U.S. politics
8: Renewable energy
9: Costs of gov't intervention
10: Environmentalism
11: Climate models
12: Solar forcing and clouds
13: Temperature data
14: Scientific misconduct
15: Gov't agencies
16: Alarmism
17: International relations
18: Agricultural industry
19: Human health
20: Corporations & the environ.
21: Urban development
22: Reuse & recycle
23: Nuclear power
24: Green jobs
25: Econ. impact of climate policy
26: Monckton
27: IPCC integrity
28: Storms
29: Emissions reduction
30: Plant impacts
31: Int'l development
32: Tax & spend
33: Conservation
34: Forest impacts
35: Cap & trade
36: Public transportation
37: Climate adaptation
38: EPA
39: Law
40: State climate reports
41: State clmate policy
42: Acidification
43: Disaster costs
44: International agreements
45: Pollution
46: Endangered species
47: Fuel standards

Figure 1: *Topic similarities.* The figure presents Jenson-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.

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

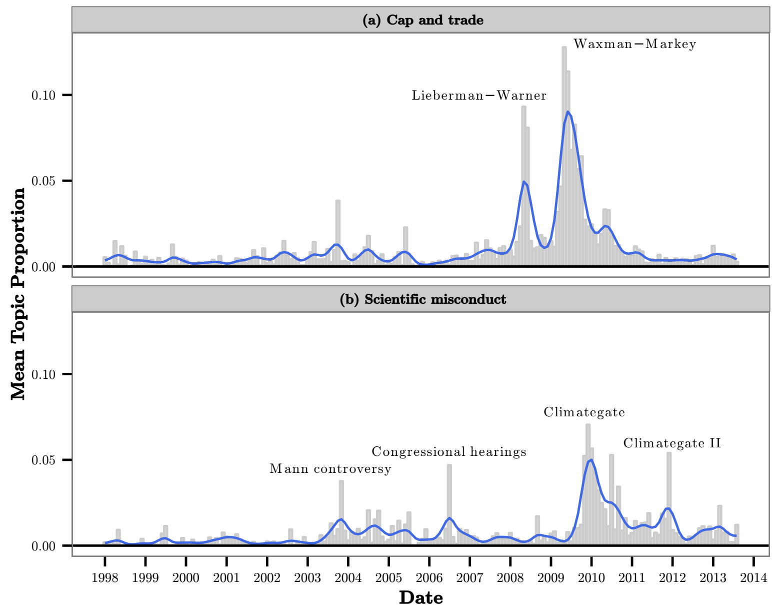


Figure 2: *Predictive validity based on external events.* The graphs illustrate the average monthly topic proportions of four topics over the period January 1998 – August 2013. A local polynomial trend line is included to assist interpretation.

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

5.2. Predictive validity and topic dynamics

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

5.2.1. *Measuring topic prevalence over time*

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

5.2.2. *Assessing predictive validity via external events*

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

Figure 2(b) displays the share of words dealing with a scientific misconduct theme. A sustained period of interest seems to cover the 2003-2005 period, with the release of papers from climate sceptics such as Stephen McIntyre, Ross McKitrick, and Hans von Storch, which criticize Michael Mann's methodology. The next substantial increase in the topic proportion is observed in July 2006, when Congressional hearings were held on the validity of Mann and colleagues' findings. However, a real break in the series occurs in November-December 2009. This is expected since this period coincides with the time when emails of researchers from the Climatic Research Unit (CRU) at the University of East Anglia were hacked, uploaded to the Internet, and subsequently scrutinized by climate sceptics. Following this flurry of attention to scientific integrity during late 2009 and early 2010, a downward trend then follows with significant peaks occurring in July 2010 when the Independent Climate Change Email Review was released and December 2011 which was just after a second round of CRU emails were uploaded to the Internet; an incident named "climategate II" by climate sceptics.

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

5.3. Assessing concurrent validity via a human "gold standard"

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

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

[2]The 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.

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

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

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

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

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

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

[3] Italics 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.

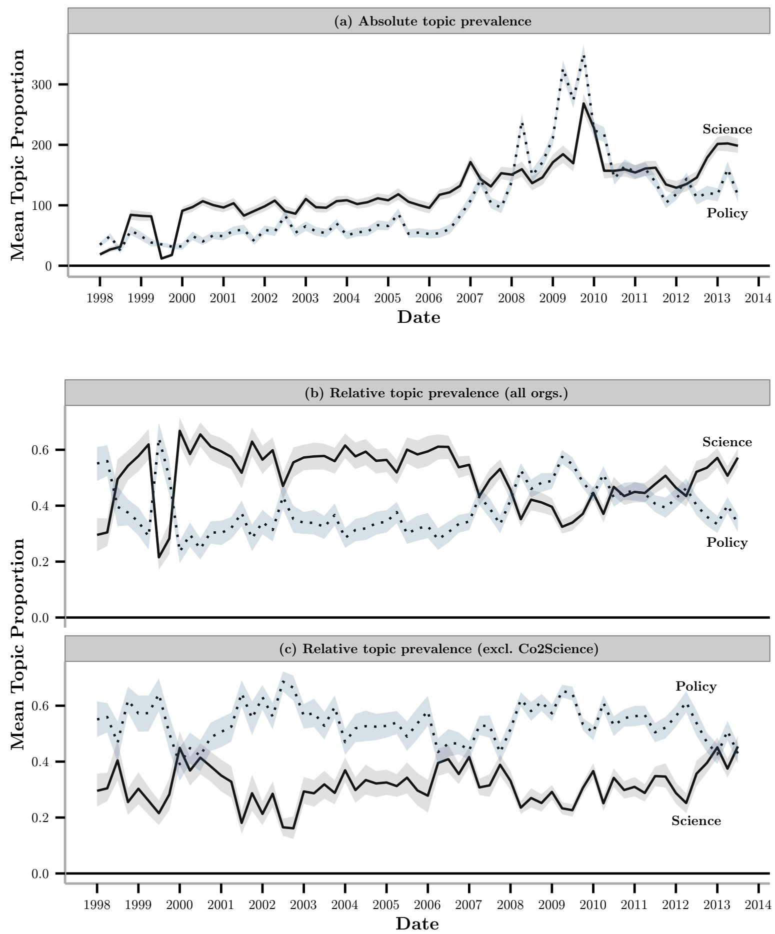


Figure 3: *The evolution of political and science-related discourse.* Panel (a) displays the summed quarterly topic probability of "science" (solid) and "politics & policy" (dotted) related themes for all CTTs in the sample over the period January 1998 – August 2013. These categories are aggregations 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.

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

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

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

[4]Note 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.

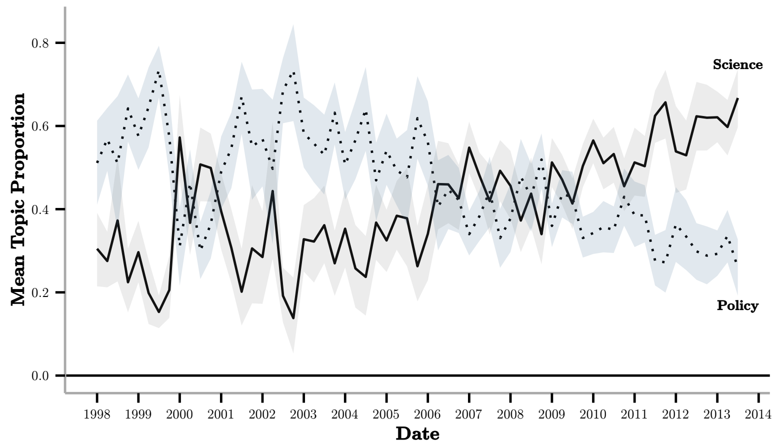


Figure 4: *The Heartland Institute's political and science-related discourse.* Displays average quarterly topic probabilities for science- and policy-related themes in documents disseminated by Heartland over the period January 1998–August 2013.

2003, Cann 2015).

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

Another potential source of heterogeneity relates to our categorizations of science and policy related discussions. It is clear that some topics labelled as "policy" are only tangentially related to "climate" policy and that there are important differences between climate science and scientific integrity. We therefore examine three themes which are directly related to climate science and policy: "Science," "Scientific Integrity," and "Energy and Emissions Policy." Figure 5

[5] http://www.desmogblog.com/directory/vocabulary/2782

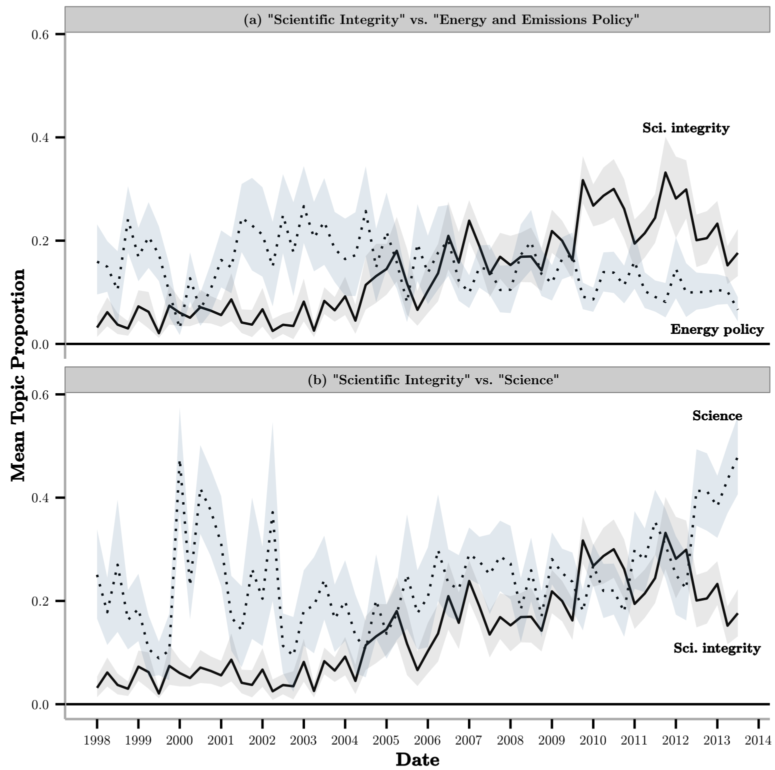


Figure 5: *Climate-specific related themes over time.* The figures show the average quarterly topic proportions of three topic clusters, which are directly related to climate science and policy, as classified in Section 5.1: "Science," "Scientific Integrity," and "Energy and Emissions Policy." Note that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

provides the results of this comparison. Several features of this figure are notable. First, considering the "Scientific Integrity" series, there has been an appreciable rise in the prevalence of integrity-related topics starting in 2004 and peaking in 2011. Second, talk of scientific integrity began to overtake that of energy policy during 2006 and 2007—which corresponds to a period dominated by *An Inconvenient Truth* and Al Gore's acceptance of the Nobel Peace Prize—and proceeded to become relatively more prevalent in the post-"climategate" era (Figure 5 (a)). Lastly, while the discussion of climate "Science" was more frequent relative to "Scientific Integrity" from 1998 to roughly 2004, the two series become intertwined for much of the sample period. This suggests that CTTs were just as likely to question the integrity of individual scientists and scientific bodies than to discuss alternative scientific viewpoints; though, there has been a perceptible break since 2012, with discussions of "Science" once again dominating the conversation.

7. Conclusion

Despite urgent calls to action among climate scientists, the U.S. government continues to avoid comprehensive climate policy action and the American public remains misinformed on key aspects of the debate. A growing literature draws attention to the influence of a well-organized and well-funded movement of climate sceptics. This study provided the first systematic update of the claims making activity of conservative think tanks—a critical piece of the climate countermovement—since the influential work of McCright and Dunlap (2000). Our key findings include:

1. The overall level of CTT claims-making has grown rapidly over the past decade and a half, reaching a peak during late 2009–early 2010;
2. The 19 CTTs studied address a wide range of topics in their written communication since McCright and Dunlap (2000), which cluster into distinct themes associated with politics, policy, science, and scientific integrity;
3. Topics questioning the integrity of individual scientists and scientific bodies appear closer (semantically) to politics than science, suggesting that claims often considered the hallmark of scientific scepticism are rooted in politics;
4. The era of climate science denial is not over. While the aggregate results demonstrate that both policy and science discussions remain stable throughout the period of study (Figure 3), a detailed analysis of a critical CTT (Figure 4) and a focus on climate change-specific themes (Figure 5) reveal the increased importance of both science and scientific integrity discussions over the sample period.
5. CTTs tend to react to the external environment—i.e., they *counter* claims—and thus studies focusing on narrow intervals of time (or a single organisation) are likely sensitive to these contextual factors.

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

References

AlSumait, L., Barbara, D., Gentle, J., Domeniconi, C., 2009. Topic significance ranking of lda generative models. In: ECML.

Anderegg, W., Prall, J., Harold, J., Schneider, S., 2010. Expert credibility in climate change. Proceedings of the National Academy of Sciences of the United States of America 107, 12107–9.

Arun, R., Suresh, V., Madhavan, C. V., Murthy, M. N., 2010. On finding the natural number of topics with latent dirichlet allocation: Some observations. In: Advances in Knowledge Discovery and Data Mining. Springer, pp. 391–402.

Austin, A., 2002. Advancing accumulation and managing its discontents: The us antienvironmental countermovement. Sociological Spectrum 22 (1), 71–105.

Beder, S., 2001. Research note-neoliberal think tanks and free market environmentalism. Environmental Politics.

Begley, S., Conant, E., Stein, S., Clift, E., Philips, M., 2007. The truth about denial. Newsweek 150 (7), 20–7.

Blei, D. M., 2012. Probablistic topic models. Communications of the ACM 55 (4), 77–84.

Blei, D. M., Ng, A. Y., Jordan, M. I., 2003. Latent dirichlet allocation. Journal of Machine Learning Research 3, 993–1022.

Boykoff, M., 2008. Lost in translation? unites states television news coverage of anthropogenic climate change, 1995-2004. Climatic Change 86, 1–11.

Boykoff, M., Boykoff, J., 2004. Balance as bias: global warming and the us prestige press. Global Environmental Change 14, 125–136.

Brulle, R. J., 2014. Institutionalizing delay: foundation funding and the creation of u.s. climate change counter-movement organizations. Climatic Change 122, 681–694.

Brulle, R. J., Carmichael, J., Jenkins, J. C., 2012. Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the us, 2002–2010. Climatic change 114 (2), 169–188.

Cann, H. W., 2015. Climate change, still challenged: Conservative think tanks and skeptic frames.

Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., Blei, D. M., 2009. Reading tea leaves: How humans interpret topic models. In: NIPS.

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., Skuce, A., 2013. Quantifying the consensus on anthropogenic global warming in the scientific literature. Environmental Research Letters 8, 024024.

Ding, D., Maibach, E. W., Zhao, X., Roser-Renouf, C., Leiserowitz, A., 2011. Support for climate policy and societal action are linked to perceptions about scientific agreement. Nature Climate Change 1 (9), 462–466.

Doran, P., Zimmerman, M., 2009. Examining the scientific consensus on climate change. Eos, Transactions American Geophysical Union 90 (3), 22–23.

Dunlap, R. E., Jacques, P. J., 2013. Climate change denial books and conservative think tanks: exploring the connection. American Behavioral Scientist, 0002764213477096.

Dunlap, R. E., McCright, A. M., 2011. Organized climate change denial. The Oxford handbook of climate change and society, 144–160.

Elgesem, D., Steskal, L., Diakopoulos, N., 2015. Structure and content of the discourse on climate change in the blogosphere: The big picture. Environmental Communication 9 (2), 169–188.

Elsasser, S. W., Dunlap, R. E., 2013. Leading voices in the denier choir: Conservativecolumnists dismissalof global warming and denigration of climate science. American Beh 57 (6), 754–776.

Endres, D. M., Schindelin, J. E., 2003. A new metric for probability distributions. IEEE Transactions on Information Theory 49 (7), 1858–1860.

Feldman, L., Maibach, E., Roser-Renouf, C., Leiserowitz, A., 2012. Climate on cable: The nature and impact of global warming coverage on fox news, cnn, and msnbc. The International Journal of Press/Politics 17 (1), 3–31.

Freudenburg, W. R., 2000. Social constructions and social constrictions: Toward analyzing the social construction of the naturalizedas well as the natural.. Environment and global modernity, 103–119.

Gower, J. C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53, 325328.

Greenpeace, 2010. Dealing in Doubt: The Climate Denial Industry and Climate Science: A Brief History of Attacks on Climate Science, Climate Scientists and the IPCC. Amsterdam: Greenpeace International.

Griffiths, T. L., Steyvers, M., 2004. Finding scientific topics. Proceedings of the National Academy of Sciences of the United States of America 101, 5228–5235.

Grimmer, J., King, G., 2011. General purpose computer-assisted clustering and conceptualization. Proceedings of the National Academy of Science Inagural Articles, 1–8.

Grimmer, J., Stewart, B. M., 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political Analysis 21 (3), 267–297.

Hart, P., 2008. Market influences on climate change frames in cnn and fox news channel broadcasts. In: International Communication Association Annual Meeting.

Hoffman, A. J., 2011. Talking past each other? cultural framing of skeptical and convinced logics in the climate change debate. Organization & Environment 24 (1), 3–33.

Holliman, R., 2011. Advocacy in the tail: Exploring the implications of climategatefor science journalism and public debate in the digital age. Journalism 12 (7), 832–846.

IPCC, 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland.

Jacques, P., Dunlap, R., Freeman, M., 2008. The organisation of denial: Conservative think tanks and environmental scepticism. Environmental Politics 17 (3), 349–385.

Knight, G., Greenberg, J., 2011. Talk of the enemy: adversarial framing and climate change discourse. Social Movement Studies 10 (4), 323–340.

Lahsen, M., 2008. Experiences of modernity in the greenhouse: A cultural analysis of a physicist trio supporting the backlash against global warming. Global environmental change 18 (1), 204–219.

Lakoff, G., 2014. The All New Don't Think of an Elephant!: Know Your Values and Frame the Debate. Chelsea Green Publishing.

Laver, M., Benoit, K., Garry, J., 2003. Extracting policy positions from political text using words as data. American Political Science Review 97 (3), 311–331.

Laver, M., Garry, J., 2000. Estimating policy positions from political texts. American Journal of Political Science 44 (3), 619–634.

Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinburg, G., Rosenthal, S., 2015. Climate change in the american mind: March 2015. Yale University and George Mason University. New Haven, CT: Yale Project on Climate Change Communication.

Mann, M. E., 2013. The hockey stick and the climate wars: Dispatches from the front lines. Columbia University Press.

McCright, A., Dunlap, R., 2000. Challenging global warming as a social problem: An analysis of the conservative movement's counter-claims. Social Problems 47 (4), 499–522.

McCright, A., Dunlap, R., 2003. Defeating kyoto: The conservative movement's impact on u.s. climate change policy. Social Problems 50 (3), 348–373.

McCright, A. M., Dunlap, R. E., Xiao, C., 2013. Perceived scientific agreement and support for government action on climate change in the usa. Climatic Change 119 (2), 511–518.

National Research Council, 2010. Advanicing the Science of Climate Change. National Academies Press.

Nisbet, M. C., 2009. Communicating climate change: Why frames matter for public engagement. Environment: Science and Policy for Sustainable Development 51 (2), 12–23.

Oreskes, N., 2004. The scientific consensus on climate change. Science 306, 1686.

Oreskes, N., Conway, E., 2010. Merchants of Doubt; How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming. New York: Bloomsbury.

Osterreicher, F., Vajda, I., 2003. A new class of metric divergences on probability spaces and its applicability in statistics. Annals of the Institute of Statistical Mathematics 55 (3), 639–653.

ONeill, S., Boykoff, M., 2011. The role of new media in engaging the public with climate change. Engaging the public with climate change: behaviour change and communication, 233–251.

Painter, J., Ashe, T., 2012. Cross-national comparison of the presence of climate skepticism in the print media in six countries, 2007-10. Environmental Research Letters 7, 1–8.

Pooley, E., 2010. The climate war: true believers, power brokers, and the fight to save the earth. Hachette UK.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., Radev, D. R., 2010. How to analyze political attention with minimal assumptions and costs. American Journal of Political Science 54 (1).

Roberts, M. E., Stewart, BrandStewart, B. D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., Rand, D. G., 2014. Structural topic models for open-ended survey responses. American Journal of Political Science 58 (4), 10641082.

Scheufele, D. A., Tewksbury, D., 2007. Framing, agenda setting, and priming: The evolution of three media effects models. Journal of communication 57 (1), 9–20.

Schmidt, A., Ivanova, A., Schäfer, M. S., 2013. Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries. Global Environmental Change 23 (5), 1233–1248.

Sharman, A., 2014. Mapping the climate sceptical blogosphere. Global Environmental Change 26, 159–170.

Union of Concerned Scientists, 2007. Smoke, Mirrors & Hot Air: How ExxonMobil Uses Big Tobacco's Tactics to Manufacture Uncertainty on Climate Science. Union of Concerned Scientists.

van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D., Maibach, E. W., 2015. The scientific consensus on climate change as a gateway belief: Experimental evidence. PloS one 10 (2), e0118489.

Wallach, H. M., Mimno, D., Mccallum, A., 2009a. Rethinking lda: Why priors matter. NIPS.

Wallach, H. M., Murray, I., Salakhutdinov, R., Mimno, D., 2009b. Evaluation methods for topic models. In: ICML.

Washington, H., Cook, J., 2011. Climate Change Denial: Heads in the Sand. Earthscan, 711 Third Avenue, New York, NY, 10017.

Weber, E. U., Stern, P. C., 2011. Public understanding of climate change in the united states. American Psychologist 66 (4), 315.

Yao, L., Mimno, D., McCallum, A., 2009. Streaming inference for latent dirichlet allocation. kdd, 2009. KDD.

Young, N., 2013. Working the fringes: the role of letters to the editor in advancing non-standard media narratives about climate change. Public Understanding of Science 22 (4), 443–459.