Global Environmental Change 36 (2016) 89–100

Contents lists available at ScienceDirect

Global Environmental Change

journal homepage: www.elsevier.com/locate/gloenvcha

Text-mining the signals of climate change doubt

Constantine Boussalis [a], Travis G. Coan [b,*]

[a] Department of Political Science, Trinity College Dublin, 3 College Green, Dublin 2, Ireland
[b] Department of Politics and Exeter Q-Step Centre, University of Exeter, Amory Building, Rennes Drive, Exeter, Devon EX4 4RJ, United Kingdom

ARTICLE INFO

Article history:
Received 27 September 2015
Received in revised form 22 November 2015
Accepted 10 December 2015
Available online 30 December 2015

Keywords:
Climate change
Scepticism
Text mining
Latent Dirichlet allocation

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 overview of conservative think tank sceptical discourse in nearly 15 years. Specifically, we (1) compile the largest corpus of contrarian literature to date, collecting over 16,000 documents from 19 organizations over the period 1998–2013; (2) introduce a methodology to measure key themes in the corpus which scales to the substantial increase in content generated by conservative think tanks over the past decade; and (3) leverage this new methodology to shed light on the relative prevalence of science- and policy-related discussion among conservative think tanks. 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.

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 10 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 change 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'" in 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 by 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: offering 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 countermovement has prompted calls for an expansion and improvement of data collection efforts on a range of climate movement and countermovement activities (Brulle et al., 2012).

Despite an active interest in CTTs, few studies have systematically analysed the nature and prevalence of contrarian themes. 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

* Corresponding author.
E-mail address: T.Coan@exeter.ac.uk (T.G. Coan).

http://dx.doi.org/10.1016/j.gloenvcha.2015.12.001
0959-3780/

little is known about how the contrarian discourse has evolved over the last decade. We seek to fill this gap in the literature by (1) compiling the largest corpus of climate sceptic literature to date, collecting over 16,000 documents from 19 organizations over the period 1998–2013; (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 themes over a 16 year period. We argue that understanding the nature and prevalence of CTT misinformation 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. Literature review

2.1. Theoretical perspectives on the "countermovement"

There is a growing literature on the theoretical processes responsible for driving the political debate on the issue of climate change (see Dunlap and McCright, 2011, 2015 for comprehensive reviews). Drawing heavily on social movement theory, the predominate framework views organized climate scepticism as a "countermovement," which evolved to combat environmental activists' early success in raising awareness on the issue of AGW (Dunlap and McCright, 2015). Brulle (2014) further situates the dynamics of contention on AGW in the context of field frame analysis (Fligstein and McAdam, 2012) and, as such, seeks to provide a coherent societal level explanation for the evolution of dominant frames which define the issue of global warming. Within the context of field frame analysis, the "field" of AGW is contested by forces supporting and opposing the dominant frame (see Meyer and Staggenborg, 1996; Levy and Egan, 2003). Specifically, the environmental movement is viewed as promoting social change, the denial countermovement is viewed as preserving the status quo (i.e., the preservation of the material interests of industry), and both sets of actors are competing to define the dominant field frame. Social change is thus viewed as a conflict to determine how climate change is understood within a wider social, political, and cultural context (Knight and Greenberg, 2011).

While field frame analysis offers a useful framework for theorizing on societal level conflict between movement and countermovement, it provides little insight into the actors, interests, and tactics of those responsible for "creat[ing] a paralyzing fog of doubt around climate change" (Begley et al., 2007). Fortunately, a well-developed literature systematically describes the core elements of the countermovement. Viewed largely as an extension of the conservative movement in the U.S., organized climate denial was born out of the deep pockets of conservative foundations and corporate interest groups committed to promoting free-market principles and rolling back government intervention in all aspects of the economy (Brulle, 2014; Dunlap and McCright, 2011). Consistent with a policy platform committed to economic growth, the conservative movement instituted a number of direct challenges to key environmental policies enacted during the 1970s (Dunlap, 1987). Yet, after suffering public defeats on environmental issues during the 1980s and early 1990s, conservatives quickly learned that directly challenging key environmental policies was fraught with risks (Bonds, 2003; Dunlap, 1987) and, as such, "shifted to a more subtle form of power characterized by non-decision-making and agenda setting" (Dunlap and McCright, 2015, p. 306). When global warming made its way onto the political agenda in the late 1980s, interest opposed to action on AGW implemented a campaign to "manufacture doubt" about the credibility of individual scientists and exaggerating scientific uncertainties (Union of Concerned Scientists, 2007; Oreskes and Conway, 2010; Greenpeace, 2010; Dunlap and McCright, 2011).

It is within the shift from direct to indirect challenges to environmental policy that the full importance of CTTs in the denial countermovement comes into view. First, relying on their image as the "alternative academia" or "counter-intellegentsia" and their ability to marginalize mainstream academia (Medvetz, 2012), CTTs play a lead role in constructing viewpoints that challenge the orthodox position 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). Second, as the engine of information in the "denial machine," CTTs are the agents actually responsible for "framing the field" of AGW. 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). As such, CTTs arguably provide the "'connective tissue' that helps hold the denial countermovement together" (Dunlap and McCright, 2015, p. 312).

Against this backdrop, a general picture is beginning to emerge on the dynamic process responsible for manufacturing uncertainty and controversy on AGW (Pooley, 2010; Oreskes and Conway, 2010; Washington and Cook, 2011; Mann, 2013). Specifically, extant literature suggests the following process: (1) conservative foundations and corporate groups provide the material base for pressing contrarian interests (Brulle, 2014); (2) CTTs transform this material base into information, generating the narrative of climate denial (McCright and Dunlap, 2000); (3) the conservative "echo chamber"—conservative media, sceptical blogs, and sympathetic policy makers—mediate and amplify key counterclaims (Dunlap and McCright, 2011); and (4) conservative politicians susceptible to the anti-climate messages seek to stymie policy changes in Congress (McCright and Dunlap, 2003). Nevertheless, despite a general understanding, considerably more research is needed to fully specify the linkages between key actors in the denial countermovement and longitudinal data is necessary to test dynamic theories of organized climate scepticism (Dunlap and McCright, 2015). The remainder of this study seeks to achieve this latter goal by generating time-series data of the sceptical discourse espoused by CTTs.

2.2. Past measures of the contrarian discourse

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" (O'Neill and Boykoff, 2011; Holliman, 2011; Knight and Greenberg, 2011; Sharman, 2014; Elgesem et al., 2015). However, few studies systematically analyse the *content* of contrarian information 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, p. 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 the conversation has evolved. Several studies provide a more recent look at the key features of the contrarian discourse more generally. Elsasser and Dunlap (2013) employ 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 sceptic 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 emphasize 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 organization 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.

3. Systematically measuring CTT climate change scepticism

3.1. The corpus

To systematically gauge contrarian themes, 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 an optical character recognition (OCR) algorithm 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 (see Section A.1.2).

The table provides a number of insights into the volume of information produced by the most important CTTs. First, these

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.

			Document type				
Organization name	Total words thous.	Total docs.	A	B	C	D	E
American Enterprise Institute (AEI)	1872.53	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)	2387.27	4592	713	0	0	1	3878
Competitive Enterprise Institute (CEI)	1743.02	1461	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	9900.54	2930	1383	1537	10	0	0
Heritage Foundation	1825.78	1652	1198	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)	3064.88	552	0	552	0	0	0
Total	24,894.58	16,028	8368	2967	105	684	3904

organizations 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 organizations 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 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, assuming 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 (see Section A.2), 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

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), 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) score, 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"). Lastly, based on an evaluation of the most

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, we code whether each topic is related to science (S) or politics and policy (P).

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

probable keywords and documents for each theme, we classify each topic as being generally related to either "science" (S) or "politics and policy" (P) (for a similar approach, see Sharman, 2014).

4.2. Topic interpretation

Looking at the full list of topics shown in Table 2, the model predicts 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. While past literature on CTTs have focused predominantly on "counter-claims" used to challenge AGW (see Section 2.2), our analysis concentrates on more general topics or themes. The two concepts, however, are closely related. Broadly speaking, we identify a number of quite specific counter-claims that are nested within our more general set of themes. In an effort to provide a more in-depth interpretation of the estimated topics, this section highlights the interplay between themes and counter-claims for a number of key areas of discussion.

As expected, the corpus is rife with documents that discuss the scientific basis of AGW and, not surprisingly, draw on the counter-claim that "[t]he evidentiary basis of global warming is weak and even wrong" (McCright and Dunlap, 2000, p. 510). For instance, 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, November 11, 2011) and any discussion to the contrary amounts to "alarmism" (Heartland Institute, May 17, 2013). Moreover, the general agreement of scientists on this relationship is repeatedly challenged within the corpus (topic 4), as demonstrated in the following assertion:

> Contrary to what you read repeatedly in daily newspapers or hear on television, most scientists do not believe there is a "scientific consensus" that man-made climate change [...] is or will be a catastrophe [...] The '97% of climate scientists' claim is garbage. Anyone who cites it ought to be ashamed. (Heartland Institute, July 7, 2013).

Appeals to long-term natural cycles in temperature (topic 5), as purportedly demonstrated by the Roman and Medieval Warm Periods, are also common. 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:

> 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, December 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 also cast into doubt. Further, the underlying assumptions of climate change models (topic 11) that are referenced in the IPCC assessments are considered to be of "dubious merit" (Fraser, July 7, 2004).

In addition to disputing the accuracy of climate science, the integrity of specific climate scientists and scientific bodies are frequently questioned. This pattern is especially true in relation to the peer-review process of the IPCC (topic 27) and is consistent with the sub-claim that "[t]he IPCC intentionally altered its reports

to create a 'scientific consensus' on global warming" presented in McCright and Dunlap (2000, p. 510). Further, the so-called "climategate" email controversy (topic 14) of late 2009, which supposedly has dealt a "death blow" to the global warming "fraud" (Heartland Institute, November 21, 2009), is presented as a watershed event in numerous documents. The following excerpt serves 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, November 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).

Table 2 also demonstrates the breadth of topics discussed in documents referencing climate change with important issue linkages across both the domestic and international political economy. International mitigation policies (topic 44) are often framed as threats to national sovereignty and are expected to have detrimental impacts on 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. The following excerpt offers an illustration of the arguments related to the economy:

> 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, then, a considerable number of documents that focus on policy-related themes rely on the general counter-claim that "[g]lobal warming policies would do more harm than good" and we see considerable evidence of the sub-claim that "[p]roposed action would harm that national economy" (McCright and Dunlap, 2000, p. 510).

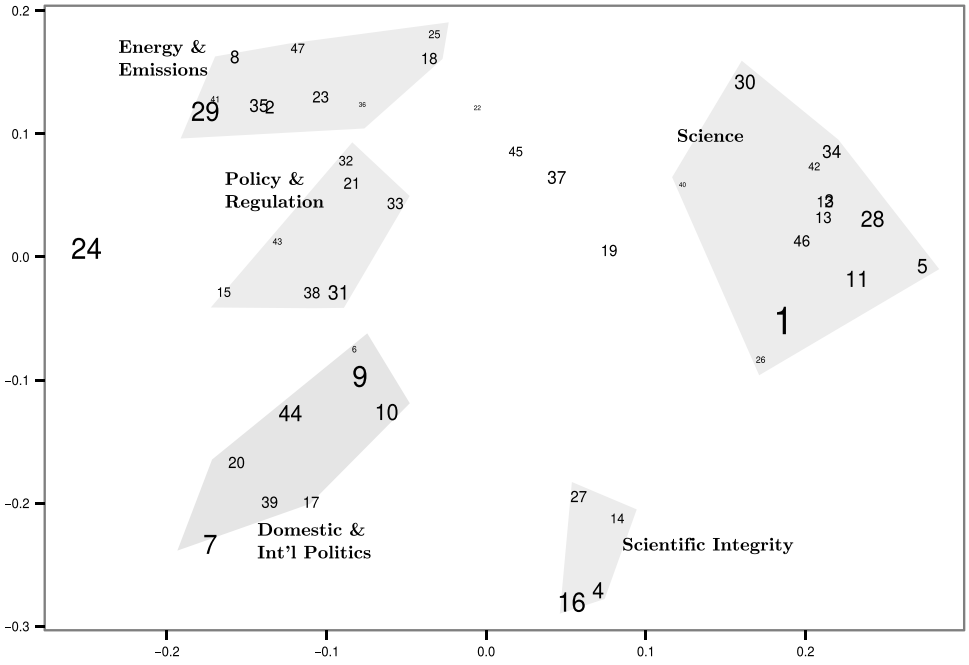
5. Assessing model quality: reliability and validity

It is crucial when coding themes to establish sufficient levels of reliability and validity. Although the benefits of employing automated methods for reliability are clear (Laver and Garry, 2000; Laver et al., 2003), 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" (p. 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 and (2) examine *concurrent* validity by comparing the model output to a human gold standard (see Section A.3 of the online appendix for an assessment of the *predictive* validity of our model via external events).

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. Fig. 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 Fig. 1, while dissimilar themes should be further away.

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. 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



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 climate policy
42: Acidification
43: Disaster costs
44: International agreements
45: Pollution
46: Endangered species
47: Fuel standards

Fig. 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.

science" to 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. 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 (see Section A.3 of the online appendix for more information on the coding procedure employed). In order to compare the quality of our model's predictions vis-à-vis the manually annotated "gold standard," we utilize "microaveraged" versions of recall and precision, which are commonly employed to judge classifier quality in "multi-class" (i.e., multiple, distinct categories) classification problems (Manning et al., 2008; see Section A.3 for more information on these evaluation metrics). We find that the microaveraged 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 "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 is 0.78 if one considers the first two most probable topics based on the model.

6. Policy vs. 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* (Luntz, 2002).

If indeed the window of opportunity for scientific scepticism has closed, this would be a welcome development for proponents of climate action. 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 organized 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. Fig. 2(a) presents an "absolute" measure of topic prevalence by relying on the sum of the topic proportions for "science" and "politics & policy" related topics for each quarter over the Q1/1998–Q3/2013 period, while Fig. 3(b) and (c) provides a "relative" measure of prevalence by using the mean topic probabilities (see Section A.4.1 of the online appendix for more information on measuring topic prevalence). Each time series also includes an estimate of uncertainty, as measured by a bootstrapped 95% confidence interval. 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 1000 replicates for each series. Further, the categories utilized in the figure are aggregations of topics following the coding scheme for "science" and "politics & policy" presented in Table 2.

Several aspects of Fig. 2 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

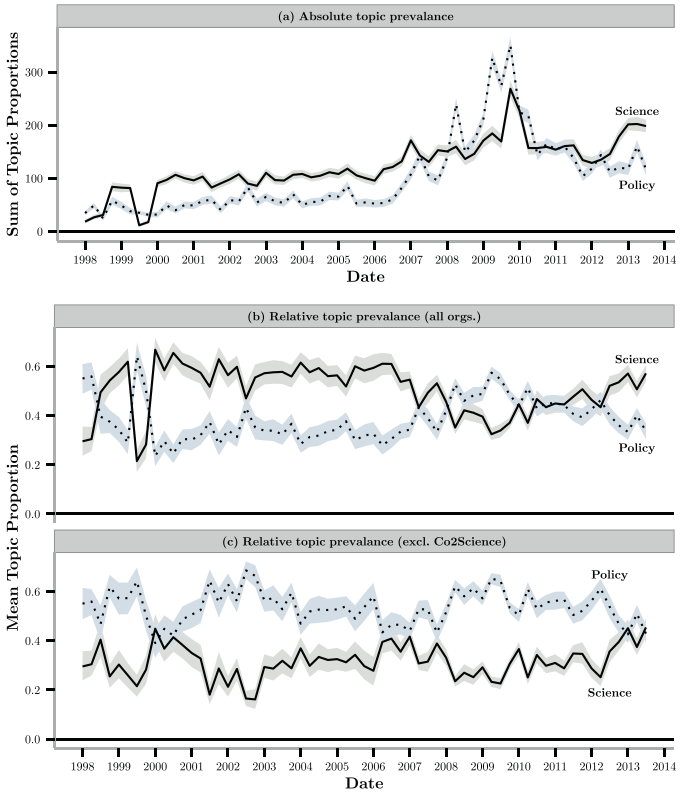


Fig. 2. The evolution of political and science-related discourse. Panel (a) displays the summed quarterly topic probability of "science" (solid) and "politics & policy" (dotted) related themes for all CTTs in the sample over the period January 1998–August 2013. These categories are aggregations of the 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.

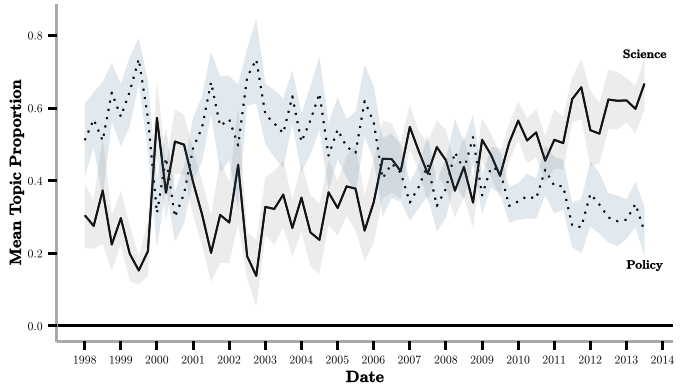


Fig. 3. The Heartland Institute's political and science-related discourse. Displays average quarterly topic probabilities for science- and policy-related themes in documents disseminated by Heartland over the period January 1998–August 2013.

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 Fig. 2(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 Fig. 2(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 Fig. 2, 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 latter category, the Heartland Institute stands out as an important countermovement organization 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 organized climate scepticism (McCright and Dunlap, 2003; Cann, 2015).

How, then, does Heartland's discourse on "science" and "politics & policy" related themes compare to the general trend illustrated in Fig. 2? We narrow our focus on Heartland in Fig. 3. As shown in the figure, beginning in 2002, we 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, 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." Fig. 4 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 (Fig. 4(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.

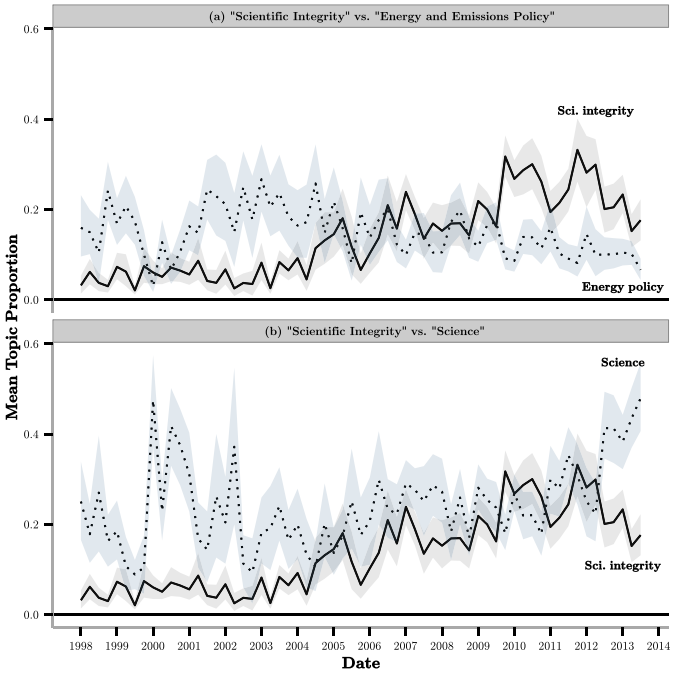


Fig. 4. Climate-specific related themes over time. The figures show the average quarterly topic proportions of three topic clusters, which are directly related to climate science and policy, as classified in Section 5.1: "Science," "Scientific Integrity," and "Energy and Emissions Policy." Note that Co2Science has been excluded from this analysis. The series covers the period Q1/1998–Q3/2013.

7. Discussion

Theoretical progress in the field of organized climate change denial, among other things, "demands the collection and rigorous, systematic examination of longitudinal data" on the "discourses, claims, and frames employed by key components in the denial countermovement" (Dunlap and McCright, 2015, p. 322). More generally, theoretical development is contingent upon finding appropriate measures of (1) the material support flowing from industry and conservative foundations Brulle (2014), (2) the misinformation emanating from CTTs, and (3) the ways in which this information is interpreted in conservative media, sceptical blogs, and by sympathetic politicians. Our study speaks directly to CTT-produced misinformation by offering a systematic description of the key topics or themes driving contrarian discourse. In this section, we conclude with a discussion of how our findings speak to the theoretical and empirical literature on organized climate change denial.

First, looking within the denial countermovement, our analyses offer insight into the prevalence and nature of climate-related themes. Judging from the exponential increase in the documents produced by CTTs, it is clear that the misinformation campaign has been escalating over time. As shown in Fig. 2a (see also Fig. A.1 in the online appendix), the post-2009 period has produced, on average, significantly more documents than the previous period (78 vs. 119 docs/mo.) and thus these data suggest that the countermovement has entered into a new equilibrium with respect to the volume of misinformation. Beyond just the total volume of sceptic content, however, we are further able to determine the dynamics of different *types* of misinformation since the late 1990s. The analysis in Section 6 suggests that the prevalence of science-related content has remained relatively constant throughout the roughly 15 year sample period and is, if anything, increasing amongst key actors in the countermovement. This finding casts a shadow on the hope that the denial machine

has somehow changed the focus of its strategy from attacking climate science to debating the utility of climate policy; the "era of climate science denial," in other words, is not over. This is an unfortunate development, as 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).

Next, considering the wider field frame conflict over the issue of AGW, our evidence further emphasizes that CTTs are highly reactive to the external environment. The overall volume of information does not fluctuate randomly over the sample period, but rises at expected times and in expected ways. For instance, the release of Al Gore's *Inconvenient Truth* and the subsequent Nobel Prize award of both Gore and the IPCC are defining moments in the advocacy of climate policy (e.g., Jacobsen, 2011), placing AGW squarely on the political agenda (Daly et al., 2015). It is thus not surprising that the denial countermovement responded with a proportional increase in rhetorical appeals to "alarmism," as well as a general questioning of both the factual basis of AGW and the integrity of climate scientists (see "Scientific Integrity" series in Fig. 4). More generally, these data suggest that as the evidence continues to mount regarding the anthropogenic causes of global warming, the level of contrarian information is also rising in a similar fashion. Further, when American environmentalists subsequently turned their attention to comprehensive climate change policy—first with the Liberman-Warner bill in 2008 and then the Waxman-Markey bill in 2009—the countermovement once again shifted its focus, arguing that any restriction on carbon emissions would be detrimental to the national economy. The evidence, therefore, clearly suggests that CTTs are in the business of "countering claims:" their goal is to dominate the framing competition on 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).

Although these data shed light on one important element of the countermovement, a number of critical theoretical and empirical questions remain unanswered. One crucial question centres on how and to what extent the information produced by CTTs is diffused through traditional media, social media, and conservative political elites. Past literature offers some direction on the topic of information diffusion: research demonstrates how CTT information has made its way into the public sphere via expert congressional testimony (McCright and Dunlap, 2003), by being featured in the mass media (Boykoff and Boykoff, 2007), through op-eds in conservative newspapers (Elsasser and Dunlap, 2013), and contrarian books (Dunlap and Jacques, 2013). Regarding print media in the U.S., a recent study suggests that the "balancing norm" identified in Boykoff and Boykoff (2004) may be on the decline, as a handful of influential American newspapers—including 5 "liberal" papers, 2 "centrist" papers, and 1 "conservative" paper—appear to present the science on AGW in a manner more consistent with the scientific consensus (Schmid-Petri et al., 2015). This is an encouraging finding, suggesting that the rise of science-related information by CTTs is not being reflected among a sample of American prestige press outlets. However, it is not clear whether this finding holds for television media, conservative talk radio, social media (e.g., the blogosphere), or even in a larger sample of conservative newspapers. We must then extend the call for systematic longitudinal data to incorporate key outlets responsible for mediating the relationship between CTT information and the general public.

Yet, perhaps the most pivotal outstanding question is whether the doubt manufactured by CTTs actually influences public opinion on AGW. Experimental evidence suggests that it could indeed have an effect. In a recent experimental analysis, van der Linden et al. (2015) offer evidence that the level of perceived scientific agreement influences beliefs and concern about climate change, which in turn influences support for climate change action (see also Ding et al., 2011; Lewandowsky et al., 2013b,a; McCright et al., 2013; Aklin and Urpelainen, 2014). However, the majority of studies examining the influence of sceptical information on public opinion outside of the lab fail to find an effect. For instance, Brulle et al. (2012) using the number of articles related to climate change from six conservative magazines does not find a statistically significant sceptic advocacy effect on American public opinion (see also Scruggs and Benegal, 2012). These studies, nevertheless, use crude proxies of contrarian information and more precise measures—such as the ones produced in this study—are likely necessary to isolate variation in public opinion on climate change and energy policy.

8. Conclusion

Our study provides the first systematic content analytic update of the climate change related literature of conservative think tanks—a critical piece of the "denial machine"—since the influential work of McCright and Dunlap (2000). Specifically, using the largest corpus of contrarian documents assembled to date, we find:

1. The overall level of CTT information has grown rapidly over the past decade and a half, reaching a peak during late 2009–early 2010.
2. 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.
3. 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 (Fig. 2), a detailed analysis of a critical CTT (Fig. 3) and a focus on climate change-specific themes (Fig. 4) reveal the increased importance of both science and scientific integrity discussions over the sample period.
4. 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 organization) are likely sensitive to these contextual factors.

It is important to note 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 stressed, however, that these organizations 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, 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. Fourth, and perhaps most importantly, although we demonstrate the close connection between the estimated themes and specific counter-claims (see Section 4.2), more detailed qualitative research is necessary to determine the prevalence of both new and past claims. 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.

Acknowledgements

The authors wish to express their gratitude to John Cook, Joshua Clinton, Kenneth Benoit and participants at the 2013 "New Directions in Analyzing Text as Data Workshop" held at the London School of Economics and Political Science, the 2014 Midwest Political Science Association Conference, and the 2015 European Political Science Association meeting for helpful suggestions on previous versions of this paper. All errors are the authors' exclusively. We would like to thank Luke White for excellent research assistance. Lastly, we are grateful to Trinity College Dublin for funding support from a Research Committee Pathfinder Programme grant (204259.13549).

Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.gloenvcha.2015.12.001.

References

Aklin, M., Urpelainen, J., 2014. Perceptions of scientific dissent undermine public support for environmental policy. Environ. Sci. Policy 38, 173–177.
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. Proc. Natl. Acad. Sci. U.S.A. 107, 12107–12109.
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. Sociol. Spectr. 22 (1), 71–105.
Beder, S., 2001. Research note-neoliberal think tanks and free market environmentalism. Environ. Polit..
Begley, S., Conant, E., Stein, S., Clift, E., Philips, M., 2007. The truth about denial. Newsweek 150 (7), 20–27.
Blei, D.M., 2012. Probablistic topic models. Commun. ACM 55 (4), 77–84.
Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.
Bonds, E., 2003. The knowledge-shaping process: elite mobilization and environmental policy. Crit. Sociol. 37, 429–446.
Boykoff, M., 2008. Lost in translation? Unites States television news coverage of anthropogenic climate change, 1995–2004. Clim. Change 86, 1–11.
Boykoff, M., Boykoff, J., 2004. Balance as bias: global warming and the us prestige press. Glob. Environ. Change 14, 125–136.
Boykoff, M.T., Boykoff, J.M., 2007. Climate change and journalistic norms: a case-study of us mass-media coverage. Geoforum 38 (6), 1190–1204.
Brulle, R.J., 2014. Institutionalizing delay: foundation funding and the creation of u.s. climate change counter-movement organizations. Clim. 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. Clim. Change 114 (2), 169–188.
Cann, H.W., 2015. Climate change, still challenged: Conservative think tanks and skeptic frames. Presented at the Annual Meeting of the Western Political Science Association, Las Vegas, April 2–4, 2015, pp. 2015.
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. Environ. Res. Lett. 8, 024024.
Daly, M., Gifford, L., Luedecke, G., McAllister, L., Nacu-Schmidt, A., Andrews, K., Boykoff, M., 2015. United States Coverage of Climate Change or Global Warming, 2004–2015. Center for Science and Technology Policy Research, Cooperative Institute for Research in Environmental Sciences, University of Colorado. Web: http://sciencepolicy.colorado.edu/media_coverage (accessed 17.11.15).
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. Nat. Clim. Change 1 (9), 462–466.
Doran, P., Zimmerman, M., 2009. Examining the scientific consensus on climate change. Eos, Trans. Am. Geophys. Union 90 (3), 22–23.
Dunlap, R.E., 1987. Polls, pollution, and politics revisited public opinion on the environment in the Reagan era. Environ. Sci. Policy Sustain. Dev. 29 (6), 6–37.
Dunlap, R.E., Jacques, P.J., 2013. Climate change denial books and conservative think tanks: exploring the connection. Am. Behav. Sci., 0002764213477096.
Dunlap, R.E., McCright, A.M., 2011. Organized climate change denial. Oxf. Handb. Clim. Change Soc. 144–160.
Dunlap, R.E., McCright, A.M., 2015. Challenging climate change: the denial countermovement. In: Dunlap, R., Brulle, R. (Eds.), Climate Change and Society: Sociological Perspectives. Oxford University Press, Oxford, UK, pp. 300–332.
Elgesem, D., Steskal, L., Diakopoulos, N., 2015. Structure and content of the discourse on climate change in the blogosphere: the big picture. Environ. Commun. 9 (2), 169–188.
Elsasser, S.W., Dunlap, R.E., 2013. Leading voices in the denier choir: conservative columnists dismissal of global warming and denigration of climate science. Am. Beh 57 (6), 754–776.
Endres, D.M., Schindelin, J.E., 2003. A new metric for probability distributions. IEEE Trans. Inform. 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. Int. J. Press/ Polit. 17 (1), 3–31.
Fligstein, N., McAdam, D., 2012. A Theory of Fields. Oxford University Press.
Freudenburg, W.R., 2000. Social constructions and social constrictions: toward analyzing the social construction of 'the naturalized as well as' the natural. Environ. Glob. Modern. 103–119.
Gower, J.C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53, 325–328.
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. Greenpeace International, Amsterdam.
Griffiths, T.L., Steyvers, M., 2004. Finding scientific topics. Proc. Natl. Acad. Sci. U.S.A. 101, 5228–5235.
Grimmer, J., King, G., 2011. General purpose computer-assisted clustering and conceptualization. Proc. Natl. Acad. Sci. 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. Polit. Anal. 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. Organ. Environ. 24 (1), 3–33.
Holliman, R., 2011. Advocacy in the tail: exploring the implications of climate gate for 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.
Jacobsen, G.D., 2011. The Al gore effect: an inconvenient truth and voluntary carbon offsets. J. Environ. Econ. Manage. 61 (1), 67–78.
Jacques, P., Dunlap, R., Freeman, M., 2008. The organisation of denial: conservative think tanks and environmental scepticism. Environ. Polit. 17 (3), 349–385.
Knight, G., Greenberg, J., 2011. Talk of the enemy: adversarial framing and climate change discourse. Social Movement Stud. 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. Glob. Environ. 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. Am. Polit. Sci. Rev. 97 (3), 311–331.
Laver, M., Garry, J., 2000. Estimating policy positions from political texts. Am. J. Polit. Sci. 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. Yale Project on Climate Change Communication, New Haven, CT.
Levy, D.L., Egan, D., 2003. A neo-Gramscian approach to corporate political strategy: conflict and accommodation in the climate change negotiations. J. Manage. Stud. 40 (4), 803–829.
Lewandowsky, S., Gignac, G.E., Vaughan, S., 2013. The pivotal role of perceived scientific consensus in acceptance of science. Nat. Clim. Change 3 (4), 399–404.
Lewandowsky, S., Oberauer, K., Gignac, G.E., 2013. Nasa faked the moon landing therefore, (climate) science is a hoax an anatomy of the motivated rejection of science. Psychol. Sci. 24 (5), 622–633.
Luntz, F., 2002. Memo on the Environment., In: https://www.motherjones.com/files/LuntzResearch_environment.pdf (accessed 19.11.15).
Mann, M.E., 2013. The Hockey Stick and the Climate Wars: Dispatches from the Front Lines. Columbia University Press.
Manning, C.D., Raghavan, P., Schütze, H., 2008. Introduction to Information Retrieval. Cambridge University Press, New York, NY, USA.
McCright, A., Dunlap, R., 2000. Challenging global warming as a social problem: an analysis of the conservative movement's counter-claims. Social Probl. 47 (4), 499–522.
McCright, A., Dunlap, R., 2003. Defeating Kyoto: the conservative movement's impact on U.S. climate change policy. Social Probl. 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. Clim. Change 119 (2), 511–518.

Medvetz, T., 2012. Think Tanks in America. University of Chicago Press.
Meyer, D.S., Staggenborg, S., 1996. Movements, countermovements, and the structure of political opportunity. Am. J. Sociol. 1628–1660.
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. Environ. Sci. Policy Sustain. Dev. 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. Bloomsbury, New York.
Osterreicher, F., Vajda, I., 2003. A new class of metric divergences on probability spaces and its applicability in statistics. Ann. Inst. Stat. Math. 55 (3), 639–653.
O'Neill, 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, pp. 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. Environ. Res. Lett. 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. Am. J. Polit. Sci. 54 (1.) .
Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., Rand, D.G., 2014. Structural topic models for open-ended survey responses. Am. J. Polit. Sci. 58 (4), 1064–1082.
Scheufele, D.A., Tewksbury, D., 2007. Framing, agenda setting, and priming: the evolution of three media effects models. J. Commun. 57 (1), 9–20.
Schmid-Petri, H., Adam, S., Schmucki, I., Häussler, T., 2015. A changing climate of skepticism: The factors shaping climate change coverage in the us press. Publ. Understanding Sci. (Bristol, England).
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. Glob. Environ. Change 23 (5), 1233–1248.
Scruggs, L., Benegal, S., 2012. Declining public concern about climate change: can we blame the great recession. Glob. Environ. Change.
Sharman, A., 2014. Mapping the climate sceptical blogosphere. Global Environ. 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., 2009. Rethinking LDA: Why Priors Matter. NIPS.
Wallach, H.M., Murray, I., Salakhutdinov, R., Mimno, D., 2009. Evaluation methods for topic models. In: ICML.
Washington, H., Cook, J., 2011. Climate Change Denial: Heads in the Sand. Earthscan, New York, NY, pp. 10017.
Weber, E.U., Stern, P.C., 2011. Public understanding of climate change in the united states. Am. Psychol. 66 (4), 315.
Yao, L., Mimno, D., McCallum, A., 2009. Streaming Inference for Latent Dirichlet Allocation. KDD.
Young, N., 2013. Working the fringes: the role of letters to the editor in advancing non-standard media narratives about climate change. Publ. Understanding Sci. 22 (4), 443–459.